THE LARGE SYNOPTIC SURVEY TELESCOPE (LSST)

A PROPOSAL FOR RESEARCH AND DEVELOPMENT AND PRE-CONSTRUCTION ENGINEERING DESIGN

CONTENTS

R. Allsman[9], W. Althouse[16], S. Aronson[1], S. Asztalos[6], T. Axelrod[8], K. Baker[6], J. Barr[9], A. Becker[27], J. Becla[16], G. Bernstein[25], R. Blandford[16,17], G. Bowden[16], E. Bowell[7], W. Brandt[11], P.Burchat[17],D. Burke[16], C. Claver[9], A. Connolly[26], K. Cook[6], R. Crutcher[24], F. Delgado[9], D. Dossa[6], N. Felt[3], D. Figer[15], M. Foss[16], J. Geary[13]. P. Gee[21], K. Gilmore[16], W. Goldstein[6], Z. Haiman[2], L. Hale[6], A. Harris[14], S. Hawley[27], E. Hileman[9], C. Hogan[27], K. Honscheid[10], M. Huffer[16], Z. Ivezic[27], B. Jain[25], M. Jarvis[25], G. Jernigan[20], S. Kahn[16,17,**], J. Kantor[8], P. Kim[16], D. Kirkby[23], L. Knox[21], V. Krabbendam[9], D. Larson[11], E. Lee[16], M. Liang[9], R. Lupton[12], V. Margoniner[21], S. Marshall[16], P. Marshall[16], M. May[1], B. McLeod[3], D. Monet[18], J. Mould[9], D. Neill[9], M. Newcomer[25], S. Nikolaev[6], M. Nordby[16], P. O'Connor[1], J. Oliver[3], S. Olivier[6], K. Olsen[9], P. Osmer[10], M. Perl[16], J. Peterson[16], P. Pinto[19], R. Plante[24], T. Poehler[4], V. Radeka[1], A. Rasmussen[16], D. Rich[16], C. Roat[21], W. Rosing[5], A. Saha[9], J. Schaefer[8], T. Schalk[22], R. Schindler[16], R. Schmidt[9], D. Schneider[11], G. Schumacher[9], J. Sebag[9], L. Seppala[6], L. Simms[16], C. Smith[9], L. Stiles[19], M. Strauss[12], P. Strittmatter[19], C. Stubbs[3], D. Sweeney[8], A. Szalay[4], P. Takacs[1], J. Thaler[24], A. Tyson[21,*], R. Upton[9], R. Van Berg[25], S. Wang[1,2], L. Wang[20], B. Winer[10], D. Wittman[21], S. Wolff[9], H. Zhan[21]

[1] Brookhaven National Laboratory
[2]Columbia University
[3] Harvard University
[4] John Hopkins University
[5]Las Cumbres Observatory
[6]Lawrence Livermore National Laboratory
[7]Lowell Observatory
[8]LSST Corporation
[9]National Optical Astronomy Observatory
[10]Ohio State University
[11]Pennsylvania State University
[12]Princeton University
[13]Smithsonian Astrophysical Observatory
[14]Space Science Institute
[15]Space Telescope Science Institute
[16]Stanford Linear Accelerator Center
[17]Stanford University
[18]United State Naval Observatory
[19]University of Arizona
[20]University of California, Berkeley
[21]University of California, Davis
[22]University of California, Santa Cruz
[23]University of California, Irvine
[24]University of Illinois at Urbana-Champaign
[25]University of Pennsylvania
[26]University of Pittsburg
[27]University of Washington

* Project Director
** Deputy Project Director

Contents

1 Introduction ... 1
2 LSST as a Facility for Fundamental Physics Research ... 3
2.1 Introduction ... 3
2.2 Cosmic Shear ... 5
2.3 Surveys of Clusters of Galaxies ... 11
2.4 Baryon Acoustic Oscillations ... 13
2.5 Type 1a Supernovae ... 15
2.6 Gravitationally Lensed Supernovae ... 18
2.7 Photometric Redshifts ... 21
2.8 The Value of Multiple Probes of Dark Energy ... 23
3 Science Requirements Flowdown ... 25
3.1 Science Requirements for the LSST ... 25
3.1.1 Introduction ... 25
3.1.2 Constraining Dark Energy and Dark Matter ... 25
3.1.3 Taking an Inventory of the Solar System ... 27
3.1.4 Exploring the Transient Optical Sky ... 28
3.1.5 Mapping the Milky Way ... 29
3.1.6 Summary of Science Requirements ... 30
3.2 Requirements Traceability Matrix ... 37
3.3 Subsystems budgets ... 39
3.3.1 Introduction ... 39
3.3.2 Pointing and tracking ... 39
3.3.3 Image Quality ... 40
3.3.4 Ellipticity ... 42
3.3.5 Optical efficiency ... 43
4 Current Design ... 45
4.1 System Engineering and Integrated Design ... 45
4.1.1 Trade Studies ... 45
4.1.2 Optical Simulation ... 52
4.1.3 Cadence Simulations ... 57
4.1.4 End-to-End Simulations and Integrated Modeling ... 62
4.1.5 Optical Design ... 65
4.1.6 Observatory Controls ... 70
4.1.7 Calibration Strategy ... 73
4.2 Telescope/Site ... 75
4.2.1 Overview and Requirements ... 75
4.2.2 Facilities ... 79
4.2.3 Telescope Mount ... 89
4.2.4 Optical System ... 94
4.2.5 Alignment and Active Optics System ... 110
4.2.6 Telescope Control System ... 129
4.2.7 Observatory Telemetry System ... 133
4.2.8 Telescope Integration ... 136
4.3 Camera ... 140
4.3.1 Overall Description ... 140
4.3.2 Sensor Design ... 144

4.3.3 Camera Electronics ... 164
4.3.4 Packaging and Focal Plane Assembly ... 171
4.3.5 Mechanical Structures and Mechanisms ... 180
4.3.6 Thermal System ... 193
4.3.7 Corrector Optics and Color Filters ... 214
4.3.8 Controls and DAQ ... 220
4.3.9 Camera Integration and Testing ... 225
4.4 Data Management ... 231
4.4.1 Overall Description ... 231
4.4.2 Data Products ... 242
4.4.3 Application Layer ... 246
4.4.4 Middleware Layer ... 259
4.4.5 Infrastructure Layer ... 273
5 R&D Investigations ... 291
5.1 Simulations ... 291
5.2 Telescope R&D; Site Investigations ... 292
5.2.1 Site Evaluation and Selection ... 292
5.2.2 Mount Development ... 295
5.2.3 Telescope Optics ... 297
5.2.4 Wavefront Sensing and Alignment ... 298
5.3 R&D Investigations for the Camera ... 299
5.3.1 Focal Plane Development Plan ... 299
5.3.2 ASIC Development Program ... 303
5.3.3 Assembly and Alignment of the Focal Plane ... 305
5.3.4 Filter Development Program ... 307
5.3.5 Camera Integration and Testing ... 311
5.4 Data Management R&D ... 312
5.4.1 Summary of Research Activities and Primary Technical Risks ... 312
5.4.2 R&D Investigations ... 314
6 Project Organization and Management ... 339
6.1 Introduction ... 339
6.2 LSST Project Management ... 339
6.2.1 Management Structure ... 339
6.2.2 Integrated Project Management Control System (PMCS) ... 340
6.2.3 Configuration Management and Change Control ... 341
6.2.4 Contingency Management ... 342
6.3 Camera Project Management ... 342
6.3.1 General Camera Management Flowdown ... 342
6.3.2 Institutional Camera Management Flowdown ... 343
6.3.3 Work Breakdown Structure ... 343
7 Cost and Schedule ... 345
7.1.1 Cost and Schedule Analysis ... 345
7.1.2 Camera Development Schedule ... 346

List of Figures

Figure 2.1.1-1 Survey power is proportional to the étendue (the $A\Omega$ product) of the telescope aperture and camera field of view in m^2 deg^2. This plot compares various imaging survey projects, assuming 100% of nights are spent in survey mode. Only LSST, PS, and VST will operate 100% in survey mode. The LSST will open up a qualitatively new regime in survey science. A unique result of a very high optical étendue is that many science programs can proceed in parallel with the same high quality data. 5
Figure 2.1.2-1 An illustration of the geometry of weak gravitational lensing of a background galaxy by a concentration of mass along the line of sight. Both the displacement and distortion of the galaxy image are proportional to the enclosed mass within the lens. Thus weak lensing provides a means of mapping the dark matter distribution in the intervening medium. 6
Figure 2.1.2-2 An illustration of the effect of weak gravitational lensing due to cosmic shear on the appearance of background galaxies. 7
Figure 2.1.2-3 The lensing power spectra constructed from 5 redshift bins. Only the 5 auto-power spectra of each redshift bin among the available 15 spectra are displayed, and the solid curves show the predictions for the concordance ΛCDM model. The boxes show the expected one-sigma measurement error due to the sample variance and intrinsic ellipticities (the sample variance is dominant at about l<1000, while the intrinsic ellipticities are dominant at l>1000). In fact, a larger number of redshift bins will be enabled by LSST leading to a much larger number of auto- and cross-spectra that can be computed. 8
Figure 2.1.2-4 (*Left)* The relevant geometry for three-point correlations of weak lensing shear. The correlations are a function of three parameters (q, r, ψ) as shown in the figure. There are two components of shear at each vertex, yielding eight separate correlation functions. (*Right*) Plots of these eight functions for ΛCDM as a function of ψ for particular values of q and r. 9
Figure 2.1.2-5 The green, gray and blue contours show the constraints expected from the power spectrum tomography, the three-point correlation, or bispectrum tomography and the joint tomography of combining the two. It is clear that the bispectrum tomography improves parameter constraints by a factor of 2 compared to just power spectrum tomography, reflecting that the non-Gaussian signal in weak lensing provides additional cosmological information that cannot be extracted by the power spectrum. Shown here are two separate sets of constraints: The lower left is for ΛCDM, and the upper right is for a SUGRA model. 10
Figure 2.1.3-1 Plots of the number of real galaxy clusters per square degree that will be detected by LSST as peaks in the shear distribution and the number of false detections, as a function of the signal-to-noise threshold. 12
Figure 2.1.3-2: Constraints on dark energy equation of state parameters from the LSST sample of clusters of galaxies. 13
Figure 2.1.4-1: Simulations of the ratio of the measured galaxy power spectrum to a featureless reference power spectrum in various redshift bins, for the full LSST survey. 14
Figure 2.1.4-2: Error contours for dark energy equation of state parameters for LSST baryon acoustic oscillation measurements in different redshift bands, with and without the assumption of zero curvature. 15
Figure 2.1.5-1 Simulated lightcurves from the LSST deep survey for a Type 1a supernova at z = 0.832. The solid lines are the input light curves. Such detailed lightcurves, combined with knowledge derived from the exceedingly large number of nearby supernovae observed, will likely allow the determination of redshifts to better than 1% based on photometric data alone. The detailed multicolor lightcurves are also important for extinction corrections and for the control of systematics arising from variations intrinsic to the supernovae themselves. 16

Figure 2.1.5-2 Simulated Hubble diagram of a sample of 30,000 supernovae derived from LSST observations with redshifts determined photometrically. 17
Figure 2.1.5-3 Error contours in the equation of state parameter and the matter density derived from a representative LSST Type 1a supernova sample and from the baryon oscillation measurements. The magenta and blue contours correspond to optimistic and pessimistic priors on the accuracy of the photometric error distributions, respectively 18
Figure 2.1.6-1: Geometry of a strong gravitational lens 19
Figure 2.1.6-2: Redshift distributions of *measurable* lensed supernovae in the LSST survey area. The middle of the two lens-galaxy curves corresponds to the lensed core-collapse supernovae, while the lower curve corresponds to the quasar lenses. For this plot we assume consistent 0.7 arcsec seeing, a magnitude limit of 23 for each measurement, and a 10-year 20000 square degree survey, BUT with 50% visibility at any one time, and a completeness of 15% reflecting the fraction of data expected to be useful for time delay measurement. 20
Figure 2.1.6-3: Example light curves of four-image lensed supernova images for a system in a well-sampled field. The input time delays with respect to the first (black) curve are 32, 43 and 55 days. This plot shows why the peak magnitude required for a good time delay measurement (23.0) is brighter than the LSST limiting magnitude for an exposure pair (24.5): the SN has to be visible below peak, and in the fainter images too 21
Figure 2.1.7-1: LSST filter band transmission curves. 21
Figure 2.1.7-2 Photometric versus spectroscopic redshifts for grizY filter (*Left*) and for the adopted ugrizY filter set (*Right*) for the LSST survey after only 5 years. No brightness priors were used in this analysis. 22
Figure 2.1.8-1 Principle components of *w(z)*. The eigenvalues (*Left*) and the first three eigenmodes (*Right*) of the *w(z)* covariance matrix for LSST measurements of cosmic shear + Planck, and for a JDEM measurement of 2000 supernovae + Planck. 24
Figure 3.2.1-1: Principle sections and subsections of the traceability matrix showing the relationship between science requirements, engineering requirements and ultimately the specific implementation of the LSST. 37
Figure 3.2.1-2 A specific example of how the traceability matrix follows a science mission (Dark Energy/Dark Matter) through an observing program, techniques, science requirements to engineering requirements. 38
Figure 3.3.2-1 Example of trailing in the LSST FOV caused by a 10 arcsec error in one of the term of the pointing model (the OTA/el nonperpendicularity CA) 40
Figure 3.3.3-1 LBT 1 surface error (nm) over LSST clear aperture. Spherical aberration and 6 flexible bending modes have been subtracted. RMS surface error = 24 nm. 41
Figure 3.3.3-2 Image quality FWHM error budget allocations in arcsec 42
Figure 3.3.4-1 Atmospheric dispersion for the g and r filters 42
Figure 3.3.4-2 Ellipticity distribution for 0.8" zenith seeing in r 43
Figure 3.3.5-1 Single Exposure Duration versus Optical Efficiency to detect a 24th mag. Point source with S/N=10 44
Figure 3.3.5-2 Optical Efficiency (OE) Budget 44
Figure 4.1.1-1 Optical Design Matrix 46
Figure 4.1.1-2 Sky Availability 47
Figure 4.1.1-3 The dependence of zenith distance (ZD) on atmospheric dispersion induced ellipticity for each of the LSST filters griz Y in 0.4, 0.6, 0.8 and 1.0 arcsecond seeing (as measured at zenith). For each case three models were computed for altitudes of 1000m (upper dotted lines), 2000m (solid lines) and 3000m (lower dotted lines). 48
Figure 4.1.1-4 Successful image difference (right) showing no systematic effects between images with ADC tracking active (left) and ADC inactive by nulling the correction (middle). Note the asteroid moving between exposures as indicated by a positive-negative signature in the difference image. 49

Figure 4.1.1-5 Fabrication process for the two wedges of the conceptual ADC 50
Figure 4.1.1-6 Conceptual Mechanical Design of the LSST ADC 50
Figure 4.1.1-7 An illustration of the possible location of an ADC within the camera body. 51
Figure 4.1.1-8 Zenith dependence of differential atmospheric dispersion induced ellipticity in grzY spectral bands (see text)..... 52
Figure 4.1.2-1 Components of the LSST simulation chain. The Operations Simulator is not directly part of Optical simulation..... 53
Figure 4.1.2-2 Arroryo predicted PSFs at the LSST image plane for a realistic Cerro Tololo atmosphere. Each image is a sequential 0.1 second integration. (Asztalos)..... 55
Figure 4.1.2-3 Amplitude (left) and phases (right) at the LSST image plane from a point object. 55
Figure 4.1.2-4 Ray tracing the LSST optics. 56
Figure 4.1.2-5 Array of points ray traced through LSST optics. Without mirror perturbations (left), with mirror perturbations (right). 57
Figure 4.1.2-6 Hubble Deep Field N (left). HDF field propagated through atmosphere and ray traced (including mirror perturbations and wind shake). 57
Figure 4.1.3-1 Three programs coverage of the sky from Cerro Pachon using real CTIO seeing corrected to Cerro Pachon. 59
Figure 4.1.3-2 Weak lensing appropriate sky coverage for 339 days of CTIO weather from Cerro Pacho with SN/NEA..... 60
Figure 4.1.3-3 Transient fields with completed sequences observed (NEA and SN programs) for 339 days CTIO weather from Cerro Pachon with WL/SN/NEA programs running. 61
Figure 4.1.3-4 Distribution of visits per field in five filters for simulations using a 3.0 degree FOV (left panel) and 3.5 degree FOV (right panel). For multiple science programs, the 3.0 FOV barely meets (does not completely cover the accessible area) the WL visit requirement, while the 3.5 degree FOV covers the available area with visits to spare. 62
Figure 4.1.4-1: The function block diagram for the LSST end-to-end simulator..... 64
Figure 4.1.5-1: LSST baseline optical design..... 65
Figure 4.1.5-2 The polychromatic image sizes are less than 0.2 arcsec across the field for 50% energy collection. Except for the g-band, the polychromatic image sizes are less than 0.24 arcsec across the field for 80% energy collection. Filter bands: u: 320nm – 400nm g: 400nm-560nm, r: 540 nm-720 nm, i: 685 nm-870 nm, y: 840 nm-950 nm, z: 945 nm-1030 nm. 68
Figure 4.1.5-3 LSST optical layout *(top)* and camera corrector optical layout *(bottom)* 69
Figure 4.1.6-1 The following diagram gives an overview of the LCS system architecture. 71
Figure 4.1.7-1 On the right side, sky emission spectrum with plot of grizY filters. Top: total sky emission flux across 50A bins. Bottom: the sky emission spectra sampled into 5A bins (from UVES instrument, VLT). On the left side, schematic view of physical processes for atmospheric and instrumental effects..... 74
Figure 4.1.7-2 Calibration process schematic. 75
Figure 4.2.1-1 Rendering of the LSST Telescope and Summit Facilities. 76
Figure 4.2.1-2 Candidate LSST site locations..... 78
Figure 4.2.2-1 Mirror & Top End Handling as driver of dome size and shape..... 83
Figure 4.2.2-2 Summit Support Facility Conceptual Plan..... 85
Figure 4.2.2-3 Potential Site for LSST at San Pedro Mártir Observatory..... 87
Figure 4.2.2-4 Potential site for LSST at Cerro Pachón..... 88
Figure 4.2.2-5 Potential site for LSST at Las Campanas 89
Figure 4.2.3-1 Front View of Telescope Assembly..... 90
Figure 4.2.3-2 Auxiliary Telescope Components..... 90
Figure 4.2.3-3 Mount Vibration Characteristics..... 91
Figure 4.2.3-4 Section view of top end assembly..... 92
Figure 4.2.3-5 Top end assembly spider arrangement..... 92

Figure 4.2.3-6 Instrument assembly 93
Figure 4.2.3-7 Secondary mirror assembly 94
Figure 4.2.4-1 Telescope Optical Design 95
Figure 4.2.4-2 Telescope Ray Trace..... 95
Figure 4.2.4-3 Primary/tertiary monolithic mirror optical surfaces 96
Figure 4.2.4-4 Primary/tertiary monolithic mirror design 97
Figure 4.2.4-5 Secondary Mirror Baseline Design, 100 mm Thick Solid Meniscus..... 98
Figure 4.2.4-6 Secondary Mirror, Active Support Locations 98
Figure 4.2.4-7 Secondary Mirror Design, Structured Meniscus Option..... 99
Figure 4.2.4-8 The swing arm profilometer has been used to measure giant secondary mirrors, such as the 1.7-m mirror shown here. The accuracy of this instrument is demonstrated by comparing data from the profilometer with an interferometric measurement..... 100
Figure 4.2.4-9 Demonstration of software for combining multiple swingarm scans to create a surface map. The figure shows (*left*) a nominal surface with simulated errors, (*center*) a simulation of a measurement of the surface using a combination of 6 swingarm scans and 2 circumferential scans, with 0.1 um rms simulated measurement noise added at each point, and (*right*) the computed residual error..... 100
Figure 4.2.4-10 Demonstration of 4-m profilometer, measuring a 1.8-m mirror 101
Figure 4.2.4-11 Sub aperture Interferometric test layout 102
Figure 4.2.4-12 Layout for demonstration of vibration insensitive Fizeau test..... 102
Figure 4.2.4-13 Interferogram and reduced surface data provided by the vibration insensitive Fizeau interferometer. 103
Figure 4.2.4-14 Rigid body motions and sign conventions 104
Figure 4.2.4-15 Field points used in sensitivity analysis. All points were equally weighted..... 105
Figure 4.2.4-16 80% EE Spot Diameter versus decenter 106
Figure 4.2.4-17 80% EE Spot Diameter versus defocus 107
Figure 4.2.4-18 LSST model in FRED..... 108
Figure 4.2.4-19 Stray light path (blue rays) and imaging path (red rays) 108
Figure 4.2.4-20 Annular ring baffles to block undesirable specular ray paths..... 109
Figure 4.2.4-21 Reflectivity curves after three reflections 110
Figure 4.2.5-1 Coordinate system and sign convention used for defining the degrees of freedom for rigid body displacements in the LSST alignment analysis. 112
Figure 4.2.5-2 Optical footprint on each of the three mirrors for field points on axis (blue) and at full field of view at 1.75 degrees off axis (green). 113
Figure 4.2.5-3 Two examples showing the field dependencies of astigmatism caused by a higher order aberration, trefoil (left) and tetrafoil (right), on the tertiary. The blue sticks represent the orientation and strength of astigmatism in the focal plane as a function of field position. 114
Figure 4.2.5-4 An example decomposition of astigmatism in the focal plane to component sources using the Zernike representation. 115
Figure 4.2.5-5 Functional diagram of the LSST Active Optics System 116
Figure 4.2.5-6: The initial merit function histogram of RMS image size across the field of view from the Monte Carlo simulations of the perturbed optical system. The optical system was perturbed in all degrees of freedom within the tolerances in Table 4.2.5-1. 117
Figure 4.2.5-7: The cumulative distribution (in %) of all the models versus the averaged value of the merit function after a first iteration of correction through the reconstructor. In this case, 50% of all the Monte Carlo models reach an averaged merit function of 4.37 microns. 117
Figure 4.2.5-8: The limiting merit function (mean RMS image size over full FOV) as a function of wavefront sample points. The noise was applied to each Zernike term at each field point using a uniform distribution having the width indicated..... 119
Figure 4.2.5-9 Z (Hi) process flow for initial alignment wavefront curvature data. 120
Figure 4.2.5-10 Phase retrieval simulations with weak and strong turbulence. 121

Figure 4.2.5-11 Illustration denoting hypothetical Hartmann spots from a plane wave, dashed lines, and a severely aberrated beam, solid line. 122
Figure 4.2.5-12 Principle of operation of each lenslet. The angular displacement of the spot for each individual lenslet indicates the average slope vector of the wave-front across the given lenslet..... 122
Figure 4.2.5-13 Preliminary design of a Shack-Hartmann wavefront sensor. It is composed of nine optical elements with a numerical aperture (NA) of 0.4. It reimages the pupil of the telescope onto a lenslet array. 123
Figure 4.2.5-14 Wavefront sensor configuration with 3x3 Shack-Hartmann sensor. The total field available is 112", having each sensor with a field of view of 37.5"..... 123
Figure 4.2.5-15 Number of photons per pixel of the Shack-Hartmann wavefront sensor (assumption of a quadcell per spot) versus star magnitude for a 15sec exposure 124
Figure 4.2.5-16 Shack-Hartmann simulation images of the applied and reconstructed phase. The image on the left shows the spot distribution on the Shack-Hartmann sensor. 124
Figure 4.2.5-17 Curvature sensing consist of taking the difference between the illuminations observed in 2 planes Pin (intra-focal) and Pex (extra-focal) separated from the focal plane by the same distance d..... 125
Figure 4.2.5-18 Concept of a CWFS where two detectors, located next to each other and in between the science detectors of the focal plane array (FPA), are placed in an extra-focal position (Pex) and in an intra-focal position (Pin) with a total separation of 2d..... 125
Figure 4.2.5-19 Example of intra and extra-focal images of a single star. The defocus distance is 500 microns. The reconstructed phase is displayed on the right. 126
Figure 4.2.5-20 Effect of defocus distance on the variance of the reconstructed phase for the iterative technique 126
Figure 4.2.5-21 Pupil overlap simulations. The intra and extra-focal images on the left show two stars separated by 30 pixels, with an intensity ratio of 40%, and with a defocus distance of 500 microns. On the right side are displayed the input phase and the reconstructed phase for comparison. 127
Figure 4.2.5-22 LSST Model in Spatial Analyzer..... 128
Figure 4.2.5-23 Laser tracker measurement modeling in Spatial Analyzer software..... 129
Figure 4.2.6-1 Control Architecture. 130
Figure 4.2.6-2 Overview of the TCS dataflow 131
Figure 4.2.7-1 Observatory Telemetry System Architecture 135
Figure 4.2.8-1 Example of Laser tracker head and retro reflectors under consideration for integration and operational alignment 137
Figure 4.2.8-2 Factory assembly and testing of SOAR telescope mount (structure and control system). A similar approach will be used for LSST..... 138
Figure 4.2.8-3 LSST summary integration Gantt chart for critical tasks. 139
Figure 4.2.8-4 Optical configuration showing two positions for prime focus cameras to test M1 and M3 and the commissioning camera in place of the LSST Camera to test M2 and the three optic system 139
Figure 4.2.8-5 LSST optical design delivers spherical wavefront at camera entrance on axis. .. 140
Figure 4.3.1-1 Optical layout of the LSST telescope (left) and camera elements (right)..... 141
Figure 4.3.1-2 Cross-sectional views of camera showing major camera components. In the upper illustration, the filter carousel and changing mechanism have been omitted for clarity; the lower illustration has been rotated 90° to show two views of the shutter mechanism and its relationship with the filter change mechanism..... 142
Figure 4.3.1-3 Cutaway view of the camera housing and components 143
Figure 4.3.2-1 Internal quantum efficiency of silicon as a function of wavelength, for thicknesses of 50, 75, 100, 150, and 250 μm..... 146

Figure 4.3.2-2 Contour plot showing the dependence of QE on temperature and silicon thickness, for 1000 nm wavelength. The blue bar shows the expected operating temperature range for the LSST camera. 147
Figure 4.3.2-3 Temperature coefficient of quantum efficiency as a function of wavelength. Thickness = 100 μm. At long wavelengths the temperature coefficient approaches 1% per degree C. 147
Figure 4.3.2-4 Diagram illustrating two detector-related contributions to the point spread function. (a) Photogenerated charge experiences diffusion as it drifts to the collecting electrodes, broadening the PSF approximately as the square root of the drift distance. (b) At longer wavelengths the optical beam penetrates deep into the silicon and forms a conical volume over which charge is generated. The conical half-angle is arctan(1/2nf), where n is the index of refraction of silicon and f is the focal ratio of incident light. 148
Figure 4.3.2-5 Contour plots showing the dependence of PSF on sensor thickness, operating voltage, resistivity, and conductivity type. Solid black lines are contours of constant g-band diffusion (FWHM shown). Temperature 173 K. Red lines mark the boundary of full depletion. The dashed blue lines show the contours of maximum electric field in kV/cm. Note that the high resistivity n-type material allows small PSF to be obtained at the lowest E-field. 150
Figure 4.3.2-6 Effect of displacement of the focal plane: (a) position of best focus for short-wavelength light; (b) focal plane displaced 10 μm in direction of incoming rays. Refraction causes position of focal point to move about 5 times farther than sensor displacement. 151
Figure 4.3.2-7 Absorption of long-wavelength light in a 100 μm-thick sensor. Focal plane position shown displaced by 0, and ±10 μm from best short-wavelength focus. Spot diagrams (with no diffusion) shown on the left. 152
Figure 4.3.2-8 PSF as a function of thickness. Effects of diffusion and beam divergence have been included. The focal plane position is varied at each wavelength until the best PSF is achieved. Sensor temperature 173 K; average electric field 2 kV/cm. 153
Figure 4.3.2-9 Point spread function dependence on focal plane position. Displacement of 0 μm corresponds to the point where light rays come to focus at the silicon surface. Positive displacements move the focal point into silicon volume. Sensor thickness = 100 μm; average electric field 2 kV/cm; operating temperature 173 K. 154
Figure 4.3.2-10 Limiting magnitude (S/N=10) for 10-second LSST exposures, for the filter set shown in Table 2. Sensor operating temperature 173K; substrate resistivity 10kΩ-cm, n-type; average electric field 2000 V/cm; focal plane position adjusted for best focus in each band. Reflection losses at sensor and atmospheric extinction not included. (a) Atmospheric seeing not included; (b) including seeing of 0.7" FWHM. 156
Figure 4.3.2-11 Outline of the 16 Mpixel strawman CCD, showing the partitioning and charge movement for the hardwired split parallel and serial registers. All pinout for the device is along the left and right edges. The fill factor achieved in this design iteration is 96.5%. 159
Figure 4.3.2-12 Serial registers for two of the interior imaging areas, with a 90 degree bend introduced in order to provide more separation between output FETs. 160
Figure 4.3.2-13 A simplified sketch of a hybrid Si PIN-CMOS senor. (Large pixel pitch, as used in IR sensors are indicated). 162
Figure 4.3.3-1 – Layout of camera electronics 164
Figure 4.3.3-2 – Readout architecture for single sensor 167
Figure 4.3.3-3 – Readout Architecture for groups of sensor modules. The designations Zone 1 and Zone 2 refer to thermal zones discussed in Section 4.3.3.2.7. 168
Figure 4.3.3-4 – Test stand architecture 171
Figure 4.3.4-1 Proposed sensor package showing a CCD on aluminum nitride carrier with electrical and mechanical interfaces. 172
Figure 4.3.4-2 Backside of CCD package 173
Figure 4.3.4-3 Sensor packages assembled into “raft” structure. 173

Figure 4.3.4-4 Cross-section through focal plane showing one packaged CCD on raft assembled to FPA integrating structure. Kinematic mounts at the raft-FPA integrating structure interface may include adjustments for raft alignment. 174
Figure 4.3.4-5 An example of a possible arrangement of 3×3 sensor rafts in the focal plane with areas provided for wavefront sensors by raft offset. 174
Figure 4.3.4-6 Raft module with integrated front-end electronics and thermal connections.[need to add label for front end elevtronics] 175
Figure 4.3.4-7 Focal plane assembly concept..... 176
Figure 4.3.4-8 Allocation of focal plane flatness specification. Of the total camera allocation of 10 μm peak-to-valley (p-v),.half is reserved for the sensor module fabrication, and the remainder allocated to the combination of fabrication, thermal and mechanical load errors in the raft substrates and focal plane integrating structure, and assembly tolerances. 177
Figure 4.3.4-9 The finite element model (upper left) is constrained at three points, each with tangential and axial displacement constraints about a cylindrical coordinate system (i.e., kinematic constraint). The color fringe plot (upper right) shows displacement of the structure due to gravity along the camera axis (Z). Deviation from flatness (or sag) is derived from this displacement information. The two (lower) animation plots show mode shapes that the image stabilization system would tend to excite. These modal frequencies should be well above the desired system bandwidth for robust control..... 179
Figure 4.3.4-10: Example of a planar-motion flexure mechanism. Three folded-hinge flexures allow planar motion and each provide one out-of-plane constraint at the fold line. Together with three actuators, coordinated X-Y-θ_Z motion is possible. 180
Figure 4.3.5-1: Concept-level design of the LSST camera. 181
Figure 4.3.5-2 Finite element model (left) and plot of maximum principal stress (right). Positive stress is tensile in the normal sign convention used here. The plot looks unsymmetrical because the upper compressive stress shows up in a plot of minimum principal stress, which is the mirror image of this one. 183
Figure 4.3.5-3 Proposed design for the shutter, incorporating both the one-sheet and the two-sheet alternatives. All dimensions in mm. The side frame has been removed for full view of the rollers..... 185
Figure 4.3.5-4: Time vs. position of a constant-width aperture through a closed-open-closed sequence. The curves correspond to the leading and trailing edges of the aperture. The next sequence would occur in the opposite direction..... 186
Figure 4.3.5-5: Time vs. position of a variable-width aperture through a close-open-close sequence. The curves correspond to the leading and trailing edges of the aperture. The next sequence would occur in the opposite direction..... 186
Figure 4.3.5-6 Mechanical tension mechanism for the one-sheet shutter. A motor (not shown) drives both rollers with equal torque while the preload force causes equal and opposite torque to tension the sheet. The geometry of this mechanism keeps these two functions orthogonal to one another..... 187
Figure 4.3.5-7: Possible edge stiffening beams formed on the sheet. Further though is needed on whether a formed beam is sufficiently sharp for the aperture. Instead, a separate formed beam could be spot welded to the sheet leaving a sharp edge. 188
Figure 4.3.5-8: Two-sheet design using two pinch rollers to tension each sheet on its main roller. 188
Figure 4.3.6-1 Schematic of the Camera Thermal and Vacuum System..... 195
Figure 4.3.6-2 Camera Thermal / Vacuum System Zone Definitions..... 196
Figure 4.3.6-3 FPA Assembly — top & bottom..... 196
Figure 4.3.6-4 FPA Assembly with Break-out..... 197
Figure 4.3.6-5 Raft Module with Break-outs 197

Figure 4.3.6-6 Detector Module (DDC Module 2B) Thermal Analysis Using A Single Central Strap 204
Figure 4.3.6-7 Example of Integrating Structure Thermal Distortion Analysis 205
Figure 4.3.6-8 Thermal/Convection Analysis of Camera (a) Thermal profile in L1, L2, L3, Filter, and Dry N_2 gas (b) Convective flow pattern in camera. 206
Figure 4.3.6-9 Thermal/Convection Analysis of Camera - L1 Termperature and Distortion. (a) Temperature profile of L1. (b) Thermal distortion pattern in camera. 207
Figure 4.3.6-10 Cryostat Cut-a-way Indicating Vacuum Zone Barriers 208
Figure 4.3.6-11 Raft Assembly Cut-a-way indicating Vacuum Zone Barriers 209
Figure 4.3.6-12 Conceptual Vacuum System Implementation 209
Figure 4.3.6-13 Chiller System 211
Figure 4.3.7-1: The optical design defines the size and placement of the optical elements and ray bundles define stay-out zones for support hardware. This concept-level design of the LSST camera shows three corrector optics, storage of five color filters, the focal plane array and other camera hardware. 215
Figure 4.3.7-2 Schematic layout of current LSST camera optics and detector array. 216
Figure 4.3.7-3 Dimensioned drawing of an early version LSST camera optics and detector array layout. 217
Figure 4.3.7-4 One example of a detailed mechanical drawings that were produced for early versions of each of the LSST camera optics. 217
Figure 4.3.7-5 Simple null test configurations for the LSST corrector lenses using a common 1.7 m diameter spherical mirror with a 4 m radius of curvature. 218
Figure 4.3.7-6 Photo of a large (2 m class) conformal window produced by commercial U.S. vendors. 219
Figure 4.3.8-1 Logical view of Camera Control System modules and communications. The CCM is the master control module. Each of the control modules labeled CM_j is responsible for managing a single camera subsystem. The message routing system transports control and status data messages between the control modules. 221
Figure 4.3.8-2 Model architecture for camera control. The top row shows Ethernet attached components, (C_j's), with control modules (CMc's) running on the same host as the CCM (right). The middle row shows control modules (CMa's) running on processors embedded in the subsystems labeled SSa_j. The bottom row illustrates subsystems (SSb's) on a common bus with control modules (CMb's) running on the bus master. 222
Figure 4.4.1-1 LSST DMS Operational Concept 232
Figure 4.4.1-2 The LSST Data Management Layered Architecture 238
Figure 4.4.1-3 The Application Layer 239
Figure 4.4.1-4 The Middleware Layer 239
Figure 4.4.1-5 The Infrastructure Layer 240
Figure 4.4.2-1 LSST Data Products Domain Model 242
Figure 4.4.3-1 Application layer view of the DMS pipelines 247
Figure 4.4.3-2 Use case diagram for IPP 249
Figure 4.4.3-3 Robustness for "determine WCS" 251
Figure 4.4.3-4 Robustness for "determine PSF" 252
Figure 4.4.3-5 Rapid Mover Associate Pipeline 257
Figure 4.4.4-1 Overview of Middleware Layer 259
Figure 4.4.4-2 The Middleware Component Diagram 260
Figure 4.4.5-1 LSST DM infrastructure with tradeoff points 276
Figure 4.4.5-2 Compute-Intensive Architecture 278
Figure 4.4.5-3 Network/Storage-intensive Architecture 279
Figure 4.4.5-4 LSST Communications Links by Site 280

Figure 4.4.5-5 IBM Multiprocessor module with 4 dual-core Power 5 CPUs and 4 L3 cache chips. 285
Figure 5.2.1-1 Goes-8 Satelite image at 11:45UT on October 25, 2000. Left, infra-red window channel (10.7micron) and right, water vapor channel (6.7 micron). LSST study evaluates 25 pixels above each site to establish level of cloud cover. 293
Figure 5.2.1-2 Image from CTIO All-Sky camera system called TASCA (Chile). The buildings in the background are the observatories located near the camera. The red point represents the pointing of the main telescope..... 294
Figure 5.2.1-3 Example of a Differential Image Motion Monitor (DIMM)..... 295
Figure 5.3.1-1 Measurement of as-built CCD flatness in focal plane of Hires imager mounted on the Keck II telescope. 302
Figure 5.3.4-1 Illustration of generic filter characteristics 308
Figure 5.3.4-2 Photo of a large (3 m class) coating chamber operated by a commercial U.S. vendor (left) along with a coated NOVA laser optic of similar size to the LSST optics (right). 308
Figure 5.4.2-1 Illustration of a 3-D concept for handling queries. 325
Figure 5.4.2-2 Structure of the LSST pipeline simulator 334
Figure 6.2.1-1 LSST Management Structure..... 340
Figure 6.3.3-1: Camera Organization Chart 344
Figure 7.1.2-1 LSST Project Timeline 347
Figure 7.1.2-2: LSST Camera Development Schedule 348

List of Tables

Table 3.1.6-1 LSST Science Requirements Summary Table 31
Table 4.1.1-1 Necessary Zenith Angle to Observe to Reach Levels of Sky coverage versus Latitude Position of the Observatory and levels of Sky Exclusion Surrounding Galactic Plane. . 47
Table 4.1.5-1 LSST baseline optical prescription for r band filter: all units are mm except as noted 67
Table 4.2.1-1 Telescope and Site Requirements 76
Table 4.2.1-2 LSST Sites (in order of north latitude to south latitude) 78
Table 4.2.2-1 Space requirements for lower enclosure and attached building 80
Table 4.2.2-2 Space requirements for summit support facility (detached)..... 80
Table 4.2.2-3 Space requirements for summit dorm/dining facility..... 81
Table 4.2.2-4 Space requirements for base support facility 81
Table 4.2.4-1 Parameters for LSST Optical System 94
Table 4.2.4-2 Actuator error forces measured on the baseline LBT actuators. Each value is the rms of the actuator error forces. The error force for an actuator is the maximum measure on that actuator. 97
Table 4.2.4-3 Summary of the error budget for LSST interferometric surface test. 103
Table 4.2.4-4 Sensitivity Coefficients 106
Table 4.2.4-5 Ranges for valid quadratic approximation 106
Table 4.2.5-1 Assembly tolerances used in simulation of initial alignment..... 117
Table 4.2.5-2: Estimated capture range requirement for maintaining operational alignment. 118
Table 4.2.8-1 Commercial Laser Tracker Measurement Performance..... 137
Table 4.3.2-1 Science requirements driving sensor design 144
Table 4.3.2-2 Requirements table (from "THE LARGE SYNOPTIC SURVEY TELESCOPE DESIGN AND DEVELOPMENT PROPOSAL" submitted to the National Science Foundation by the LSST Corporation, December, 2003)..... 145
Table 4.3.2-3 Allowable detector contribution to point spread function..... 148
Table 4.3.2-4 Sensor features 158
Table 4.3.3-1 Distribution of camera electronics functionality 165
Table 4.3.3-2 Focal plane array and readout specifications 166
Table 4.3.3-3 Preliminary ASIC specifications..... 169
Table 4.3.4-1 Key results from the finite element analysis using steel and scaled results for SC-30 and Invar-36. 179
Table 4.3.6-1 Cryostat Heat Extraction Summary. 198
Table 4.3.6-2 - FPA Thermal Loads and Heat Flow Estimate. 199
Table 4.3.6-3 – BEE Assembly Thermal Loads and Heat Flow Estimate 200
Table 4.3.6-4 – Camera Enclosure Thermal Loads and Heat Flow Estimate..... 200
Table 4.3.6-5 Requirements For Temperature Control Within The Cryostat..... 202
Table 4.3.6-6 Requirements For Temperature Control: Integrating Structure, Camera Body, and Camera enclosure excluding Cryostat 203
Table 4.3.6-7 Requirements For Temperature/Pressure Control: Optical Elements 207
Table 4.3.6-8 Requirements for Vacuum & Contamination Control 210
Table 4.3.7-1 Parameter details for the three corrector lenses and (i band) filter. 216
Table 4.3.7-2 Baseline LSST filter band-pass FWHM points..... 220
Table 4.4.5-1 LSST Communications Links Options 281

Table 4.4.5-2 Current supercomputer performance..... 286
Table 4.4.5-3 Projected cost performance trends in computing 286
Table 4.4.5-4 Projected cost of 1 Petabyte of disk storage..... 287
Table 4.4.5-5 Cost Comparison of Media Shipping versus Network..... 289
Table 5.3.1-1 CCD sensor development schedule..... 300
Table 5.3.2-1 Set of milestones for the signal processing ASIC development 304
Table 5.3.4-1 Detector response curve (QE) 310
Table 5.4.1-1 Summary of Data Management Research Activities and Primary Technical Risks 312
Table 7.1.1-1 Resources Required for Preconstruction R&D and Engineering 345

Executive Summary

The discovery that 95% of the universe is made from mysterious “dark energy” and “dark matter” has generated tremendous excitement in the scientific community, and piqued strong interest in the public at large. Most experts agree that these discoveries have taken us to the brink of a revolution in the standard views of cosmology and particle physics. The Large Synoptic Survey Telescope (LSST) will address these fundamental questions with observations and measurements that are sensitive to the nature of space and time, the evolution of energy and matter, and the relation between these kinematical and dynamical properties of our universe. The LSST will provide precise characterizations of dark energy and dark matter through studies of multiple phenomena in a single deep survey of one half the astronomical sky. Specific measurements will include tomographic ($z \leq 3$) projections of correlations in spatial shear observed in weak lensing of galaxies, baryon acoustic oscillations, and cluster distributions and power spectra of spatial distributions of galaxies. These quantities depend on cosmological distances and distributions of energy and matter in differing ways. Combined with studies of "standard candles" (supernovae at $z \leq 1$) and "synchronized clocks" (time delays of multiply-lensed supernovae) detected by LSST, these measurements will over-constrain and determine fundamental cosmological parameters. For example, six eigenmodes of the dark energy equation of state will be determined by LSST data alone, and when combined with measurements (i.e. from Planck) of anisotropies in the cosmic microwave background ($z \cong 1100$), the lowest eigenvalues will be determined with percent-level systematic accuracy.

The LSST will enable scientific programs that would require centuries to complete with existing telescopes. The total LSST system optical throughput (étendue ≡ aperture × field of view = 320 $m^2 deg^2$) will be two orders of magnitude greater than any existing facility, and will allow a survey of the sky of unprecedented depth and width. The LSST mission is a multi-pass survey with 2000 observations of each 10 square degree patch of the sky spanning six photometric bands (0.3 – 1.0 μm) to magnitudes 26.5-27 AB. In each year of the survey 250,000 Type Ia supernovae ($z \leq 1$) will be detected and prompt alerts issued to the international observing community. Surface brightness shapes of over 3 billion galaxies ($z \leq 3$) will be measured in the course of the ten-year mission. Precise determination of the PSF across each image, accurate photometric calibration, and continuous monitoring of system performance and observing conditions will lead to unprecedented control of systematic errors.

With its 8.4m primary aperture, the LSST will join the present generation of telescopes with “8-meter class” mirrors – the Large Binocular Telescope (LBT), Gemini, Subaru, and the Giant Magellan Telescope (GMT) that is presently being constructed. The unique LSST 3-mirror optical design, combined with a large (65cm diameter) focal plane, produces an extraordinary field of view (3.5° FOV). The large aperture and wide FOV are optimum for a survey instrument like LSST; together they define much of the technical challenge in the telescope and camera. The resulting short focal length (a “fast” f/1.2 focal ratio) and correspondingly short depth of focus (10 μm confocal length) challenge the imaging quality. A contract to fabricate the monolithic LSST primary-tertiary mirror is in place with the Steward Mirror Laboratory, and fabrication of the LSST mirror is to be integrated with production of the seven 8.4m GMT mirrors. The telescope active optics, wave-front sensing, guiding, and observational monitoring systems are modern technologies with counterparts working in the field today. The planned LSST cadence of observation will require careful engineering to achieve fast telescope slew rates and settle times, and special attention to reliability and maintainability are needed to maximize time-on-the-sky.

The focal-plane detector of the LSST camera will be a 3.2 Giga-pixel mosaic with 10μm pixels in 4K × 4K sensor packages (CCD and PiN-CMOS options are being developed). The camera

will require 6400 channels of electronics to achieve fast read-out (2 sec) of each 15 sec exposure. Sensor R&D is underway at Brookhaven National Laboratory (BNL) in collaboration with potential vendors. The short confocal length of the optics and the required image quality create tight requirements on flatness and thermal stability of the camera focal plane array; these are subjects of active R&D at BNL and SLAC. The LSST optical photometric bands (u,g,r,i,z,Y) will be similar to those used in the Sloan Digital Sky Survey (though SDSS does not use Y). The LSST filters will be the largest ever used on a telescope, so potential vendors are involved in early study of design and fabrication techniques. Final assembly of the full camera system, and complete operational tests and calibrations, will be done at SLAC before transport to the observatory site for mounting on the telescope.

The LSST will acquire nearly 2000 images and produce ~120 Tbytes of raw and preprocessed image and catalog data per full night of observing. Images will be acquired every 15 seconds, image analysis for stringent quality control and detected transient alerts are to be generated within 60 seconds, and data must be archived and accessible for decades. This dynamic range poses challenges to the design of the LSST data acquisition and management systems similar to those encountered in applications in aerospace, intelligence, and high-energy physics. The LSST data management design will use a layered architecture (infrastructure, middleware, and applications layers) to segregate tasks and allow forward compatibility with future developments in hardware and software technologies. Software and hardware specification and design will use modern Iconix process with Unified Modeling Language diagrams and strict coding standards. The envisioned networking and computing structure will allow data management to be distributed between facilities on the observatory mountain, at a local base facility, and at archive and data centers in the United States. Data products that will be created automatically as images are acquired (level one products) include raw images, calibrated science images with instrument signatures removed, source and object catalogs, and alerts to a variety of classes of transient events. Higher level data products, driven by off-line science analyses, will also require extensive computing support. The LSST Collaboration is committed to providing open access by the scientific community to all data products with no proprietary period and no delay in publication except as needed to commission the facility and assure the quality of the data.

Construction and operation of the LSST is proposed to be a joint initiative of the U.S. Department of Energy, the National Science Foundation, and the privately funded LSST Corporation, Tucson, AZ – a non-profit entity. An experienced project central management team is headquartered at LSST Corporation, with leading members at the major institutions in the project. The Stanford Linear Accelerator Center (SLAC) is proposed as the lead laboratory for the DOE contribution to the project.

Major milestones being proposed for the project include recommendation of a site for the facility in FY2006, R&D and engineering design through FY2008, and start of construction in FY2009. With timely approvals and appropriate funding, completion of construction and “first light” can occur in FY2012 with commissioning of hardware and software continuing through 2013. The LSST can be providing data products to the scientific community early in the next decade.

1 Introduction

- "The Committee supports the Large Synoptic Survey Telescope project, which has significant promise for shedding light on the dark energy." *Connecting Quarks with the Cosmos*

- "The SSE [Solar System Exploration] Survey recommends [the construction of] a survey facility, such as the Large-Aperture Synoptic Survey Telescope (LSST)…to determine the contents and nature of the Kuiper Belt to provide scientific context for the targeting of spacecraft missions to explore this new region of the solar system…" *New Frontiers in the Solar System*

- "The Large-aperture Synoptic Survey Telescope (LSST) will catalog 90 percent of the near-Earth objects larger than 300-m and assess the threat they pose to life on Earth. It will find some 10,000 primitive objects in the Kuiper Belt, which contains a fossil record of the formation of the solar system. It will also contribute to the study of the structure of the universe by observing thousands of supernovae, both nearby and at large redshift, and by measuring the distribution of dark matter through gravitational lensing." *Astronomy and Astrophysics in the New Millennium*

These studies, conducted by the National Research Council to recommend research priorities for the coming decade, have all endorsed the construction of a wide-field telescope, the LSST, that will survey the entire visible sky every few days to extremely faint limiting magnitudes. Advances in microelectronics, large optics fabrication, and computer hardware and software now make it possible to build a system that will address a broad range of problems in cosmology, astrophysics and solar system exploration in qualitatively and quantitatively new ways.

The LSST system will obtain sequential images of the entire observable sky every few nights. These images can be co-added to provide unprecedented depth and area coverage. The same images can also be subtracted from each other to highlight celestial sources that change in brightness, position, or both. Repeat imaging on a variety of timescales from 15 seconds to years will open a new "time window" on the universe. A distinguishing feature of the experimental design is that multiple science programs can be carried out in parallel; a common set of images will address a wide diversity of science goals. Indeed, all the science objectives are addressed by a single LSST survey database. The LSST facility will enable programs that would take a century on current facilities. The data will be reduced in real time and the resulting images, database, search tools, and software will be made publicly available.

Because of its unprecedented capabilities and its promise for discovery at the frontiers of astronomy and physics, the LSST has brought together scientists and engineers from many universities, Department of Energy laboratories, the National Optical Astronomy Observatory (NOAO), and the private sector. Together, this group has devised a system concept that will meet the requirements of the three decadal surveys: an 8.4-m telescope, a camera system with a 10 square-degree field of view, and a suite of image-processing pipelines that will produce and provide access to images in real time. There are engineering challenges in fields ranging from device physics to data mining. The 3.2 billion pixel camera will be the world's largest imager. The acquisition, real-time processing, cataloging, and accessing of data at the extraordinary rates that will be realized by the LSST (anticipated to be ~ 20 Terabytes per night) will catalyze significant developments in computational science and engineering.

At its last meeting in 2003, we presented to the SLAC EPAC a *Letter of Intent* proposing SLAC involvement in the research and development phase of the LSST project. As envisioned at that time, a team supported by DOE-OS, led by SLAC with significant involvement at BNL, LLNL, and university-based HEP groups, would take overall responsibility for the LSST camera, the data acquisition system, and aspects of the pipeline software and simulations systems. The telescope itself, the enclosure, the site, and other components of the software and operations, were to be developed with funding from the NSF and from private sources. The LSST Corporation, whose members at that time included the Research Corporation, the Universities of Arizona and Washington, and NOAO, was to assume overall management responsibility for the LSST collaboration, leading to the production of a system design that is mature enough to allow the project to proceed to pre-construction review.

In the ensuing two years, most elements of the plan we presented have indeed been realized. The LSST collaboration has grown to over 100 members, with significant representation from both astrophysics and high energy physics groups at both national laboratories and at universities. There are now fourteen institutional members of the LSST Corporation, and new applications for membership are coming in monthly. SLAC has indeed taken leadership for the development of the LSST camera: Both the Lead Scientist (Kahn) and the Project Manager (Gilmore) for the camera are SLAC affiliates, as is the Systems Engineer (Althouse) for the overall LSST project. Kahn has also been appointed Deputy Project Director for LSST as a whole. The technical maturity of the design of the telescope, camera, and data management system has also increased substantially. We are presently engaged in the development of a project execution plan which should lead to a detailed cost estimate of the project through construction. We expect to have that completed by the Fall of 2006.

In September 2005, the LSST Corporation was awarded a $14.2M contract from the Astronomy Division of the NSF for LSST design & development. That funding is primarily being utilized to support the work being carried out on the telescope and data management subsystems. Here we submit a formal proposal for DOE-OS support of the research & development and the engineering design of the LSST camera. This is for pre-construction funding.

The DOE Office of High Energy Physics recently announced CD-0 approval for the development of a "ground-based dark energy experiment". The LSST is a natural candidate for that opportunity, and we seek the EPAC's endorsement of our proposal for consideration by the DOE and its various advisory committees.

The following sections of this proposal provide the scientific motivation for DOE-OHEP involvement in LSST; the requirements placed on the telescope, instrument, and software by the scientific goals; the technical descriptions of the current designs of the telescope, camera, and data management subsystems; a plan for outstanding research & development activities associated with these subsystems; a project management plan; and an overview of the anticipated costs and schedule.

A summary of the key parameters for the LSST survey is given in the table below.

6-band Survey: *ugrizy* 320–1050 nm

- Sky area covered: 20,000 deg^2 0.2 arcsec / pixel
- Each 9.6 sq.deg FOV revisited >300 times/band
- Time resolution: >20 sec
- Limiting magnitude: 26.5 AB magnitude @10σ (24.5 in u)
  24 AB mag in 15 seconds
- Photometry precision: 0.01 mag requirement, 0.005 mag goal
- Galaxy density: 50 galaxies/sq.arcmin
- 3 billion galaxies with color redshifts
- Time domain: Log sampling, seconds – years

2 LSST as a Facility for Fundamental Physics Research

2.1.1 Introduction

In the last decade, cosmologists have converged on a "standard model'' of cosmology, the Concordance Cosmological Model, which successfully unifies a wide array of diverse observational constraints, but requires the existence of two rather exotic new forms of matter and energy. In particular, the matter density of the Universe appears to be dominated by some form of non-baryonic dark matter, while the energy density is dominated by a new form of vacuum energy field with negative pressure, which has been dubbed "dark energy." Within the context of this model, 96 percent of the mass-energy of the Universe is not only "dark", but cannot be easily accommodated by the Standard Model of Particle Physics.

Explaining these phenomena has emerged as among the greatest challenges to modern theoretical physics. Non-baryonic dark matter implies the existence of a totally new sector of particles. There is widespread belief that this may be a manifestation of supersymmetry, but direct evidence for such a connection is lacking. The evidence for dark energy poses an even greater problem. While dark energy is compatible with and may be related to Einstein's famous cosmological constant, Λ, the inferred value of Λ, is many orders of magnitude below the "natural" values expected from quantum field theory. In particular, the mass scale implied by Λ is in the milli-electron-Volt range, a regime where we were not expecting to find evidence for "new physics".

While there is reason to be optimistic that hints to the solutions to these puzzles will come from future accelerator-based experiments like the Large Hadron Collider and the International Linear Collider, it is also clear that more refined cosmological measurements will be extremely important. Current data provide an accurate value for the average mass density of the Universe in dark matter, but still provide only very weak constraints on how dark matter behaves dynamically, and how it interacts with itself or with baryonic matter. There is marginal evidence that the clumpiness observed in the dark matter haloes of galaxies is less pronounced than would be expected for purely gravitational interactions, but this still needs to be confirmed and quantified with much higher precision.

As for dark energy, current data merely constrain the existence of the effect but little else. Of particular interest is an understanding of the dynamical behavior of dark energy, i.e. how it behaves with cosmic time or with redshift. It has become common to characterize this evolution in terms of an "equation of state parameter": $w \equiv p/\rho$, where p is the pressure, and ρ is the energy density. For a pure cosmological constant, $w = -1$, and is constant in time. If the dark energy is associated with a new scalar field, one might expect w to be a function of the scale factor of the Universe, usually represented by a, where $a = 1$ in the current epoch ($z = 0$), and $a = 0$ at the time of the Big Bang. A simple parametrization is then to characterize the evolution of the dark energy in terms of w_0, the value of the equation of state parameter now, and w_a, its derivative with respect to the scale factor a. Current data only constrain w_0 to be consistent with -1 to within 10 – 20% (depending on assumptions), and place no meaningful constraints on w_a.

The deep, very wide-field, multi-color imaging survey of the sky that LSST will produce will be a "goldmine" for these kinds of cosmological investigations, in that it will enable a large

number of distinct and complementary analyses that independently constrain dark matter and dark energy models with much higher precision than has been possible before. Key measurements that LSST will make include:

- The two- and three-point power spectra of cosmic shear as a function of redshift in both the linear and non-linear regimes
- The number density and power spectrum of clusters of galaxies as a function of redshift
- A detection of baryon acoustic oscillations in the galaxy power spectrum as a function of redshift
- The Hubble diagram of a large number of well-sampled light curves of Type 1a supernovae out to $z = 1.2$
- The discovery of a sizeable number of gravitationally lensed supernovae

These various techniques probe the Concordance Cosmological Model in a multitude of ways, using "standard candles", "standard rulers", and "standard clocks" as a function of redshift to constrain the expansion history of the universe, and by measuring the growth of structure with cosmic time to constrain the dynamical history of gravitational interactions on a wide range of spatial scales. In the subsections below, we provide short overviews of each of these techniques, and present estimates of the quantitative constraints that will come from the LSST survey. However, we want to emphasize that this should not be interpreted as a complete list of possible cosmological science investigations with LSST: This field is still in its infancy, and new ideas for possible analyses of the LSST database are being discovered all the time. It is quite likely that what will emerge as the most interesting and most constraining investigations have not yet even been envisioned.

For all of the analyses highlighted above, the relevant figure-of-merit by which to evaluate an imaging survey is what astronomers and optical designers call the étendue of the system: $A\Omega$ - the product of the effective collecting area of the telescope and the field of view sampled by the camera. The étendue is inversely proportional to the time it takes to sample a given solid angle of sky down to a given depth, or equivalently the depth that can be reached for a patch of sky in a given time. The product $A\Omega T$, where T is the total observing time devoted to the survey, is analogous to the integrated luminosity of accelerator-based experiments. For high precision measurements like cosmic shear and baryon acoustic oscillations, the statistical error bars are propotional to the square root of $A\Omega T$. For the study of rare events like lensed supernovae, the number of expected detections is linearly proportional to $A\Omega T$.

Figure 2.1.1-1 below shows a comparison of the étendue of LSST to that of an array of survey instruments on existing, planned, or in some cases, proposed facilities. With the exception of Pan-STARRS (PS1 and PS4) and the VLT Survey Telescope (VST), these are not on dedicated telescopes, so the effective étendue is really significantly lower than what is plotted – it should be reduced by the fraction of telescope time allotted to the survey (typically only 10 – 30%). In any case, it is easily seen that LSST represents a dramatic increase in capability for this field. It represents a two-order-of-magnitude improvement in étendue over existing surveys, and will have at least a factor five higher $A\Omega T$ product than any other experiment which has even been proposed. In fact, with the exception of LSST, none of these facilities propose to undertake the single homogeneous deep wide-area survey that is essential to address the problem of dark energy. PS1, PS4, and VST are planning sequential, very different types of surveys.

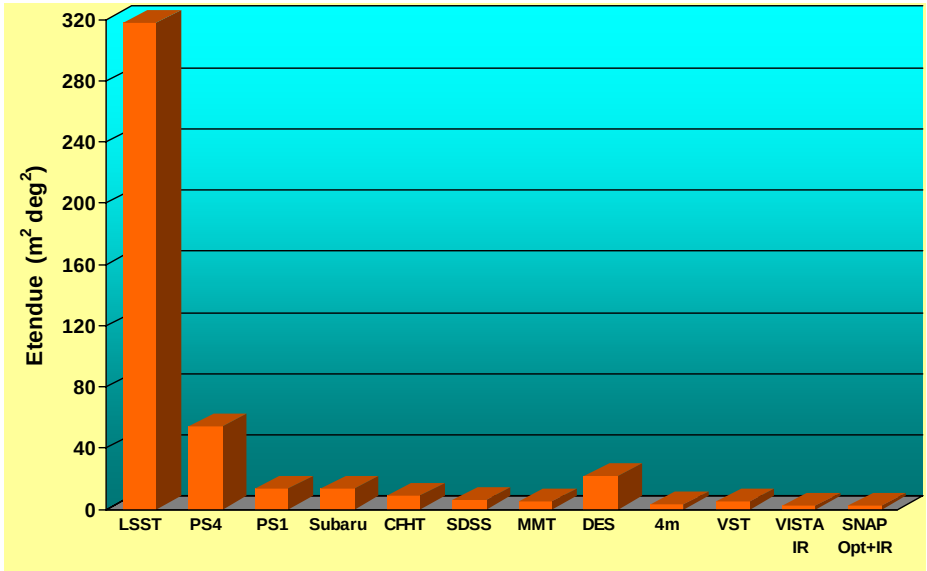


Figure 2.1.1-1 Survey power is proportional to the étendue (the $A\Omega$ product) of the telescope aperture and camera field of view in m^2 deg^2. This plot compares various imaging survey projects, assuming 100% of nights are spent in survey mode. Only LSST, PS, and VST will operate 100% in survey mode. The LSST will open up a qualitatively new regime in survey science. A unique result of a very high optical étendue is that many science programs can proceed in parallel with the same high quality data.

2.1.2 Cosmic Shear

The term "cosmic shear" refers to the systematic and correlated distortion of the appearance of background galaxies due to weak gravitational lensing by the clustering of dark matter in the intervening universe. The basic principles are illustrated in Figure 2.1.2-1 and Figure 2.1.2-2. As light rays from a background source pass through the gravitational potential of an intervening dark matter clump, they get subtly deflected. For a point source of light, this merely produces a displacement of the apparent position of the object. For an extended source like a galaxy, the image of the source also gets distorted or sheared. The degree of shear is roughly given by the angular displacement, which is equal to the ratio of the distance from the source to the lens, D_{LS}, to the distance from the source to the observer, D_S, times a gravitational potential factor: $4GM/bc^2$, where M is the enclosed mass of the lens, b is the impact parameter that the light rays make relative to the lens center, and c is the speed of light. Thus the measured ellipticity or shear of a background galaxy provides information about the column density of dark matter along the line of sight and the geometry of the Universe.

Of course individual galaxies can have elliptical shapes, even with no shear due to lensing. Therefore, the measurement of this effect for a single galaxy is not easily interpretable. However, the alignments of background galaxies (especially at different redshifts) are completely random and uncorrelated. Therefore, one can distinguish the component of shear due to lensing by measuring correlations in the shear as a function of angular scale. For an excess of mass at some point in space, one will tend to see the major axes of background galaxies aligned circumferentially around the center of the mass distribution. For a local deficit of mass relative to the average density, one will tend to see the major axes of background galaxies aligned radially outward from the mass minimum.

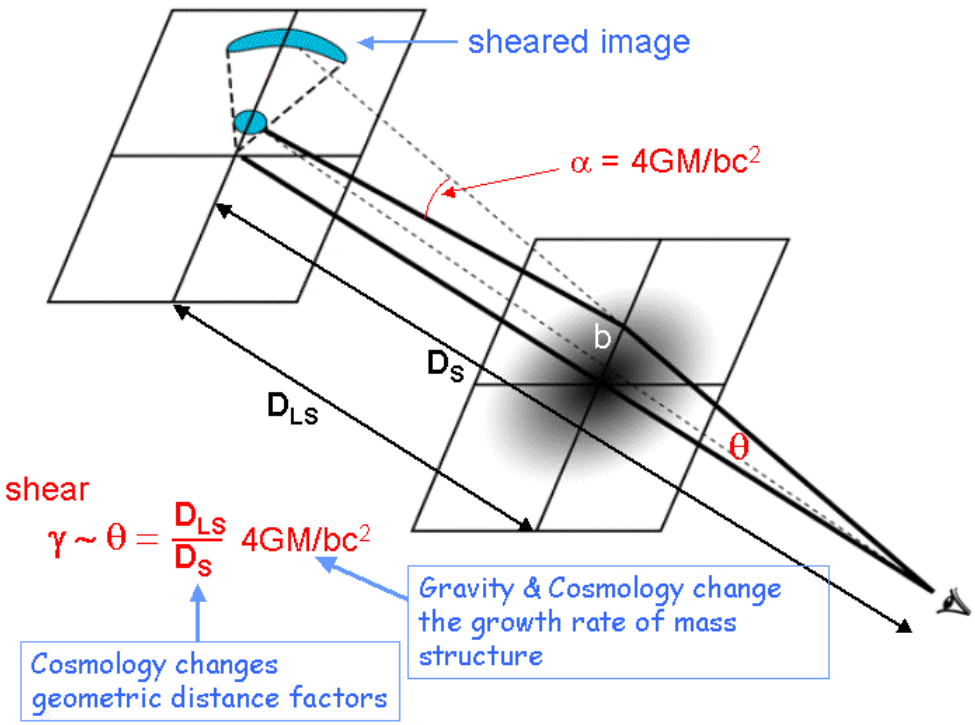


Figure 2.1.2-1 An illustration of the geometry of weak gravitational lensing of a background galaxy by a concentration of mass along the line of sight. Both the displacement and distortion of the galaxy image are proportional to the enclosed mass within the lens. Thus weak lensing provides a means of mapping the dark matter distribution in the intervening medium.

As the Universe expands, concentrations of dark matter grow with time due to gravitational accretion. Smaller concentrations grow first and then gradually coalesce to form larger and larger structures. This "growth of structure" produces predictable statistical patterns in the distribution of dark matter as a function of redshift, Those predictions can be tested by measuring the statistical properties of the weak lensing shear of background galaxies. That concept is illustrated in Figure 2.1.2-2. The random concentrations of dark matter are represented by the spidery gray structures in the cartoon. The statistical properties of those structures are imprinted in correlations of the shapes and orientations of background galaxies on the observed sky. This statistical weak lensing effect is usually referred to as "cosmic shear".

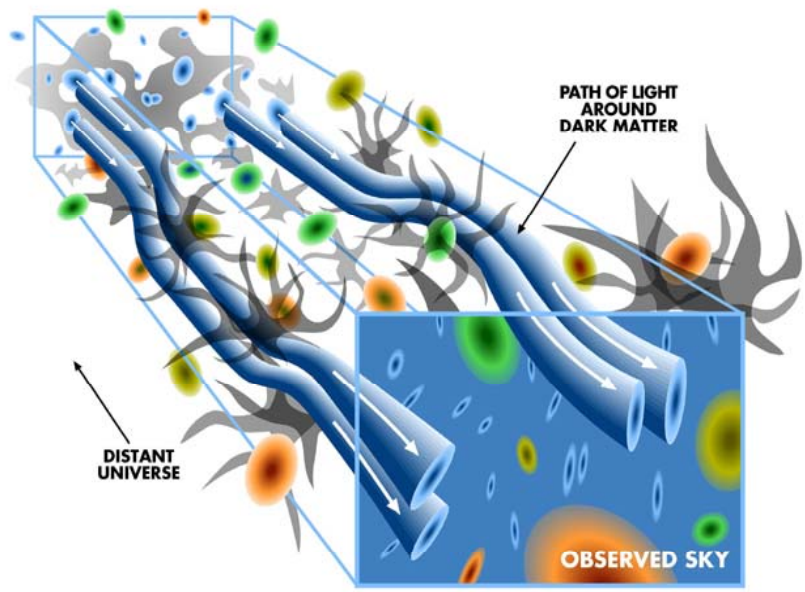


Figure 2.1.2-2 An illustration of the effect of weak gravitational lensing due to cosmic shear on the appearance of background galaxies.

Cosmic shear is sensitive to the underlying cosmology in a variety of ways. First, as indicated above, the lensing signal depends on geometric factors – ratios of distances whose scaling with redshift in turn depends on the expansion history of the Universe. Second, the amplitude of the lensing signal is related to the growth of structure. This is a kind of cosmic clock. If the expansion of the Universe is too fast, larger structures cannot form at the early times implied by higher redshifts. To exploit the full power of these data, it is crucial to measure the cosmic shear as a function of redshift, a technique known as "power spectrum tomography". Here the background galaxies are separated into redshift bins using the ratios of their intensities in different spectral bands (photometric redshifts, or photo-z's – see Section 2.1.7 below). One then measures the statistical properties of the lensing for these individual redshift slices. The result allows us to isolate structures at low redshift from those at higher redshift along the line of sight.

The statistic that has been most often discussed in the context of cosmic shear is the "lensing power spectrum". This is the Fourier transform of the two-point shear correlation function, i.e. the average product of the shear of two different galaxies separated by some angle θ, and calculated as a function of θ. Because the sky is a sphere, the lensing power spectrum is usually calculated and plotted in a spherical harmonic basis, similar to what is done for the fluctuations in the cosmic microwave background (CMB). Predicted lensing power spectra for the LSST database are plotted in Figure 2.1.2-3. These are given in five redshift bins. As the redshift of the background galaxies increases, the lensing signal is stronger because there are more intervening dark matter concentrations. The dependence on redshift of the amplitudes of these curves provides a measure of the growth of structure with cosmic time, as discussed above.

Note that these curves are not simple straight lines – there is information in their shapes as well as their amplitudes. Of particular significance is the point of inflection which occurs at $l \sim$ few hundred (corresponding to $\theta \sim 0.5 - 1$ degree). This represents the transition from the "linear" to the "non-linear" regime. In the linear regime (larger angular scales, i.e. lower l), the structures

are still small perturbations on the overall background mass density. The growth of structure in the linear regime can be calculated analytically using techniques that are as good and as robust as those used to predict the fluctuations in the CMB. The interpretation of the shear power spectrum in this regime in terms of constraints on the underlying cosmology is thus especially straightforward. In the non-linear regime, the structures have grown into significant perturbations on the background density, and the calculation of their subsequent evolution requires N-body calculations to model the gravitational interactions of the dark matter particles. In principle, the data are more sensitive to the details of the underlying cosmology in this regime, but possible uncertainties in our understanding of dark matter interactions and/or the effects of baryons can affect the interpretation of the measurement. Given this complementarity, it is crucial to measure lensing power spectra in both regimes. *Only a survey experiment with an étendue as large as that of LSST will have the power to measure lensing power spectra across this transition!* Present measurements of cosmic shear have been limited to only the smallest angular scales (few arc-minutes, $l \sim$ several thousand), where the effects are strongest.

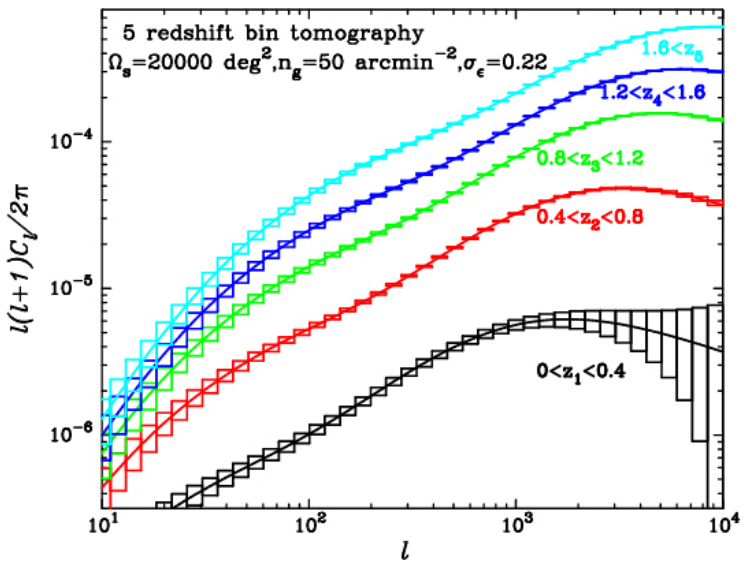


Figure 2.1.2-3 The lensing power spectra constructed from 5 redshift bins. Only the 5 auto-power spectra of each redshift bin among the available 15 spectra are displayed, and the solid curves show the predictions for the concordance ΛCDM model. The boxes show the expected one-sigma measurement error due to the sample variance and intrinsic ellipticities (the sample variance is dominant at about l<1000, while the intrinsic ellipticities are dominant at l>1000). In fact, a larger number of redshift bins will be enabled by LSST leading to a much larger number of auto- and cross-spectra that can be computed.

The curves plotted in Figure 2.1.2-3 are "auto-power spectra". They represent shear correlations between galaxies in the same redshift bin. One can also compute "cross-power spectra" between different redshift bins. These contain additional information. Further, one can cross correlate with the CMB and with the density of foreground galaxies. Such multiple probes yield tight constraints on the expansion history of the Universe and thus on the nature of dark energy.

It is also possible to measure higher-moment correlation functions of the shear field. The shear three-point function is an independent measurement from the two-point function and thus adds to the total signal-to-noise obtainable from weak lensing data. Furthermore, the constraints on cosmological parameters are along somewhat different degeneracies than the two-point function, so the combination of the two statistics is significantly more powerful than either one individually. Finally, the three-point function can probe aspects of the shear field, such as non-Gaussianity, which the two-point functions cannot.

The three-point shear correlation function, however, is distinctly more complicated for weak lensing studies. First, the geometry of triangles dictates that the three-point function is a function of three parameters: for example (q, r, ψ) in the diagram shown in Figure 2.1.2-4. Second, the shear at each vertex has 2 components. Thus, the full three-point function has 8 combinations of these, leading to 8 separate correlation functions. These can be divided into parity-odd and parity-even functions corresponding to whether they change sign under the transformation $\psi \rightarrow 2\pi - \psi$. The predicted values of the 8 functions for ΛCDM cosmology are plotted in Figure 2.1.2-4 as a function of ψ for two values of (r,q). The symbols + and x refer to the two shear components relative to the center of the triangle.

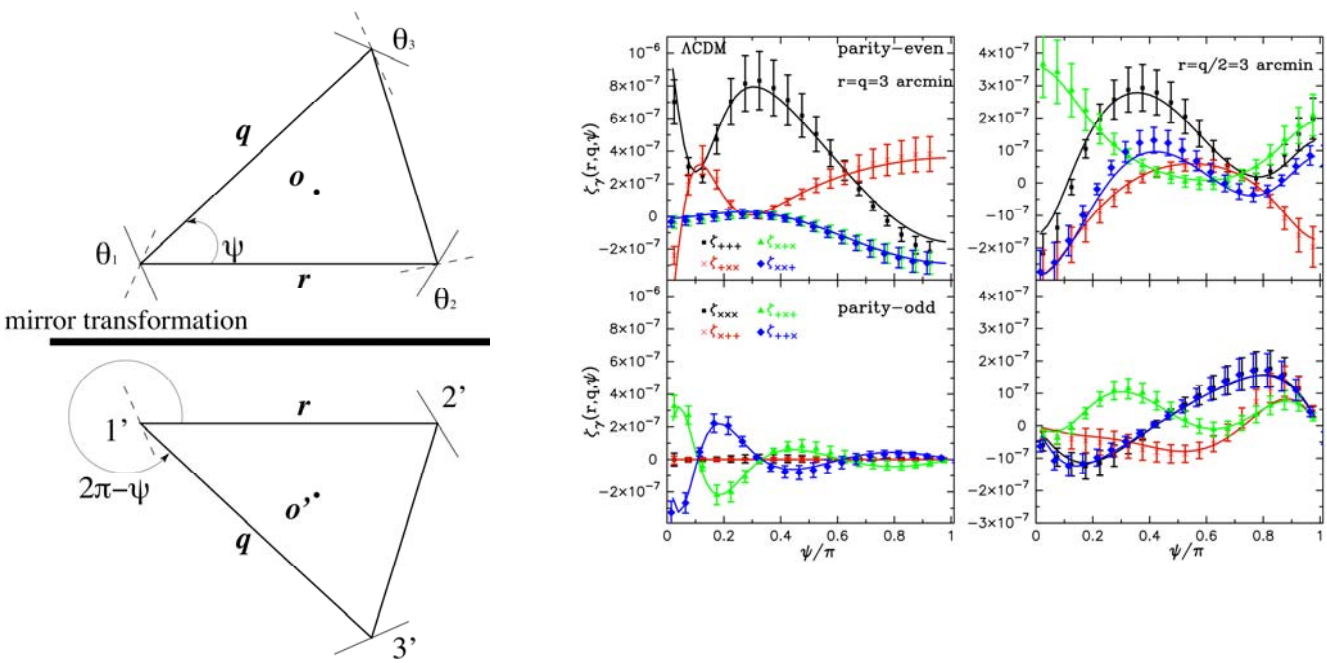


Figure 2.1.2-4 (*Left)* The relevant geometry for three-point correlations of weak lensing shear. The correlations are a function of three parameters (q, r, ψ) as shown in the figure. There are two components of shear at each vertex, yielding eight separate correlation functions. (*Right*) Plots of these eight functions for ΛCDM as a function of ψ for particular values of q and r.

In Figure 2.1.2-5, we show the constraints on cosmological parameters that will be derivable from LSST's lensing data. We show the 68% confidence limit contours for the two-point and

three-point correlation functions separately (green and gray, respectively), as well as the combined constraints using both measurements (blue). These constraints take advantage of the LSST's ability to measure photometric redshifts for the lensed galaxies. Not only does this improve the calibration of the source population compared to what is possible with current surveys, but, as discussed above, it also allows us to perform separate auto- and cross-correlations of galaxies in different redshift slices. (We use 5 redshift slices here.) The cross-correlations improve the signal-to-noise in general by a factor of two or so, and they improve the constraints on w_a by a factor of about 10.

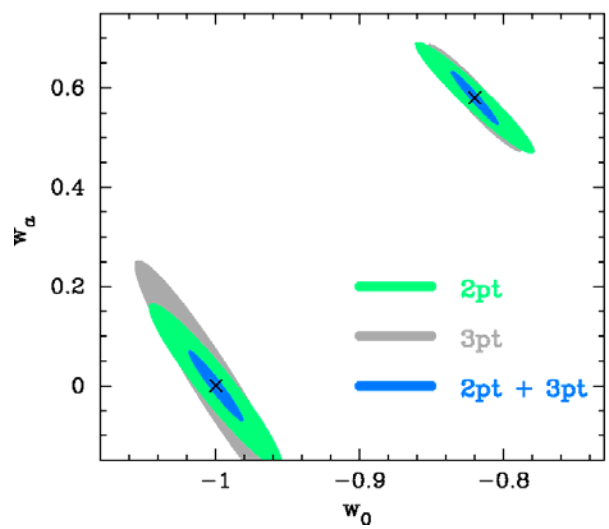


Figure 2.1.2-5 The green, gray and blue contours show the constraints expected from the power spectrum tomography, the three-point correlation, or bispectrum tomography and the joint tomography of combining the two. It is clear that the bispectrum tomography improves parameter constraints by a factor of 2 compared to just power spectrum tomography, reflecting that the non-Gaussian signal in weak lensing provides additional cosmological information that cannot be extracted by the power spectrum. Shown here are two separate sets of constraints: The lower left is for ΛCDM, and the upper right is for a SUGRA model.

Cosmic shear measurements with imaging survey data require great attention to systematics since the weak lensing signal (particularly at large angular separations) is at a very low level. The LSST has been designed from the start with this measurement in mind. Particular attention has been paid to the control of the point spread function (PSF) of the telescope and the ability to monitor effects within the telescope and camera which could lead to PSF variations across the field. However, of central importance is the fact that LSST will enable a large number (~ 300 per color) of separate exposures to be assembled for each region of sky. Spurious shear correlations introduced by the atmosphere or by environmental conditions around the telescope will cancel out when the correlations are performed between galaxies in different exposures. This provides a crucial check on our analysis. The ability to assemble many exposures per field is a direct consequence of the large aperture of the the telescope and the very fast readout speed of the camera (2 s for the entire 3 Gpixel array), features that are not available on any other existing or planned survey experiment.

2.1.3 Surveys of Clusters of Galaxies

Clusters of galaxies are the most massive, gravitationally bound structures in the Universe, and thus an especially good probe of the growth of structure. Their number density ($dN/dzd\Omega$) and their spatial power spectrum ($P(k)$ – the Fourier transform of the cluster-cluster correlation function) are exponentially sensitive to the cosmological expansion. LSST has the potential to identify over a hundred thousand galaxy clusters. This cluster sample will have statistically well-controlled mass estimates, and can place precise and robust constraints on cosmological parameters. By combining measurements of $dN/dzd\Omega$ and $P(k)$, degeneracies among cosmological parameters, and also between cosmological parameters and systematic errors in the analysis, can be broken, yielding percent-level constraints on individual parameters. The LSST cluster constraints are complementary to those from LSST measurements of cosmic shear, and from supernova studies. The only caveat with this approach is that since clusters form in a highly nonlinear regime, these constraints rely on the validity of N-body simulations of cluster properties.

LSST will detect clusters as peaks in the shear distribution resulting from weak lensing measurements. In using clusters for cosmology, it is essential that the observable we use to find them provide an accurate estimate of the total mass. Via weak lensing, we are looking directly at the dark matter, which dominates the mass of the cluster. Other methods of finding clusters (X-ray surveys, measurements of the Sunyaev-Zeldovitch effect) rely on baryon tracers of the mass, which are biased and may be less reliable. The mass-shear relation for shear peaks can be accurately calibrated from simulations for any assumed underlying cosmology.

Galaxy clusters will be selected as a set of peaks in a smoothed two-dimensional shear map. Using a filter with particular angular scale, we identify peaks above a threshold corresponding to a multiple of the noise. The correspondence between peaks and clusters is imperfect due to (a) missing a fraction of the real clusters, and (b) false detection of overdense structures, due to projected lower mass structures along the ling of sight. These effects are quantified by the fraction of real clusters detected, and the purity of the sample, i.e. the fraction of peaks that correspond to real clusters.

Using the N-body simulations for an underlying cosmological model, we can determine these two statistics. Some sample results are shown in Figure 2.1.3-1, in which we have plotted the number of real clusters detected and the number of false detections as a function of noise threshold for a particular spatial filter, corresponding to one arc-minute. As can be seen, with a signal-to-noise threshold of five, the sample is ~ 75% pure, and the contamination can be reliably estimated.

Our simulations show that LSST will detect ~ 200,000 clusters. The resulting constraints on dark energy parameters are illustrated in Figure 2.1.3-2. We find that w_0 can be constrained to better than 4% and w_a can be constrained to ~ 0.1 from the cluster sample alone. $dN/dzd\Omega$ contains most of the information on w_a, while $P(k)$ substantially improves the constraints on w_0. Adding CMB constraints from Planck results in relatively modest improvements.

The cluster statistics will also yield tight constraints on the sum of the neutrino masses. The Universe is filled with a relic background of neutrinos, thermally produced in the Big Bang. Massive neutrinos behave differently from massless ones in two cosmologically important ways. First, because they have mass, their energy density dilutes with redshift less rapidly. This extra energy density means a Universe with massive neutrinos expands more rapidly than one with massless neutrinos (assuming the same number of species in each). The increased expansion is a drag on the growth of structure, since more rapid expansion makes it harder for matter to cluster.

The second effect of mass is that the neutrino free-streaming length decreases with increasing mass. On length scales larger than the neutrino free-streaming length, neutrinos can collapse into

gravitational potential wells. This added contribution to gravitational instability cancels out the drag from the increased expansion, meaning there is no net effect on large scales. On scales below the neutrino free-streaming length, the neutrinos cannot cluster, and thus the increased expansion rate does not suppress the matter power spectrum. The amplitude of the suppression is proportional to the sum of the neutrino masses, whereas the free-streaming length is inversely proportional to the individual masses. The sum can be determined from the former, whereas the individual masses (in principle) can be determined from the latter.

Current astrophysical limits on the sum of the neutrino masses are ~ 1 eV. The LSST cluster sample can determine this sum with an error of 0.02 eV. This is a very interesting mass range, since the atmospheric neutrino oscillations require that at least one of the active neutrinos have a mass in the range 0.04 – 0.1 eV. More detailed considerations show that the sum of the active neutrino masses should be at least 0.06 eV.

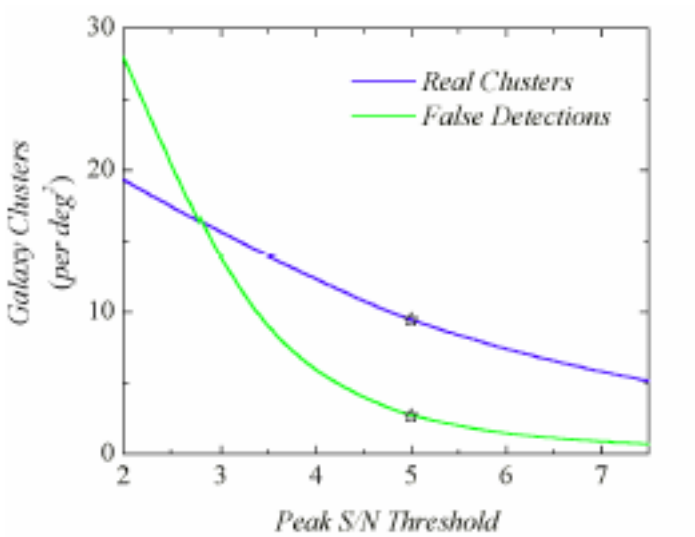


Figure 2.1.3-1 Plots of the number of real galaxy clusters per square degree that will be detected by LSST as peaks in the shear distribution and the number of false detections, as a function of the signal-to-noise threshold.

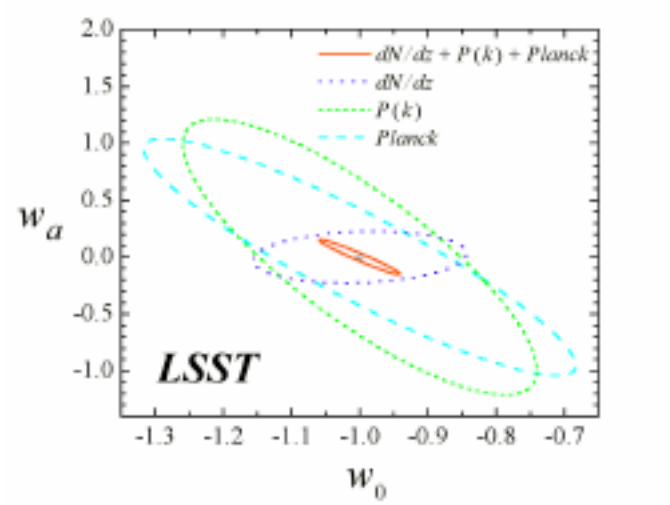


Figure 2.1.3-2: Constraints on dark energy equation of state parameters from the LSST sample of clusters of galaxies.

2.1.4 Baryon Acoustic Oscillations

In the tightly coupled photon-plasma-dark matter fluid prior to recombination, acoustic waves, supported by the photon pressure, create a characteristic scale – the sound horizon R_S in the matter distribution. After recombination, the sound speed of the neutral gas practically drops to zero, and thus the imprint of R_S at recombination becomes frozen in the matter, and subsequently, the galaxy distributions. These same acoustic waves give rise to the peaks seen in the CMB temperature fluctuation spectrum, allowing R_S to be accurately determined (its value is ~ 150 Mpc). Due to the dark matter dominance, the signature of R_S is a set of peaks in the galaxy power spectrum (the Fourier transform of the galaxy-galaxy correlation function), which are known as "baryon acoustic oscillations" (BAOs). Since the BAOs are associated with a well-defined commoving distance, but are measured as a function of angle, they can be used to provide a "standard ruler" for determining the angular diameter distance as a function of redshift, and thereby constrain the cosmic expansion.

The LSST survey will provide a sample of over three billion galaxies, far larger than has ever been previously assembled. Using photo-z's, these can be collected into redshift bins, and the galaxy power spectrum can be computed as a function of redshift. Simulations of the quality of the data we expect for this effect are given in Figure 2.1.4-1. As can be seen, the BAO peaks are cleanly detected for redshifts ranging from $z \sim 0.2 - 2.7$. The implied angular diameter distance can be measured to ~ 0.4% accuracy, especially at higher redshift.

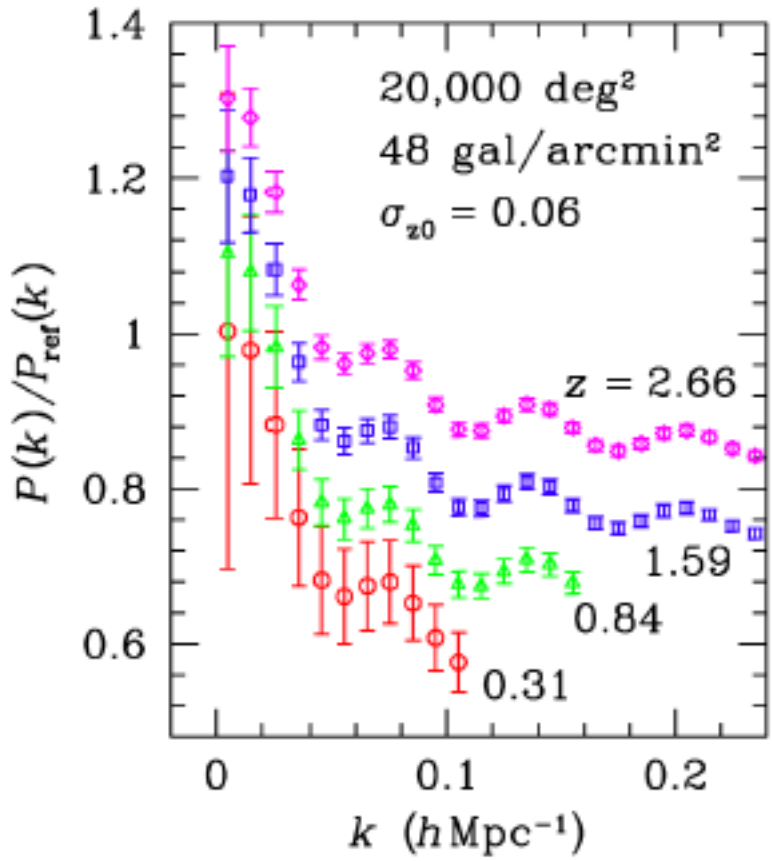


Figure 2.1.4-1: Simulations of the ratio of the measured galaxy power spectrum to a featureless reference power spectrum in various redshift bins, for the full LSST survey.

If the *P(k)* distribution can be measured in the longitudinal direction (along the line-of-sight), as well as in the transverse direction, we can obtain a direct measurement of the Hubble constant as a function of redshift, *H(z)*. This would provide even more sensitive constraints on dark energy parameters. However, the error distribution in photo-z's suppresses the line-of-sight modes in *P(k)*. If the nature of that error distribution can be accurately modeled or measured, we can still reconstruct information about *H(z)*, since the degree of suppression is exponentially sensitive to the Hubble parameter. A spectroscopic sub-survey training set is required. The viability of this technique still needs to be demonstrated, but it is an exciting possibility.

In Figure 2.1.4-2, we show the error constraints on dark energy parameters resulting from BAO measurements in different redshift bands. Assuming a flat Universe ($\Omega_k = 0$), the high-z BAO data do not provide strong additional constraints. However, if we allow Ω_k to be a free parameter, the errors on w_0 and w_a increase considerably if only the low-z constraints are included. Hence, high-z BAO data are useful for breaking the degeneracy between curvature and dark energy. LSST measurements can constrain Ω_k to $\pm 10^{-3}$.

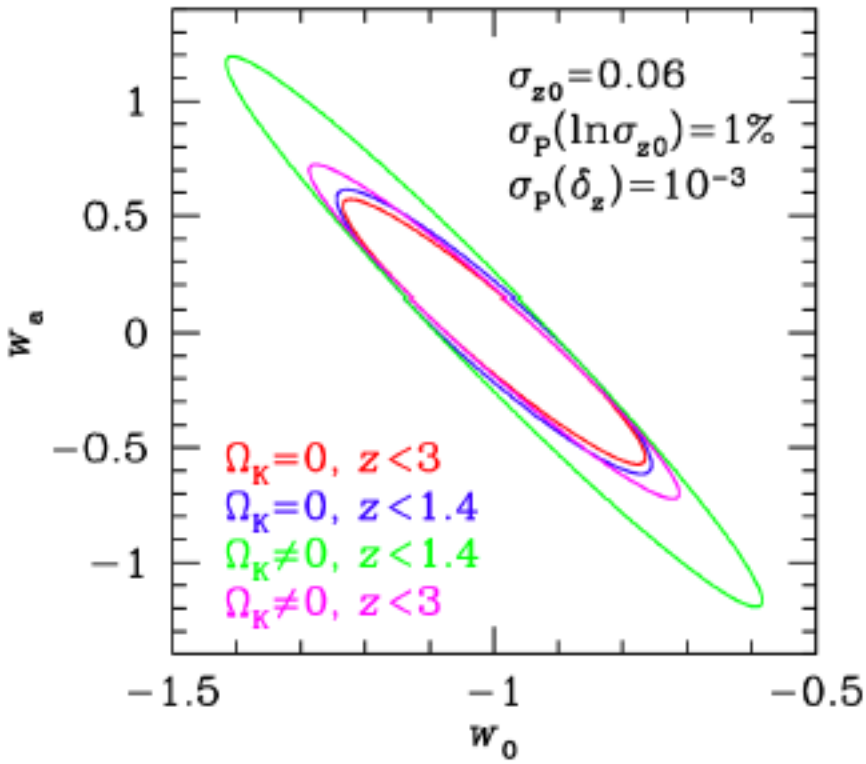


Figure 2.1.4-2: Error contours for dark energy equation of state parameters for LSST baryon acoustic oscillation measurements in different redshift bands, with and without the assumption of zero curvature.

2.1.5 Type 1a Supernovae

The use of Type 1a supernovae as calibrateable standard candles for cosmology led to the initial discovery of the acceleration of the expansion of the Universe, and this still represents one of the most promising techniques for further constraining dark energy parameters if residual systematics issues can be successfully resolved. LSST will find supernovae in two ways. The first is a result of its normal operating mode, in which each field within the 20,000 square degrees covered is sampled in the various color bands every few days. That cadence will enable the discovery of roughly 280,000 Type 1a supernovae per year. The supernova sample will have a mean redshift ~ 0.45, and will extend out to z = 0.8. The lightcurves will typically be sampled every five days in the main search filter (r-band), and there will be ~ two observations per month in the other filters. In addition, there will be a deep, pointed search in three 10 square degree fields, which will yield the discovery and close monitoring of 30,000 supernovae per year out to z ~ 1.2. A primary goal of the supernova analysis program will be to detect systematics affecting the interpretation of supernova data for cosmology, while, at the same time, providing the derivation of constraints on cosmological parameters. This will be feasible because the extremely large sample size of the LSST supernova database allows for multiple parameter fits, which can self-calibrate systematics in ways not accessible to current or other planned surveys. The systematic relations deduced from these supernovae will be helpful for current and future space-based projects targeting supernovae at even higher redshifts.

LSST will detect roughly 800 supernovae every night between its shallow and deep searches. The lightcurves (simulated examples of which are shown in Figure 2.1.5-1) will be far more

detailed than those obtained by present-day searches. For the deep search, observations are made for ten minutes per night, dividing this time in a five-day cadence among the five filters resident in the camera. The resulting lightcurves will have unprecedented time and color sampling, following tens of thousands of supernovae throughout their evolution with over 100 photometric points per lightcurve.

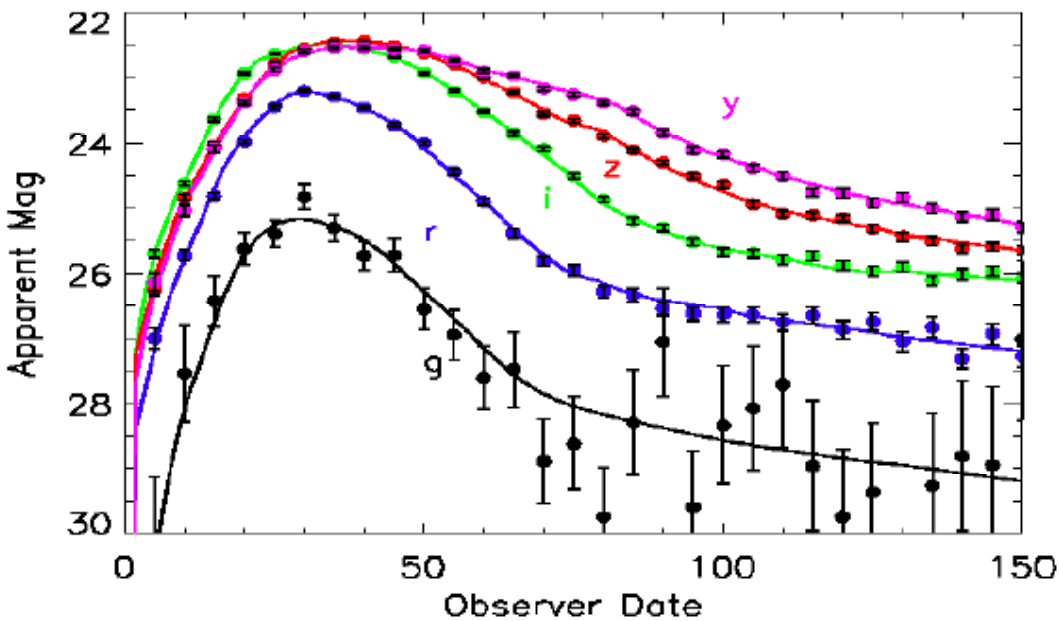


Figure 2.1.5-1 Simulated lightcurves from the LSST deep survey for a Type 1a supernova at z = 0.832. The solid lines are the input light curves. Such detailed lightcurves, combined with knowledge derived from the exceedingly large number of nearby supernovae observed, will likely allow the determination of redshifts to better than 1% based on photometric data alone. The detailed multicolor lightcurves are also important for extinction corrections and for the control of systematics arising from variations intrinsic to the supernovae themselves.

A large sample of nearby supernovae can be followed up spectroscopically by wide field spectroscopic survey facilities, such as the Large sky Area Multi-Object Spectroscopic Telescope (LAMOST) being constructed in China. With a field of view of 20 square degrees and an aperture of 4 m, there are about 11 Type 1a supernovae at any time in the LAMOST field that are at redshift below 0.3 – bright enough for spectroscopic observations. With typical exposure times of 1.5 h, LAMOST can observe about 10,000 supernovae during its survey mode every year.

Spectroscopy of supernovae at redshifts higher than 0.3 is likely to remain difficult for the foreseeable future. A sub-sample of these supernovae may be selected for spectroscopic studies, but the total number of such observations is likely to be small. The LSST lightcurves will be far more detailed than those obtained from present-day searches, however. This makes it possible to obtain photometric redshifts from the supernovae themselves, in addition to those obtained from their host galaxies. A Type 1a supernova spectrum has many strong spectral features, which provide the same opportunity for photometric redshift determination as in galaxies. Unlike a galaxy spectrum, however, the supernova spectrum evolves with time in a very specific way – one highly correlated with the width parameter as measured from the lightcurve. Thus, supernovae should be better-suited than galaxies for photometric redshift determination. The

large number of supernovae at redshifts below 0.3 can be used to calibrate these photometric redshifts.

To test the hypothesis of deducing photometric redshifts from multi-epoch supernova lightcurves, we have performed a simulation of a deep supernova search. Synthetic lightcurves are obtained from a fairly complete simulation of the observing and data reduction process, including the effects of weather, variable seeing, photon statistics, and other sources of photometric errors. Each lightcurve is then subjected to a five-parameter fit, for time of explosion, width parameter, redshift, host-galaxy reddening, and distance modulus. The preliminary results show that it is possible to derive supernova redshifts to better than 1%.

In Figure 2.1.5-2, we show our simulation of the Hubble diagram (a plot of the distance modulus, a parameter related to the logarithm of the distance, versus redshift) for 30,000 Type 1a supernovae obtained over three years in a single field, with redshifts determined photometrically. Regions of apparently increased scatter actually contain quite a small number of cases (~ 100), where the fitting procedure did not perform as well as usual. The cut-off near $z \sim 1.3$ occurs as the supernovae redshift out of the spectral range covered by the LSST filter set.

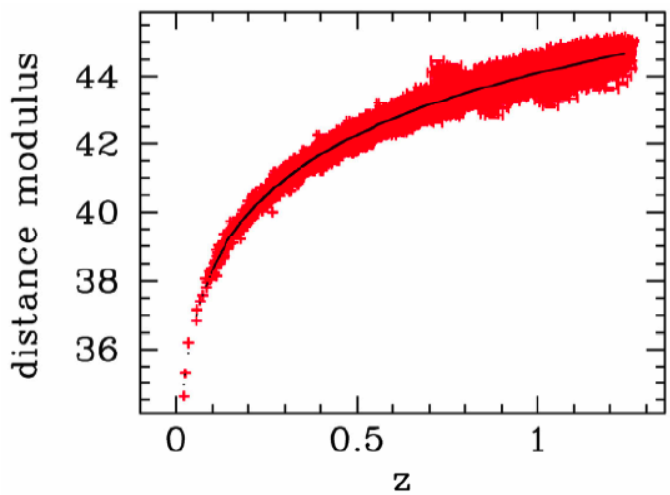


Figure 2.1.5-2 Simulated Hubble diagram of a sample of 30,000 supernovae derived from LSST observations with redshifts determined photometrically.

One-sigma error contours in the w_0-Ω_m plane derived from the LSST supernova survey are shown in Figure 2.1.5-3. Here we assume a flat universe with a constant value of w. The supernova constraints are derived for a sample of 15,000 supernovae out to $z \sim 0.85$. Also shown are the constraints imposed by the BAO measurements discussed earlier. Note that the supernvova and BAO constraints are nearly orthogonal to one another, so that the combination of the two yields a much tighter joint constraint.

LSST supernovae will also constrain angular variations of cosmological parameters across the sky. Such measurements serve several purposes: They help control systematics due to incomplete understanding of the local universe, they provide observational constraints on large-scale velocity fields in the local universe, and they constitute a direct test of the homogeneity of the cosmological parameters.

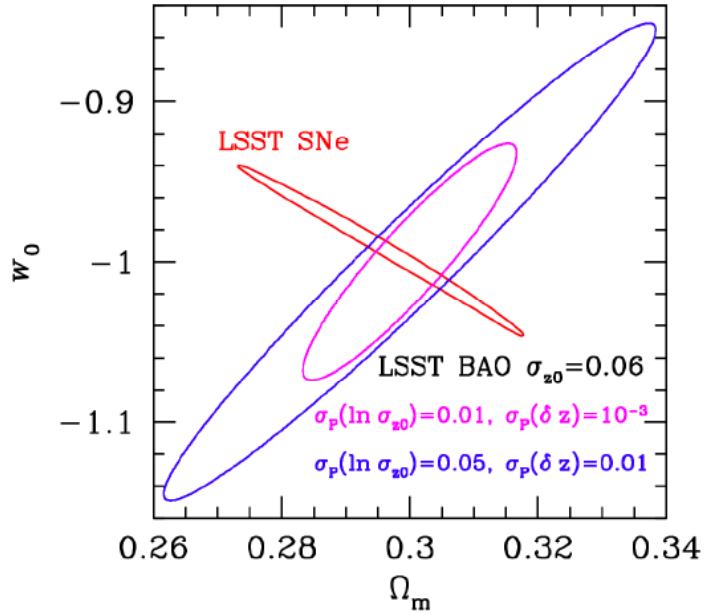


Figure 2.1.5-3 Error contours in the equation of state parameter and the matter density derived from a representative LSST Type 1a supernova sample and from the baryon oscillation measurements. The magenta and blue contours correspond to optimistic and pessimistic priors on the accuracy of the photometric error distributions, respectively.

2.1.6 Gravitationally Lensed Supernovae

When a background source lies very close to the center of a foreground distribution of mass, gravitational lensing can produce multiple images, as illustrated in Figure 2.1.6-1. This phenomenon is conventionally called "strong gravitational lensing". The positions, shapes, and ratios of intensities of the various images contain detailed information on the mass distribution of the lens, as well as on the ratio of distances to the source and the lens. Strong gravitational lenses are therefore useful for constraining the distribution of dark matter in the haloes of galaxies and clusters, and for helping us to understand the interactions of dark matter.

Strong gravitational lenses have traditionally been thought of as rare objects. However, the extreme surveying power of LSST will still produce an extremely large sample. With ~ one in 1000 high redshift objects expected to be multiply imaged, there will be literally millions of lensed objects (mostly galaxies) lying within the survey area. While only a fraction of these will be observable (due to the finite angular resolution and depth of the survey), LSST should still find some very useful exotic lenses, of the kind that would not be detectable with anything other than such a high-étendue cadence system.

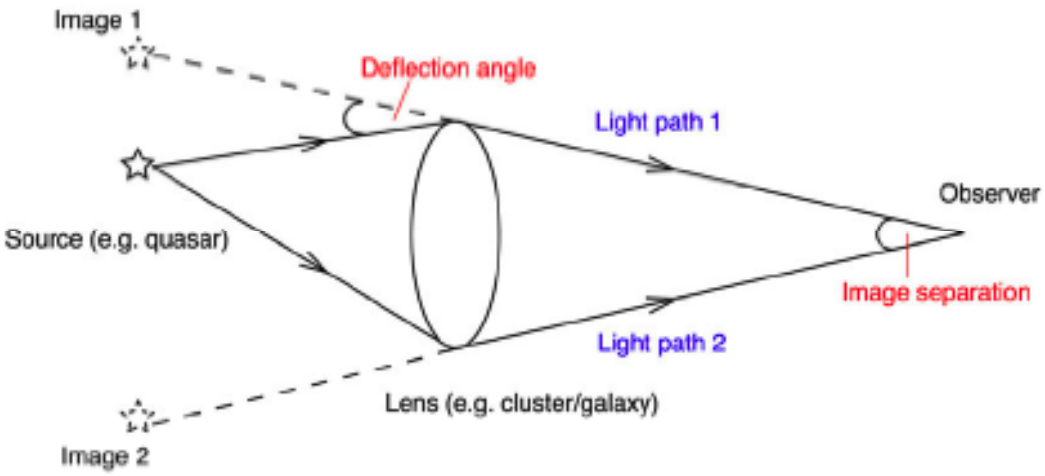


Figure 2.1.6-1: Geometry of a strong gravitational lens.

The relative arrival times of photons in the multiple images of a strong gravitational lens are also dependent on the mass distribution: If the source is variable, the time delays between the variations in each image can be measured. With the lens mass constrained by the image positions and other observables, the time delays can be predicted. Comparing the predicted delays with the measured delays gives a value for the Hubble constant. Conversely, if the Hubble constant is assumed to be known from other measurements, then the time delays provide valuable information on the mass distribution in the lensing galaxy.

Most strong lenses in the Universe are massive elliptical galaxies lensing high redshift galaxies. These systems typically have time delays of a few months. For the time delay to be observable, the source must be variable. Therefore, monitoring efforts to date have concentrated on observations of lensed quasars, which vary on timescales of days. Time delays have now been measured in over ten separate systems.

Supernovae make much better sources for measuring lens time delays: Their lightcurves show obvious, template-following peaks. These peaks last for ~ one month (more when observed at high redshift due to time dilation), requiring a monitoring system with a cadence of 1 – 2 weeks. LSST is ideally suited for this purpose. Figure 2.1.6-2 shows that, over a ten year period, and in a given area of sky to the depth that will be sampled by LSST, more supernovae occur than there are quasars. From this figure, we may anticipate observing and extracting lens time delays from a few hundred lensed supernova systems. At present, no such event has yet been observed, and it is unlikely that any will before LSST comes on line.

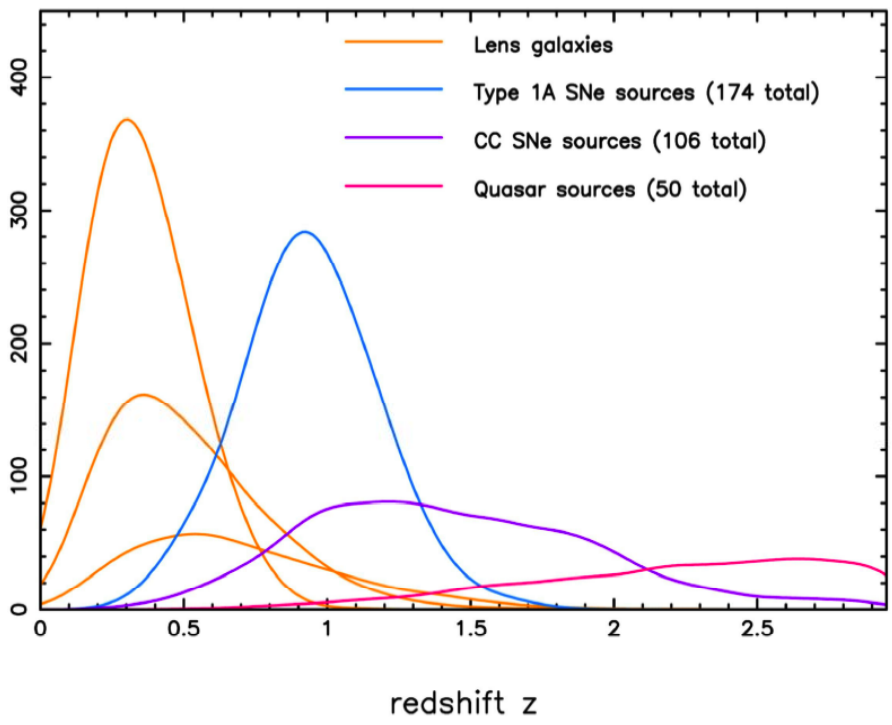


Figure 2.1.6-2: Redshift distributions of *measurable* lensed supernovae in the LSST survey area. The middle of the two lens-galaxy curves corresponds to the lensed core-collapse supernovae, while the lower curve corresponds to the quasar lenses. For this plot we assume consistent 0.7 arcsec seeing, a magnitude limit of 23 for each measurement, and a 10-year 20000 square degree survey, BUT with 50% visibility at any one time, and a completeness of 15% reflecting the fraction of data expected to be useful for time delay measurement.

The time delays that can be measured will typically have a fractional error of ~ 10%. The most useful measurements will come from the r-filter. Typical lightcurves for multiple images are shown in Figure 2.1.6-3. Measurements of this kind will provide an intriguing probe of dark matter, dark energy, and the expansion history of the Universe.

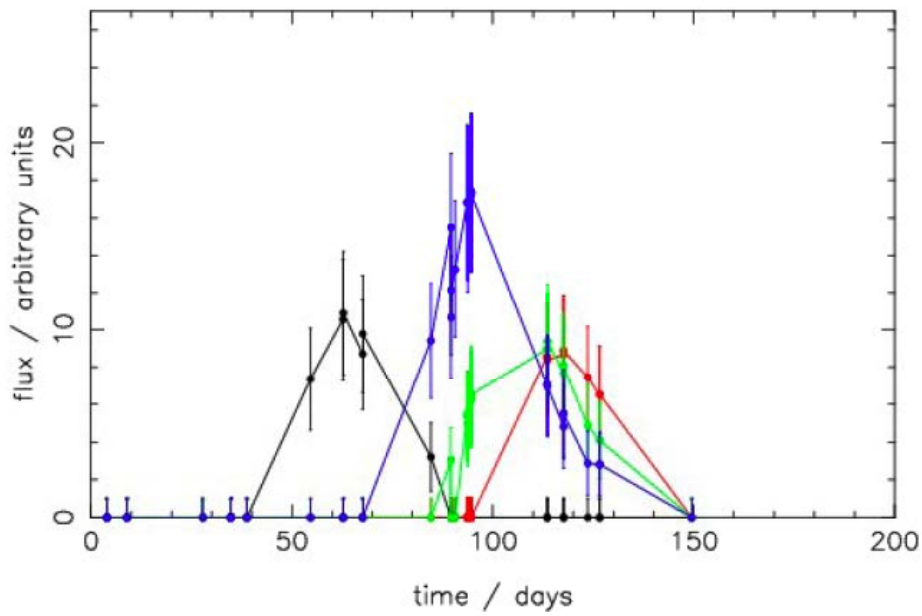


Figure 2.1.6-3: Example light curves of four-image lensed supernova images for a system in a well-sampled field. The input time delays with respect to the first (black) curve are 32, 43 and 55 days. This plot shows why the peak magnitude required for a good time delay measurement (23.0) is brighter than the LSST limiting magnitude for an exposure pair (24.5): the SN has to be visible below peak, and in the fainter images too.

2.1.7 Photometric Redshifts

The redshift distribution of galaxies plays a fundamental role in the de-projection of the 2-dimensional distribution of galaxies on the sky into the real 3-dimensional Universe, as well as in any study of galaxy evolution, and in most methods useful in constraining cosmological parameters, as we have discussed above. LSST will make photometric measurements in six colors defined by the filter bands plotted in Figure 2.1.7-1. A great strength of LSST is that from these photometric measurements, we will be able to derive redshifts for each galaxy with a typical precision of ~0.07(1+z) per galaxy, without any pre-selection. Better precision may be obtained by adding combinations of magnitude and surface brightness priors or by restricting the sample to the red sequence.

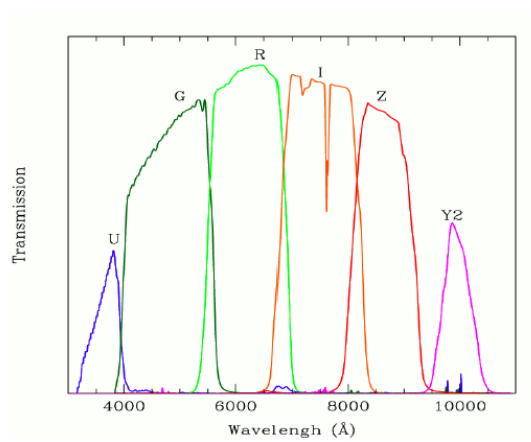


Figure 2.1.7-1: LSST filter band transmission curves.

How accurate can photometric redshifts be? Intrinsic limitations to photometric redshifts arise from constraining *z* from a limited number of observations, with limited prior knowledge about the galaxy spectral energy distribution (SED) (and its evolution, reddening, etc.). The Hubble Deep Field North (HDFN) has been widely used as a standard data set for testing photometric redshift techniques due to its depth, wavelength coverage, and number of photometric bands (7 total: HST~UBVI and ground JHK), and the best results obtained so far for this data set are $z \sim 0.06\ (1+z)$. The accuracy improves dramatically when brightness priors are invoked and can be further improved if the SED is somehow known (for example focusing on luminous red galaxies, which have uniform properties). The difficulty is cleanly pre-selecting a certain type of galaxy, and the trade-off is that the density of objects decreases dramatically. Representative spectroscopic training sets are difficult to obtain, especially for a deep survey. Therefore we plan to use SED fitting methods, even though training set methods may still be useful for subsets of the LSST data.

We have simulated LSST photometric redshifts in two ways. The first involves using models of different galaxy SEDs at a range of redshifts, convolving them with the LSST filter response, adding the estimated photon noise, and then analyzing the recovered photometric redshifts. The results are illustrated in Figure 2.1.7-2. As can be seen, catastrophic photo-z errors are decreased by adding the u band and can be further minimized using priors such as the luminosity function and the surface brightness.

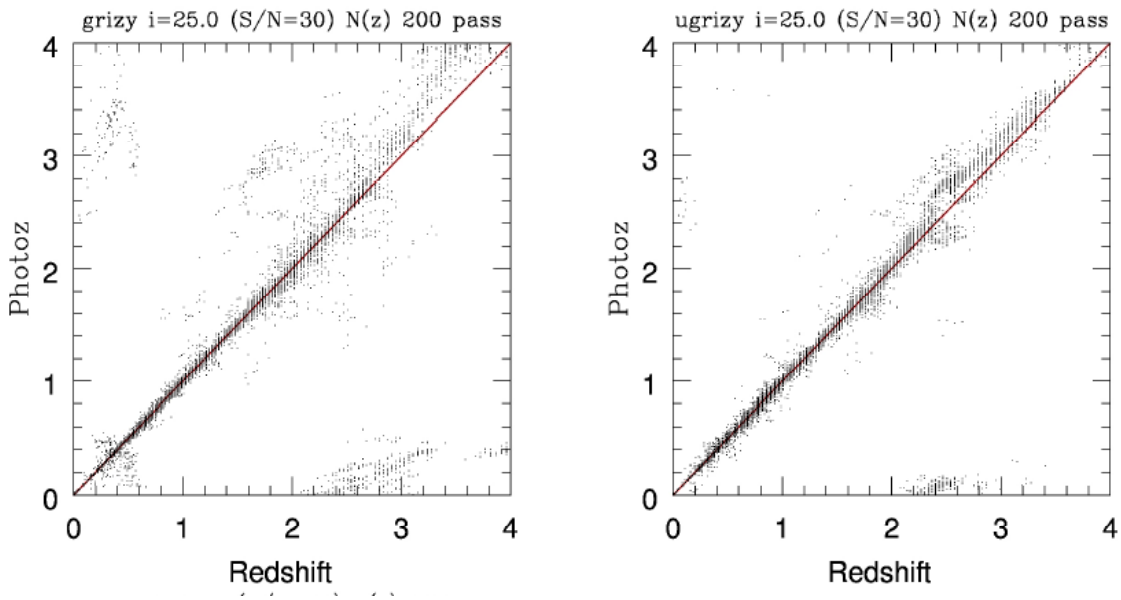


Figure 2.1.7-2 Photometric versus spectroscopic redshifts for grizY filter (*Left*) and for the adopted ugrizY filter set (*Right*) for the LSST survey after only 5 years. No brightness priors were used in this analysis.

A second approach is based on real UBVIJ data. First, we convolved the HDFN UBVI space images with 0.7" FWHM seeing. Then we re-pixelized, adding noise, and cataloged the images to match the expected data quality for the final full depth stack of exposures. The final stack will go to 26.7AB and 25.4AB (10-sigma and 30-sigma in the i band.) The photometric redshift technique used here is based on SED fitting and on a magnitude or luminosity function prior. We achieve $<(z_{phot}-z_{spec})/(1+z_{spec})> = 0.01 \pm 0.09$ for all the detected objects. If the 5% worst outliers are excluded, $<(z_{phot}-z_{spec})/(1+z_{spec})> = 0.01 \pm 0.07$.

As emphasized in the preceding sections, many of the LSST probes of dark energy depend on accurate estimates of the photo-z scatter and bias vs z. As seen in Figure 2.1.7-2, the scatter and bias are well defined, but they must be measured (calibrated). An intensive spectroscopic campaign in selected areas is required. Indeed this is needed by the community. By necessity this will involve faint spectroscopic calibration of even fainter 10-15 band ultra photo-z data on a representative sample of 100,000 galaxies.

2.1.8 The Value of Multiple Probes of Dark Energy

As covered above, LSST will enable a variety of independent probes of dark energy. There is a strong driver for this kind of multi-pronged approach. We currently have no idea what dark energy is. Today there are toy models, parametrized by w_0 and w_a, which span a range of 0.3 in w_0, and 0.8 in w_a. The multiple probes that LSST will provide lead to multiple checks for systematics. When the results from LSST analyses are combined with those from CMB measurements with Planck and complementary supernova and lensing measurements with the Joint Dark Energy Mission (JDEM), percent level precision on dark energy parameters can be reliably achieved.

One means of visualizing the complementarity of different techniques is by examining how they independently determine $w(z)$. This can be accomplished by binning $w(z)$ in redshift bins, and then diagonalizing the error covariance matrix produced by propagating errors for a given set of measurements using a particular technique. The result is a set of eigenmodes and eigenvalues, which illustrate the sensitivity of that technique for characterizing w in different regimes of redshift space.

In Figure 2.1.8-1, we show the results of that analysis for the cosmic shear measurements that will be made by LSST and for the precision supernova constraints that may come from a sample of 2000 supernovae measured by JDEM. We see a striking difference in the modes for these two probes. In particular, the LSST lensing constraints stretch to higher z. The reason for this is that lensing is less sensitive to the growth factor at the lower redshifts where the source density in a given redshift bin is small and the lensing window (for sources at higher z) is also small. Thus, the supernovae are better at detecting changes in $w(z)$ at lower z, and the cosmic shear measurements tend to be better at detecting changes at higher redshift.

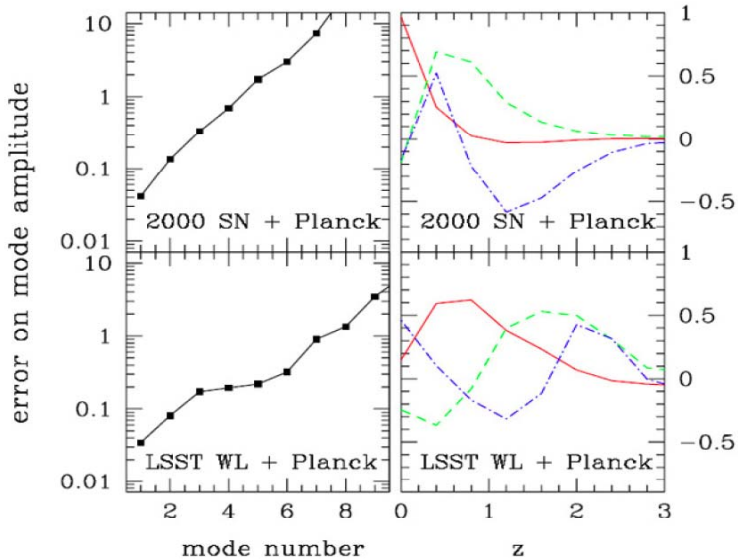


Figure 2.1.8-1 Principle components of *w(z)*. The eigenvalues (*Left*) and the first three eigenmodes (*Right*) of the *w(z)* covariance matrix for LSST measurements of cosmic shear + Planck, and for a JDEM measurement of 2000 supernovae + Planck.

The lensing and supernova techniques also have strikingly different eigenvalue spectra. The error on the amplitude of the best determined mode is quite similar for each (~ 0.03), but this degrades much faster for higher modes for the supernovae. The cosmic shear constraints yield six modes with $\sigma < 0.5$, whereas this is only true for the first three modes for the supernovae.

The use of multiple probes may also potentially allow us to discriminate whether the existence of dark energy is due to new physics in the stress-energy tensor (as commonly assumed), or due to new gravitational physics. This is because the cosmic shear and cluster measurements are not only sensitive to *H(z)*, but also to the rate of growth of the large-scale density field. In linear perturbation theory, the growth of structure can be very simply described. The growth rate depends on the gravitational force law. If the kinematic history of the universe becomes well-constrained by supernova and baryon oscillation measurements, we can use the growth of structure diagnostics to constrain departures from Newtonian gravity.

There is historical precedent for phenomena that suggested the existence of new, unseen forms of matter, but were later found to owe their true explanation to a new theory of gravity. The anomalous perihelion precession of Mercury detected in the 19th century was first attributed to unseen matter by Leverrier in 1860. Of course, we now know that this anomalous precession is due to corrections to Newtonian physics as accounted for in Einstein's theory of general relativity.

3 Science Requirements Flowdown

3.1 Science Requirements for the LSST

3.1.1 Introduction

The LSST consortium has identified four science programs as the key drivers of the science requirements for the project:

- Constraining Dark Energy and Dark Matter
- Taking an Inventory of the Solar System
- Exploring the Transient Optical Sky
- Mapping the Milky Way.

The selection of these four areas as the key science programs for the LSST was the result of discussions within the consortium and reflects the input of

- the three NRC studies that have endorsed the LSST,
- the report of the LSST Science Working Group (SWG), an independent committee formed by NOAO to represent community interests
- the scientific interests of the partners in the LSSTC, and
- the physics and astrophysics community.

The SWG report (available as http://www.lsst.org/Science/docs/DRM2) should be consulted for a more detailed discussion of the major scientific advances that can be expected from the construction of a wide-field telescope that is dedicated to repeated, deep, multi-color imaging of the sky.

For each of the four primary science drivers selected by the LSSTC, this section briefly describes the science goals and the most challenging requirements for the telescope and instrument that are derived from those. Tables are also provided that integrate the detailed requirements of these four programs. If these requirements are met by the LSST — and indications of the preliminary engineering studies undertaken to date indicate that they can be — then the LSST will not only enable all four of these major scientific initiatives but will also make it possible to pursue many other research programs. Some examples are described in the SWG report, but the long-lived data archives of the LSST will have the astrometric and photometric precision needed to support entirely new research directions which will inevitably develop during the next several decades.

3.1.2 Constraining Dark Energy and Dark Matter

Driven by observations, current models of cosmology require the existence of both dark matter and dark energy (DE). One of the primary challenges for fundamental physics is to understand these two major components of the universe. The primary DE science drivers for LSST come from a suite of two and three point cosmic shear tomography analyses coupled with galaxy power spectrum and baryon acoustic oscillation (BAO) data, as well as from the use of supernovae as standard candles. Due to its wide area coverage, LSST will be uniquely capable of measuring at least 7 parameters related to DE: the lowest 6 eigenmodes of the DE equation of state vs. redshift $w(z)$ and any directional dependence. DE exerts its largest effects at moderate redshift; LSST's redshift coverage will bracket the epoch at which DE began to dominate the cosmic expansion.

When combined with Planck CMB data, these data from LSST will sharply test models of DE, whether due to new gravitational physics, vacuum energy, or other causes.

3.1.2.1 Weak Lensing Studies

Weak lensing (WL) techniques can be used to map the distribution of mass as a function of redshift and thereby trace the history of both the expansion of the universe and the growth of structure. These use common deep wide-area multi-color imaging with stringent requirements for the shear systematics in at least two bands and photometry in all bands. These requirements are covered in more detail in the LSST DETF report and references therein.

The shear systematic errors can be mostly corrected by use of foreground stars. The PSF within each exposure must be mapped, fit, and corrected. The precision of this correction depends on how many stars are available, and thus depends on the angular scale. The overall scale of the combined errors is set by the requirement of distinguishing models of the origin of DE: unique sensitivity to the cosmic shear power spectrum from arcminute to 100 degree scales and wide redshift range, the ability to probe at least five DE eigenfunctions, and any variation over the sky. This leads to an etendue requirement for *areal coverage times depth* (several billion source galaxies to z=3), as well as photometric precision and wide angular coverage (> 90 deg). [Shear on angular scales less than a few arcminutes is compromized in cosmological interpretation due to baryon coupling.]

The power of the LSST relative to existing weak lensing surveys derives from its ability to survey much larger areas of the sky to faint limiting surface brightness while maintaining exquisite control of systematic errors in the galaxy shapes. Therefore, characterizing dark energy places particularly strong requirements on the total area of sky covered, the depth of the stacked image, the number of revisits to each field, the ellipticity and sampling of the point spread function (PSF), and the choice of filters, which must be suited to obtaining accurate photometric redshifts. At least five bands are required. Photometric precision of at least 1% is required, as well as calibration of photometric redshifts over the redshift interval 0.1 – 3.

The scale of residual shear errors should be set by the statistical error floors, not systematics. The two components of statistical shear errors vary oppositely with angular scale. On small angular scales (< few arcminutes) the source galaxy shear error is dominated by the random "shot" noise of the galaxy intrinsic ellipticities (about e=0.3 rms per galaxy) and the finite areal density of source galaxies. On large angular scales the source shear error is dominated by large scale structure cosmic variance. The cross-over point varies with source redshift. For all redshifts in projection, over the range of angular scales for LSST WL science the two errors sum to nearly a constant statistical shear power of 3×10^{-7}, or a source rms residual ellipticity of 0.001. The residual shear power systematics vs. angular scale (*after PSF corrections*) must be less than 30% of the statistical shear power. This includes correlations between angle bins. To achieve this goal, the residual shear power systematics (after corrections) must track the statistical errors at a factor of ~3 lower level. While the statistical error is uncorrelated with angular scale (source galaxies randomly oriented), systematic errors are typically correlated. Therefore, statistical errors beat down when averaged over a broad angular band, but systematics do not unless they are chopped or are stochastic from seeing.

3.1.2.2 Supernovae

Supernovae (SN) provided the first evidence that the expansion of the universe is accelerating. LSST will be a powerful SN factory. Operating in a standard mode of repeated scans of the sky with images taken every few days and with exposures of 30 seconds, LSST will discover 250,000 Type Ia SN annually. Their mean redshift will be z~0.45 with a maximum redshift of ~0.7. These data, when combined with priors from other experiments, can constrain the lowest eigenmode of

w (*i.e.* the mean value) in the nearby universe to 1 percent, and given the dense sampling on the sky, can be used to search for any dependence of w on direction, which would be an indicator of new physics. Some SN will be located in the same direction as foreground galaxy clusters; a measurement of the magnification of the SN will make it possible to model the cluster mass distribution. Core-collapse SN will provide estimates of the star formation rate during the epoch when star formation was changing very rapidly. Longer exposures (10-20 minutes/band) of a small area of the sky could extend the discovery of SN to a mean redshift of 0.7 with some objects out to z~1.4. The added statistical leverage on the "pre-acceleration" era will narrow the confidence interval on both w and its derivative with redshift.

Spectroscopic follow-up for so many SNe will be impossible. Exploitation of the data from the LSST will require light-curves which are well-sampled both in brightness and color as a function of time. This is essential to the search for systematic differences in supernova populations which may masquerade as cosmological effects as well as for determining photometric redshifts from the supernovae themselves; the development of techniques for determining photometric redshifts from supernova light-curves is currently being pursued by several community groups. Good image quality is required to separate SNe photometrically from their host galaxies. Observations in five photometric bands will be necessary to ensure that, for any given supernova, light-curves in four bands will be obtained (due the spread in redshift). Absolute photometric calibration to 1 percent is adequate, but the importance of K-corrections to supernova cosmology implies that the calibration of the zero points between filters remains a serious issue, as is stability of the response functions, especially near the edges of bandpasses where the line emission from supernovae makes this more of a problem than for stellar spectra.

3.1.3 Taking an Inventory of the Solar System

The Earth orbits within a swarm of asteroids; some small number of these objects will ultimately strike the Earth's surface. The U.S. Congress has mandated that by the year 2008, 90% of the near-Earth asteroids (NEAs) with diameters greater than 1 km be discovered and their orbits determined. Impacts of NEAs of this size have the potential to change the Earth's climate and cause mass extinctions such as the one credited with killing the dinosaurs. A NASA report published in 2003 estimates conservatively that with current search techniques, about 70% of the NEAs with diameters larger than 1 km will be cataloged by 2008. This same report quantifies the risk of impacts by smaller bodies, which have the potential of causing significant ground damage and recommends as a reasonable next goal reduction of the residual hazard by another order of magnitude. Achieving this goal would require discovery of about 90% of the potentially hazardous asteroids (PHAs) down to diameters of about 140 m. While it is unlikely that any currently planned facility could achieve this goal within a decade or two, modeling suggests that the LSST could find about 90% of the PHAs with diameters larger than 250 m within ten years.

The search for PHAs puts strong constraints on the cadence of observations, requiring closely spaced pairs of observations two or preferably three times per lunation in order to link observations unambiguously and derive orbits. Individual exposures should be shorter than about 1 minute each to minimize the effects of trailing for the majority of moving objects. Because of the faintness and the large number of PHAs and other asteroids that will be detected, LSST must provide the follow-up required to derive orbits rather than relying, as current surveys do, on separate telescopes. The observations should be obtained within ±15 degrees of the Ecliptic. The images should be well sampled to enable accurate astrometry, with absolute accuracy not worse than 0.1 arcsec. There are no special requirements on filters, although bands such as V and R that offer the greatest sensitivity are preferable. The images should reach a depth of at least 24 (5σ for point sources) in the r band in order to probe the <1 km size range at main-belt distances. Based on recent photometric measurements of asteroids by the Sloan Digital Sky Survey, the

photometry should be better that 1–2% to allow for color-based taxonomic classification and light-curve measurements.

The LSST can also make a major contribution to mapping Kuiper Belt Objects (KBOs). The orbits of KBOs provide a fossil record of the early history of the solar system; their eccentricities and inclinations contain clues to past perturbations by giant planets. The sizes of the KBOs hold clues to the accretion events that formed them and to their subsequent evolution through collisional grinding, etc. The compositions of KBOs are not identical and are correlated with their dynamical state; the reasons for these differences are not known. Light curves can be used to constrain the angular momentum distribution and internal strengths of the bodies. A more complete sample of KBOs and determination of their properties can assist with selecting targets for future NASA missions. The survey for PHAs can simultaneously provide the joint color-magnitude-orbital distribution for all bright (r<24) KBOs. The 100 or so observations obtained for each bright KBO can be searched for brightness variations, but modeling will be required to determine how well periods can be extracted from observations made at random times. At the very least, it will be possible to determine amplitudes for many thousands of KBOs, and periods can likely be derived for many of them.

Long exposures would be required to push the detection of KBOs to smaller sizes and reach the erosion-dominated regime in order to study the collisional history of various types of KBOs. KBO science would be greatly amplified if a small fraction of the observing time were devoted to hour-long observations in the ecliptic. This same mode of observation may have applications to the study of variable and transient objects. Apart from exposure time and the requirement for multiple filters (at least two) to classify objects according to composition, the requirements for the KBO science are essentially similar to the requirements for the detection and orbital determination.

3.1.4 Exploring the Transient Optical Sky

The LSST will open a new window on the variable sky. Recent surveys have shown the power of variability for studying gravitational lensing, searching for supernovae, determining the physical properties of gamma-ray burst sources, etc. The LSST, with its repeated, wide-area coverage to deep limiting magnitudes will enable the discovery and analysis of rare and exotic objects such as neutron star and black hole binaries; gamma-ray bursts and X-ray flashes, at least some of which apparently mark the deaths of massive stars; AGNs and blazars; and very possibly new classes of transients, such as binary mergers and stellar disruptions by black holes. It is likely that the LSST will detect numerous microlensing events in the local group and perhaps beyond. The LSST would provide alerts for concerted monitoring of these events, and open the possibility of discovering planets and obtaining spectra of lensed stars in distant galaxies as well as our own. LSST can also provide multi-wavelength monitoring over time of objects discovered by the Gamma-Ray Large Area Space Telescope (GLAST) and the Energetic X-ray Imaging Survey Telescope (EXIST). With its large aperture, the LSST is well suited to conducting a Deep Supernova Search in selected areas. LSST will also provide a powerful new capability for monitoring periodic variables, such as RR Lyrae stars, which can be used to map the Galactic halo and intergalactic space to distances exceeding 400 kpc. Since LSST extends time-volume space a thousand times over current surveys, the most interesting science may well be the discovery of new classes of objects.

Exploiting the capabilities of LSST for time domain science requires large area coverage to enhance the probability of detecting rare events; time coverage, since light curves are necessary to distinguish certain types of variables and in some cases infer their properties (*e.g.* determining the intrinsic luminosity of supernovae Type Ia depends on measurements of their rate of decline); accurate color information to assist with the classification of variable objects; good image quality

to enable differencing of images, especially in crowded fields; and rapid data reduction and classification in order to flag interesting objects for spectroscopic and other follow up with separate facilities. Time scales ranging from ~1 min (to constrain the properties of fast faint transients such as those recently discovered by the Deep Lens Survey) to ~10 years (to study long-period variables and quasars) should be probed over a significant fraction of the sky. It should be possible to measure colors of fast transients, and to reach $r \sim 24$ in individual visits. Fast reporting of likely transients to the community is required in order to facilitate followup observations.

3.1.5 Mapping the Milky Way

The LSST is ideally suited to answering two basic questions about the Milky Way Galaxy: What is the structure and accretion history of the Milky Way? What are the fundamental properties of all the stars within 300 pc of the Sun?

Standard models posit that galaxies form from seeds planted by the Big Bang with accretion over time playing a significant role in determining their structure. Detailed study of the Milky Way can provide rigorous tests of these ideas, and the LSST will be able to map the 3-D shape and extent of the halo of our Galaxy. Specifically, the LSST will detect F turn-off stars to distances of 200 kpc; isolate stellar populations according to color; and determine halo kinematics through measurement of proper motions at distances exceeding 10 kpc. The LSST dataset can be used to identify streams of stars in the halo that are thought to provide a fossil record of discrete accretion events. The LSST in its standard surveying mode will be able to detect RR Lyrae variables and classical novae at a distance of 400 kpc and hence can explore the extent and structure of our own halo out to half the distance to the Andromeda Galaxy. The proper motions and photometric parallaxes for these stars can be used to characterize the properties of the dark matter halo in which the Milky Way is embedded.

Is our solar system with its family of planets unique? Or are there many more that contain Earth-like planets within the so-called habitable zone? How do solar systems form? Detailed exploration of our local neighborhood is key to answering these questions. The LSST will obtain better than 3σ parallax measurements of hydrogen-burning stars to a distance of 300 pc and of brown dwarfs to tens of parsecs. These measurements will provide basic information on candidate stars that merit further study in the search for companions, including planets. Residuals from the fits for position, proper motions, and parallax will be searched for the signature of Keplerian motion to identify stars and brown dwarfs with companions and provide fundamental estimates of the mass of the primaries. LSST data will be used to determine the initial mass functions for low-mass stars and sub-stellar mass objects and to test models of brown dwarf structure. The age of the Galactic disk can be inferred from white dwarf cooling curves.

Key requirements for mapping the Galaxy are large area coverage; excellent image quality to maximize the accuracy of the photometry and astrometry, especially in crowded fields; photometric accuracy of at least 1 percent to separate main sequence and giant stars; stringent astrometric accuracy to enable parallax and proper motion measurements; and dynamic range that allows measurement of astrometric standards at least as bright as $r = 15$. In order to probe the halo out to distances of 100 kpc using numerous main-sequence stars, the total depth has to reach r~27 (assuming 5% photometry in the r band at r=25.6; the SDSS has demonstrated such studies out to distances of 15 kpc using data with r<21.5). To study the metallicity distribution of stars in the Sgr tidal stream and other halo substructures at distances out to at least ~40 kpc, the total depth in the u band has to reach ~24.5. In order to constrain tangential velocity at a distance of 10 kpc to within 10 km/s the proper motion accuracy has to be at least 0.2 mas/yr. The same requirement follows from the decision to obtain the same proper motion accuracy as GAIA at its faint end (r~20). The LSST will then represent an "extension" of GAIA astrometric

measurements to 4 magnitudes greater depth. In order to produce a complete sample of the solar neighborhood stars out to a distance of 300 pc (the thin disk scale height), with 3σ or better geometric distances, the parallax measurements accurate to 1 mas are required. In summary, these requirements imply that the LSST will enable studies of the distribution of numerous main-sequence stars beyond the presumed edge of the Galaxy's halo, of their metallicity distribution throughout most of the halo, of their kinematics beyond the thick disk/halo boundary, and will obtain direct distance measurements below the hydrogen-burning limit for a representative thin-disk sample.

3.1.6 Summary of Science Requirements

Table 3.1.6-1 provides an integrated list of the scientific requirements for these four programs. This list combines the requirements derived from the considerations listed above.

Table 3.1.6-1 LSST Science Requirements Summary Table

	Parameter	Symbol	Units	Design Spec	Minimum Spec	Stretch Goal	SRD ref.
Filter Set	Filter complement	—	—	ugrizY	ugrizY	ubgrizY	§3.3.1 Table 1
	No. filters in camera	Nfilters	—	5	3	6	§3.3.1 Table 2
	Time to exchange a filter	TDFCmax	hr	8	72	—	§3.3.1 Table 2
	Filter change interval [1]	TFmax	min	2	10	1	§3.3.1 Table 3
	Out-of-band leakage per 10 nm bandwidth	Fleak	%	0.01	0.02	0.003	§3.3.1 Table 4
	Out-of-band leakage, total	FleakTot	%	0.05	0.1	0.02	§3.3.1 Table 4
Camera Rotation			deg	±90			§3.3.6
Single Visit [2] Depth	**Ensemble distribution:** Median 5σ depth (min) [3]	D1	mag	24.5	24.2	24.7	§3.3.2 Table 5
	Fraction of images for which 5σ depth exceeds Z1 (max)	DF1	%	10	20	5	§3.3.2 Table 5
		Z1	mag	24.2	23.8	24.5	
	Spatial variation: Fraction of field for which 5σ depth is brighter than median by Z2 (max)	DF2	%	15	20	10	§3.3.2 Table 6
		Z2	mag	0.2	0.4	0.2	
Minimum Exposure Time		ETmin	sec	5	10	1	§3.3.2 Table 7

[1] Maximum elapsed time between two visits in different filters
[2] Co-added pair of 15 sec back-to-back exposures
[3] *r* band, AB magnitude scale; see SRD Table 5 caption for specifications in other filter bands

Single Image Specifications

	Parameter		Symbol	Units	Design Spec	Minimum Spec	Stretch Goal	SRD ref.
Image Quality	**PSF size distribution:** [4]							
	Median delivered seeing for atmospheric seeing of 0.44, 0.6 and 0.8 arcsec		S1(0.44)	arcsec FWHM	0.53	0.59	0.51	§3.3.3 Table 8
			S1(0.60)		0.67	0.72	0.65	
			S1(0.80)		0.85	0.89	0.84	
	Fraction of images with PSF exceeding SX times S1		SF1	%	10	10	5	§3.3.3 Table 8
			SX	—	1.1	1.2	1.1	
	PSF profile:							
	Max ratio of encircled energy diam to FWHM for encircled energy of:	80%		—	2.1	2.3	2.0	§3.3.3 Table 9
		95%		—	3.5	3.8	3.3	
		99%		—	4.8	5.3	4.6	
	e.g., for fiducial delivered seeing of 0.67 arcsec, diameter for encircled energy of (max):	80%	SR1	arcsec	1.44	1.6	1.36	§3.3.3 Table 9
		95%	SR2	arcsec	2.3	2.5	2.2	
		99%	SR3	arcsec	3.2	3.5	3.1	
	Point source ellipticity distribution: [4]							
	Median ellipticity (max)		SE1	—	0.04	0.05	0.03	§3.3.3 Table 10
	Fraction of images exceeding SE2 (max)		EF1	%	5	10	5	§3.3.3 Table 10
			SE2	—	0.07	0.1	0.05	
	Median of residuals after smoothing over field of view (max)		SE3	—	0.002	0.003	0.001	§3.3.3 Table 10
	Fraction of residuals exceeding SE4 (max)		EF2	%	10	15	10	§3.3.3 Table 10
			SE4	—	0.003	0.005	0.002	

[4] *r* and *i* bands only; other bands not specified

Single Image Specifications (con'd)

	Parameter	Symbol	Units	Design Spec	Minimum Spec	Stretch Goal	SRD ref.
Photometric Quality	**Relative photometric errors:** [5] Point source magnitude repeatability (rms, max)	PA1	millimag	5	8	3	§3.3.4 Table 11
	Fraction of measurements deviating by more than PA2 from the mean	PF1	%	10	20	5	§3.3.4 Table 11
		PA2	millimag	15	15	10	
	Effects of ghosts: [6] Excess noise in multi-observation magnitude distribution	EPErr	%	10	20	5	§3.3.4 Table 12
	Fraction of image area with ghosts with gradients (1 arcsec scale) exceeding 1/3 sky noise	GhostAF	%	1	5	0.5	§3.3.4 Table 12
	Absolute photometric errors: Width of photometric zero point error distribution (rms, max) [6]	PA3	millimag	10	15	5	§3.3.4 Table 13
	Fraction of zero point error distribution exceeding PA4 [7]	PF2	%	10	20	5	§3.3.4 Table 13
		PA4	millimag	15	15	15	
	Knowledge of band-to-band zero point correlations	PA5 (g-r)	millimag	5	10	3	§3.3.4 Table 14
		PA5 (r-i)		5	10	3	
		PA5 (all other)		10	15	5	
	Knowledge of correlation of photometric magnitudes to external physical scale	PS6	millimag	20	50	10	§3.3.4 Table 15

[5] *g*, *r* and *i* bands; PA1 and PA2 in *u*, *z* and *Y* may be 50% larger
[6] *r* and *i* bands only; other bands not specified
[7] *g*, *r* and *i* bands; PA3 and PA4 in *u*, *z* and *Y* may be factor of 2 larger

Single Image Specifications (con'd)

	Parameter	Symbol	Units	Design Spec	Minimum Spec	Stretch Goal	SRD ref.
Astrometric Quality	**5 arcmin scales (4Kx4K sensor):** [8]						
	Point source distance repeatability (rms, max)	AM1	milli-arcsec	10	20	5	§3.3.5 Table 16
	Fraction of above distribution deviating by more than AD1 from the median	AF1	%	10	20	5	§3.3.5 Table 16
		AD1	milli-arcsec	20	40	10	
	20 arcmin scales (raft): [8]						
	Point source distance repeatability (rms, max)	AM2	milli-arcsec	10	20	5	§3.3.5 Table 16
	Fraction of above distribution deviating by more than AD2 from the median	AF2	%	10	20	5	§3.3.5 Table 16
		AD2	milli-arcsec	20	40	10	
	200 arcmin scales (camera): [8]						
	Point source distance repeatability (rms, max)	AM3	milli-arcsec	15	30	10	§3.3.5 Table 16
	Fraction of above distribution deviating by more than AD3 from the median	AF3	%	10	20	5	§3.3.5 Table 16
		AD3	milli-arcsec	30	50	20	
	Color differences in astrometric mapping:						
	Difference in distances measured in r and other bands (rms, max)	AB1	milli-arcsec	10	20	5	§3.3.5 Table 17
	Fraction of above distribution deviating by more than AB2 from the mean	ABF1	%	10	20	5	§3.3.5 Table 17
		AB2	milli-arcsec	20	40	10	
	Absolute accuracy: knowledge of median error in absolute astrometric positions	AA1	milli-arcsec	50	100	20	§3.3.5 Table 18

[8] *r* and *i* bands; other bands not specified

Full Survey Specifications

	Parameter		Symbol	Units	Design Spec	Minimum Spec	Stretch Goal	SRD ref.
Survey Cadence	Median number of visits per sky location (min)	**Design Depth (ref)**						§3.4 Table 19
		24.3	Nv1 (u)	—	10	8	12	
		26.5	Nv1 (g)	—	40	32	48	
		27.8	Nv1 ®	—	400	320	480	
		26.6	Nv1 (i)	—	300	240	360	
		25.5	Nv1 (z)	—	100	80	120	
		24.7	Nv1 (Y)	—	150	120	180	
	Observing time allocated to special programs, min & max		SPTmin	%	5	1		§3.4 Table 20
			SPTmax	%	10	20		
	Area with fast (30–1,800 sec) revisits (min)		RVA1	deg^2	2,000	1,000	3,000	§3.4 Table 21
	Area with 25% of visits separated by >5 yrs (min)		RVA2	deg^2	15,000	10,000	20,000	§3.4 Table 22
	Area with 25% of visits spanning at least 4 calendar months (min)		RVA3	deg^2	15,000	10,000	20,000	§3.4 Table 22
Image Quality	**Point source ellipticity distribution after stacking:** [9]							
	Median ellipticity (max)		TE1	—	0.0001	0.0002	0.00005	§3.4 Table 23
	Fraction of images exceeding TE2 (max)		TEF1	%	15	25	15	§3.4 Table 23
			TE2	—	0.0002	0.0004	0.0001	
Data Processing	Certified data release interval		DRT1	year	1.0	2.0	0.5	§3.4 Table 24
	Optical transient alert latency		OTT1	min	1.0	2.0	0.5	§3.4 Table 24

[9] *r* and *i* bands; other bands not specified

3.2 Requirements Traceability Matrix

3.2.1 Introduction.

The LSST has developed a Traceability matrix that allows for both a visible display and careful mapping of the scientific missions to engineering requirements and implementation. The matrix is separated into 3 distinct regions, Science Requirements, Engineering Requirements, and Implementation as shown in Figure 3.2.1-1

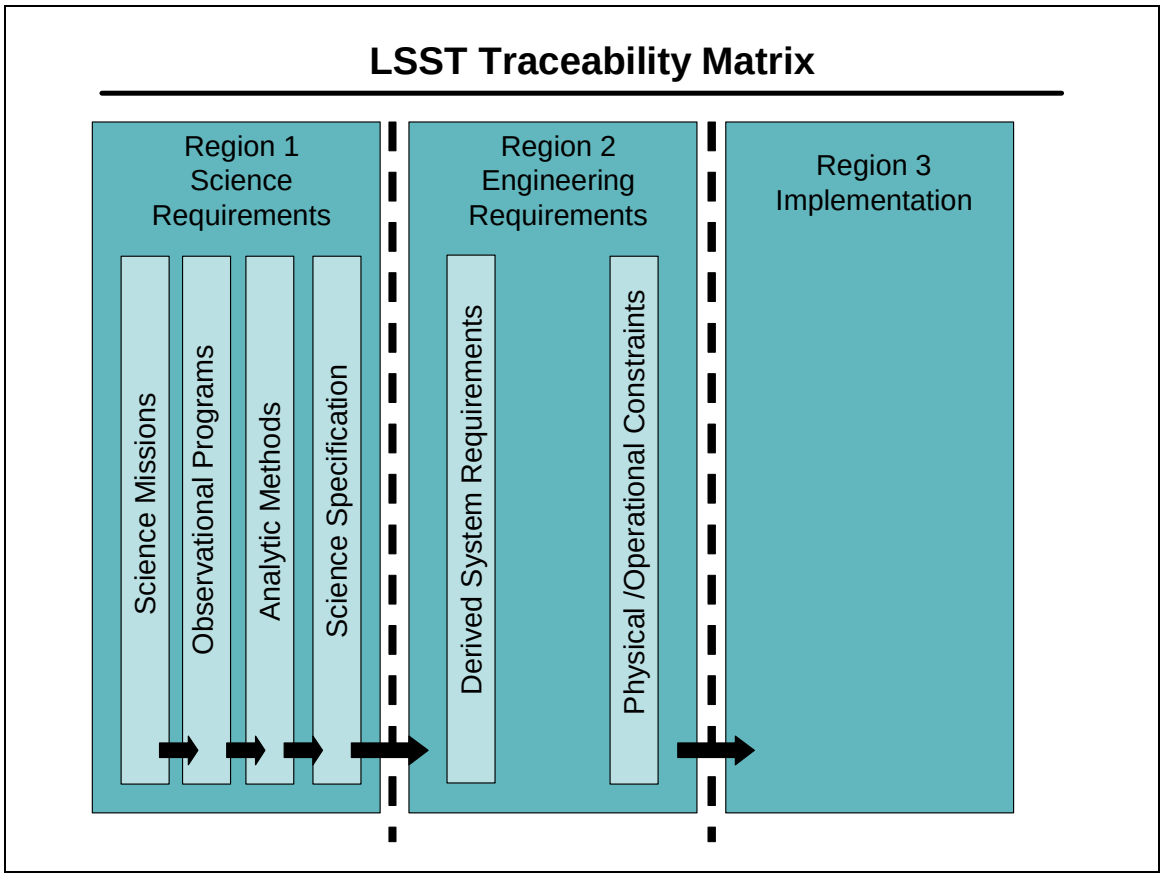


Figure 3.2.1-1: Principle sections and subsections of the traceability matrix showing the relationship between science requirements, engineering requirements and ultimately the specific implementation of the LSST.

The current focus of development is in the first two regions. In Region 1 the four science missions, used to define the overall system requirements, Dark Energy/Dark Matter, Solar System Science, Optical Transients, and Galactic Science are listed in the first column. The observational programs that address the science in each of these missions is identified in the second column with an indicator to show the mission it supports. The third column introduces the analytic methods that are used to address the observational program and its data with numerical indicators to show the second column origin. The final column is the full science specifications that are also

found in the LSST Science Requirements document. All the science specification are also numerically keyed to indicate origin in the previous column. This traceability matrix shows the basis of the derived scientific requirements and allows the details to be tracked back to its original mission with the additional fidelity of identifying the observational program and analytic methods which each introduce requirements in order that the LSST achieves the necessary performance with the served data. Figure 3.2.1-2 shows the traceability mapping used in the matrix.

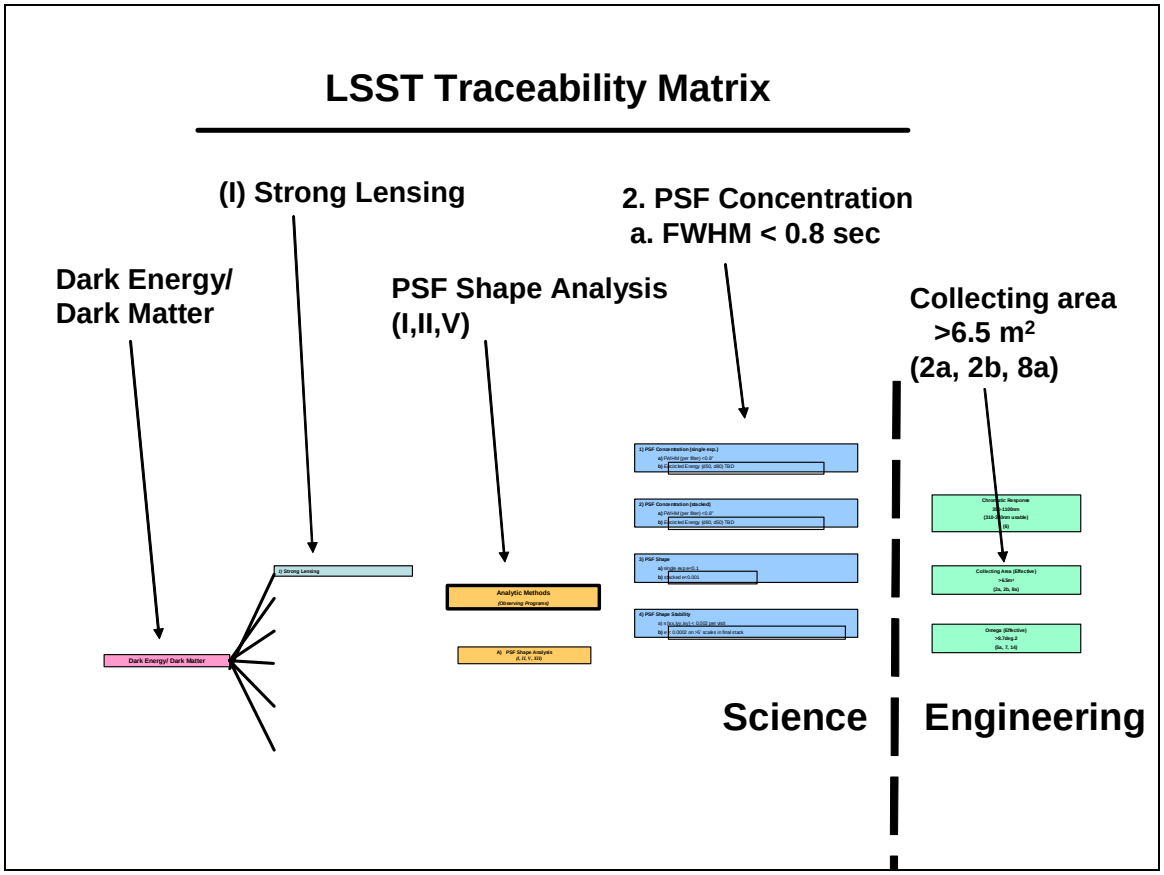


Figure 3.2.1-2 A specific example of how the traceability matrix follows a science mission (Dark Energy/Dark Matter) through an observing program, techniques, science requirements to engineering requirements.

In region 2 the requirements are transformed and tracked in engineering terms. The first column is a direct mapping of the Science Specifications of Region 1 translated into engineering parlance. The subsequent four columns in region 2 show the distribution of the key engineering specifications as they affect the major subsections of the LSST, the Camera, the Telescope, the Data Management system and the Telescope Site. The final column in region 2 is included to identify the additional constraints that must be considered in the design. These constraints are physical limitations, operational issues, financial and schedule directives and other project requirements that do not map directly to science missions.

The current version of the LSST traceability matrix includes no detail in region 3 at this time. Baseline approaches exist for nearly all the elements of the LSST but the focus for the traceability matrix now is the requirements and will focus toward the final region as the development program continues. The traceability matrix will be a living tool throughout the development phase. As the scientific tools and analytic approaches progress and develop, the impact to the

LSST requirements will be easily assessed through the matrix. As requirements are questioned, the origin can quickly be established for detailed assessment. The traceability matrix will serve the scientific, engineering and management aspects of the LSST throughout the design, development and implementation phases.

3.3 Subsystems budgets

3.3.1 Introduction

The top-down budgets are developed to support the requirement flow-down process and to demonstrate the systems performances. They are mainly based on preliminary performance analysis, subsystem error analysis, and estimation of achievable tolerances from previous projects. Initial estimations include approximation of correction possible and of perceived difficulty of a task. These preliminary budgets will continually be revised as the project progresses to provide a mechanism for matching design and manufacturing changes to the top-level requirements.

Four main top-level error budgets have been identified, from which subsystem requirements would be derived: Pointing and Tracking, Image Quality, Ellipticity and Optical efficiency. As a top-down approach, each budget is composed of sub-categories which add-up to meet the overall specification, leading occasionally to some very tight subsystem requirements. Statistical quantities are usually added in quadrature, as the square root of the sum of the squares (RSS).

These budgets are usually given for the zenith pointing case. But allowance versus zenith angle is also included for areas which may change significantly with pointing angle. In that case, allocations are degraded by the three-fifth power law of the secant of the zenith angle. The budgets allocations are after applying correction or calibration. Assumptions are taken as to how well these corrections may be performed. The net error after correction is given in these budgets.

3.3.2 Pointing and tracking

LSST, with its large field of view, requires a very accurate pointing and tracking performance. The high efficiency specified for the whole system implies that most observations will need to be acquired in blind pointing mode. In that mode, the telescope control system relies on the accuracy of its pointing model to point the telescope at the right position, without any optical feedback from the sky. Moreover, wavefront sensors will be distributed at fixed positions in the focal plane. Although the type of WFS has not been selected yet, it is probable that their field of view will be limited. Having the star fall precisely into the WFS acquisition field becomes then an important requirement. Finally, a non-negligible pointing error could have a large trailing effect on the images during tracking that would affect the image quality performances. This is particularly important if LSST exposures are unguided, relying on the accuracy of the pointing model in open-loop.

For these reasons, LSST requirements for pointing and tracking are as critical as those usually set for large telescopes with smaller field of view:

- At all elevation, the absolute pointing requirement is 1.5” RMS with a goal of 1” RMS
- After applying an offset equal to its field of view, the pointing error must not be more than 0.5” RMS with a goal of 0.1” RMS
- For smaller offset, the pointing error must not be more than 0.1” RMS
- Open loop tracking must be better than 0.02” RMS over a 1min interval, and better than 0.4” in 1 hour.

Studies done by P.T. Wallace from the Rutherford Appleton Laboratory (UK) for the LSST project has shown that an error of 10” on some of the terms of the pointing model could generate a 1.4” star trail on a 1min exposure near zenith (see Figure 3.3.2-1). These studies were based on a 3deg diameter blind spot. To insure that these errors would cause a trailing smaller than 0.1” entails that an accuracy of 0.7” is required on some of these terms.

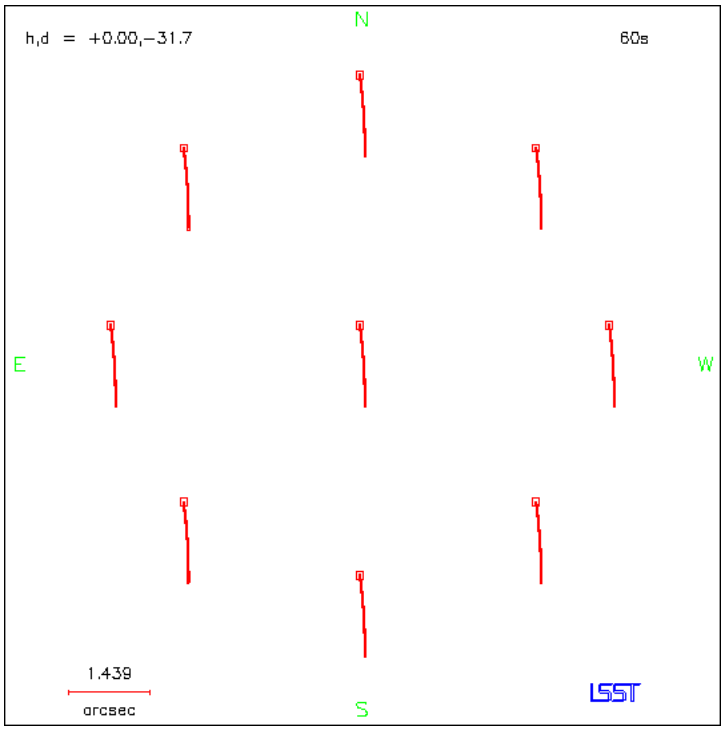


Figure 3.3.2-1 Example of trailing in the LSST FOV caused by a 10 arcsec error in one of the term of the pointing model (the OTA/el nonperpendicularity CA)

As the design progresses, we will investigate ways to mitigate the impact of errors in the pointing model in order to ease some of the requirements. For example, because the worst effects of pointing-model inaccuracy occur near the zenith, allowing for a wider blind spot diameter would relax some requirements.

3.3.3 Image Quality

For the initial analysis, an overall performance of 0.3” FWHM was selected for the top down image quality error budget. This performance would correspond to a degradation of ~10% of the image quality for a median atmospheric seeing of 0.7”. The total error is distributed among all the categories and subcategories identified as potential sources of image quality degradation. This distribution is based on information available from other projects, results from modeling studies or estimates, and all these allocations appear achievable.

For example, in Figure 3.3.3-1, is presented a surface error map of the Large Binocular Telescope (LBT) 8.4m primary mirror #1 where we added a mask in the central area to represent the large central obscuration in the LSST primary mirror. The original data was measured on the real LBT mirror at the end of its fabrication. This data is applicable to LSST as our primary mirror will be fabricated in a similar fashion as LBT primaries. The total RMS surface error is equal to 24nm on this image after removal of spherical aberration and 6 other flexible bending

modes. As a result, the M1 fabrication error has an allocation of 0.11” FWHM in the LSST image quality error budget.

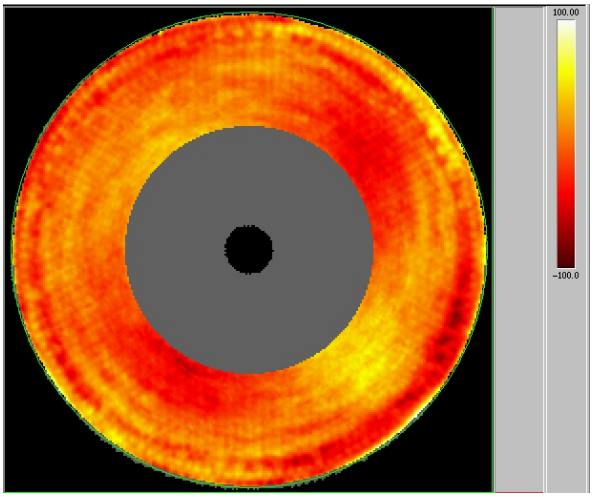


Figure 3.3.3-1 LBT 1 surface error (nm) over LSST clear aperture. Spherical aberration and 6 flexible bending modes have been subtracted. RMS surface error = 24 nm.

The allocations to the main categories and subcategories are summarized in Figure 3.3.3-2. The total allocations are 0.15” FWHM for the camera and 0.25” FWHM for the telescope.

Three main sublevels have been identified for the camera: the corrector including the lenses and filters, the detector and the optical alignment. The corrector allocation error takes into account the large diameter of the refractive elements. The detector error includes diffusion and divergence effects caused by thick detectors (thicker detectors are needed to improve the quantum efficiency in the red).

The subcategories under telescope have been grouped into static and dynamic errors. The best achievable image quality is set by the optical design. Other static errors include mirror fabrication, support and thermal errors. On the dynamic side, the active optics system is intended to correct for the mirror figure distortions at low temporal bandwidth. Wavefront sensors will be deployed in the focal plane for that purpose. The active optics error is an estimate of the accuracy of the wavefront sensors. This allocation will be revisited during R&D following the choice of wavefront sensors. Wind forces blowing on the mirrors could have the potential to distort the mirror surfaces with a non-negligible residual error even after active optics correction. For LSST, this wind buffeting on the mirrors is estimated as a small error because of the inherent stiffness of the honeycomb structure. The large secondary mirror could present a significant cross section for wind shake that could produce image jitter during an exposure. This source of tilt error would be corrected by a guiding system working at a fairly fast temporal bandwidth. The allocated image jitter error is mainly the residual windshake after guiding correction and does not include atmospheric turbulence. A detailed analysis will be done during R&D to model this error versus sampling frequency and wind speed. The camera behavior is included in these dynamic allocation errors (represented by the dashed lines in the Figure 3.3.3-2).

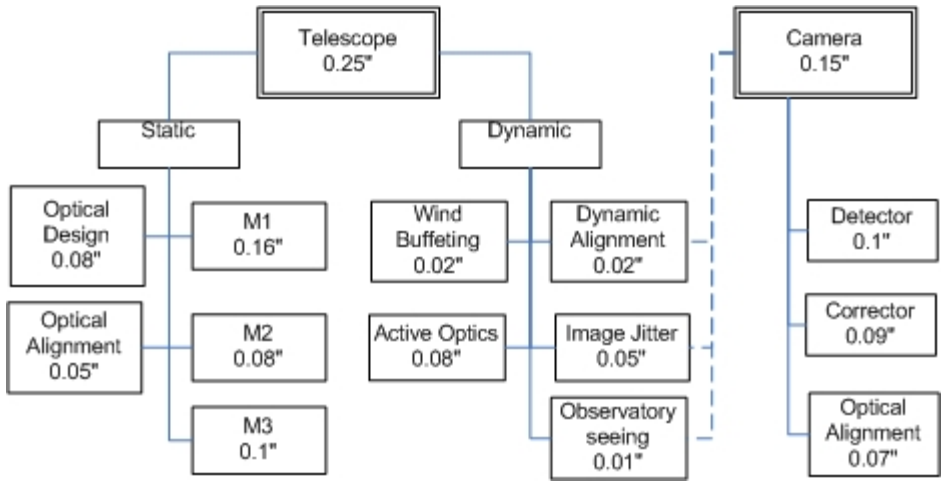


Figure 3.3.3-2 Image quality FWHM error budget allocations in arcsec

3.3.4 Ellipticity

The science requirement document specifies a PSF ellipticity better than 0.1 for each single exposure. The PSF ellipticity *e* used in the gravitational lensing community is equal to $e = \sqrt{e1^2 + e2^2}$ with $e1 = \frac{(ixx - iyy)}{(ixx + iyy)}$ and $e2 = \frac{2(ixy)}{(ixx + iyy)}$ and *ixx*, *iyy* and *ixy* being the second moments of the object intensity. For the purpose of this budget, we have assumed that the PSF ellipticity is aligned along the x direction (with *ixy*=0), that *iyy* is equivalent to the PSF semi-minor axis and *ixx* is equivalent to the PSF semi-major axis. This is a worst-case scenario. With these approximations, the PSF ellipticity becomes $e = 1 - \frac{2(FWHM)}{FWHM + \sqrt{FWHM^2 + Errors^2}}$ where *FWHM* corresponds to a Gaussian seeing distribution and *Errors* is the RSS of the error contributions that caused the elliptical PSF along the x direction.

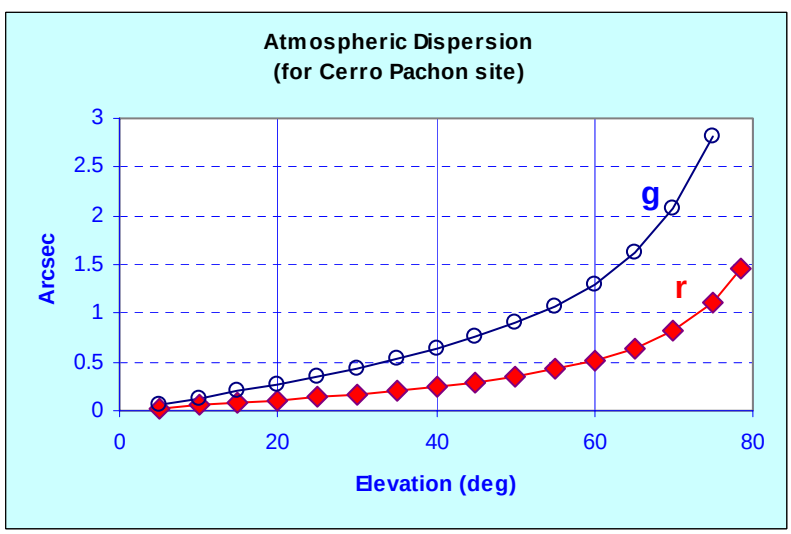


Figure 3.3.4-1 Atmospheric dispersion for the g and r filters

The atmosphere is one of the main contributors to this ellipticity budget, as the atmospheric dispersion stretches each star into a spectrum with blue at the top and read at the bottom. The atmospheric dispersion calculations are based on P.T. Wallace formulation study done for the LSST project. The major parameter in those equations, for a given site, is the range of zenith distance allowed for the telescope. As the zenith distance increases, the atmospheric dispersion becomes predominant especially in the blue wavelength (Figure 3.3.4-1). The ellipticity budget includes other contributors associated with tracking errors (that generate image trailing) and telescope errors (from vibrations and wind shake) with minor allocations.

An ellipticity distribution can be computed for this initial budget by taking an assumption on the airmass distribution of the observations. In the future this information will be extracted from the operation simulator. The resulting ellipticity distribution can then be compared to the specification. For instance, such results are shown in Figure 3.3.4-2 for seeing conditions of 0.8" FWHM. With these conditions, practically all r exposures and 80% of the g exposures have an ellipticity below 0.1.

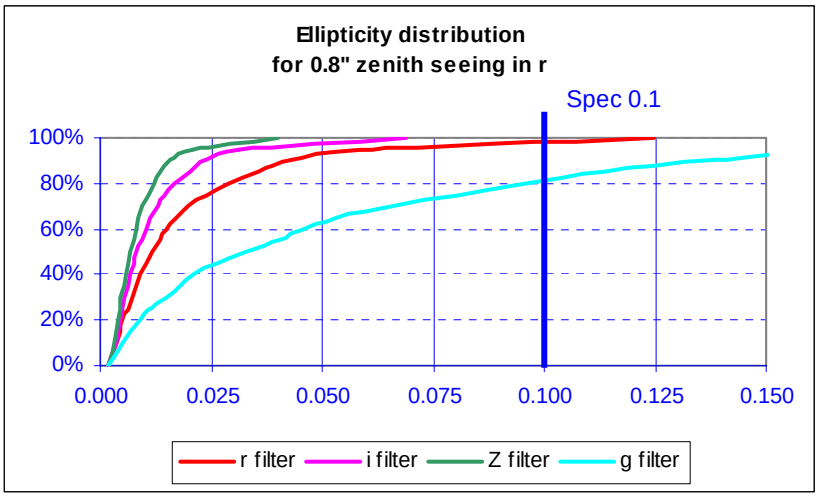


Figure 3.3.4-2 Ellipticity distribution for 0.8" zenith seeing in r

3.3.5 Optical efficiency

The optical efficiency is the product of the detector quantum efficiency times the percentage of transmission of the refractive optics and times the percentage of reflection of the mirrors. It is derived from the magnitude depth requirement in the science requirement document. The LSST exposure time duration versus optical efficiency is plotted on figure xxx for the filters g and r. These results were obtained using the LSST Exposure time calculator to detect a point source of magnitude 24 with a signal to noise ratio of 10. For this plot, we assumed a large effective aperture diameter of 7 meter and a median seeing of 0.7" to derive minimum requirements. The initial choice of 10sec exposure would have required an optical efficiency better than 0.85 in r to detect a 24th magnitude point source with a signal to noise ratio of 10. An optical efficiency better than 0.6 would be required in r for an exposure time below 20sec. The final choice of exposure times will be studied in more details during the R&D phase.

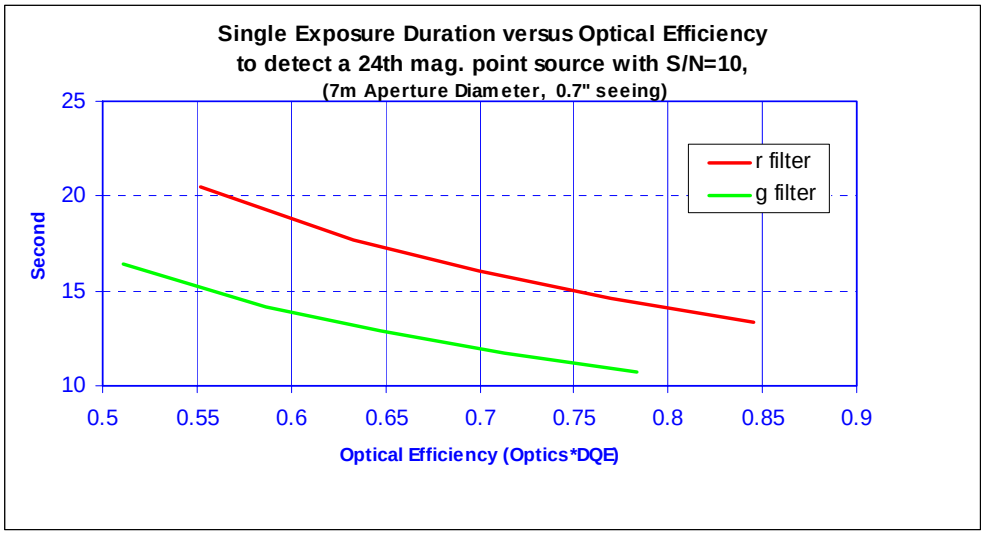


Figure 3.3.5-1 Single Exposure Duration versus Optical Efficiency to detect a 24th mag. Point source with S/N=10

An initial optical efficiency budget has also been computed using achieved reflection and transmission performances from previous projects, and using the allowable specification for the detector quantum efficiency. Based on results from the Gemini project, the mirror coatings were assumed to be protected silver. An optical efficiency better than 0.6 is achievable in the 600-800 nm wavelength band pass.

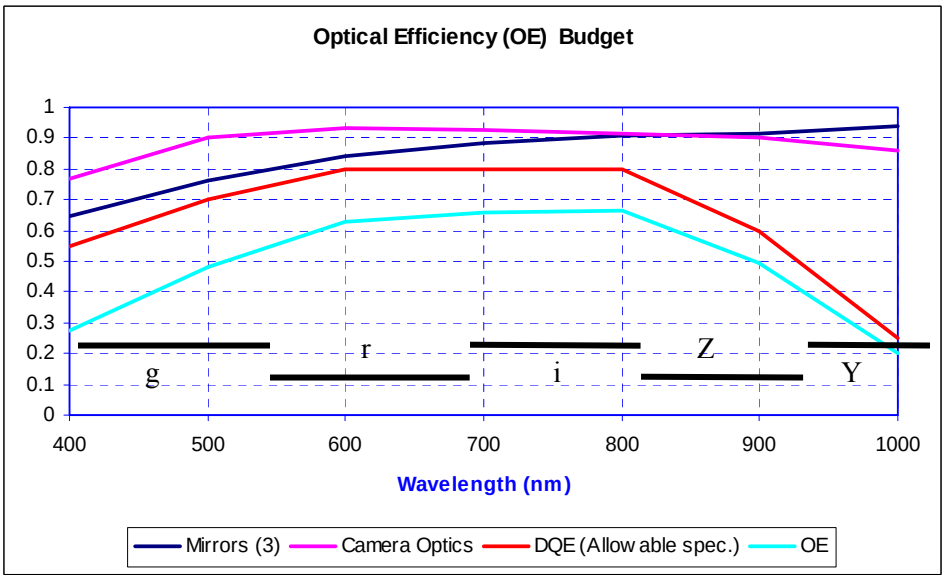


Figure 3.3.5-2 Optical Efficiency (OE) Budget

4 Current Design

4.1 System Engineering and Integrated Design

4.1.1 Trade Studies

The complexity of the LSST requires a close attention on systems engineering. The distributed nature of the LSST project further demands that close coordination be maintained among the project participants and development groups. The approach to addressing specific systems issues in this early stage of development has been through small focused, cross subsystem, and appropriate discipline teams assembled with a specific charge and schedule. Two such working groups that have completed their investigations have addressed the Optical Design and the question of an Atmospheric Dispersion Corrector.

4.1.1.1 Optical Design

In early 2004 the LSST assembled the Optical Design Tiger Team (ODTT) to establish a revised baseline design. The science requirements had matured to reveal a desire for a greater Etendue and project optical designers developed refinements opening a solution space that indicated advantages for reduced surface complexity and shorter overall design length. The ODTT was established to investigate the current requirements, potential solution space, and critical subsystem parameters that have particular impact by the baseline optical design. This team was to charged to recommend the baseline configuration balancing scientific return and technical feasibility.

To organize the vast design space a matrix of optical design families was established as shown in Figure 4.1.1-1. The columns represent designs of increasing fields of view and the rows are different families of optical optimized on specific parameters. The field of view was driven by the desire for information throughput and required number of visits to satisfy the weak lensing science requirement for a minimum number of field visits per filter over the survey length. The operations simulator described in Section 4.1.3 was used to indicate that 3.0 degree FOV could only marginally achieve the requirement with the most optimistic system performance values. With realistic performance parameters for the LSST system, the 3 degree FOV failed to meet our scientific goals. The 3.5 degree FOV was determined to meet the scientific requirement with reasonable system performance values and the 4.0 degree field of view was only more. Physical and engineering evaluations determined the 4.0 degree FOV was on the threshold of becoming very high risk and very expensive.

The families of designs focused on two parameters; vignetting and overall design compactness. Achieving constant illumination put a high degree of asphericity in the optical elements that was particularly difficult for the secondary mirror. A compact design was identified as very advantageous to maintain stiffness in the design to support the highly agile nature of this telescope and the shorter configuration greatly simplified the support of the camera and access for maintenance. Since both of these types of optical designs tended to increase secondary mirror aspherisity specific investigations were performed with fabricators to parameterize secondary mirror size, asphericity, accuracy and cost. The constant illumination design family was very advantageous in system throughput compared to the 62% average throughput for the minimal ashericity family but the necessary aspheric departure for the secondary mirror was substantially higher and outside several step function increases in manufacturing difficulty and cost.

Following the development of the optimized designs and completion of the many investigations into critical parameters the final choice was the 3.5 degree low asphericity family of designs. All fourteen of the panel members were unanimous in the choice. The rigorous review and evaluation established a new baseline adopted for the LSST project. Optical designers will continue to refine the design for final optimization but the approach to selecting the design proved very effective and advantageous to the project.

FOV / Design	3.0 deg.	3.5 deg.	4.0 deg.
Constant Illumination (short)	253 (0.65)	345 (0.64)	450 (0.64)
Low Asphericity (short)	230 (0.55)	**318 (0.52)**	393 (0.50)
Low Asphericity (long)	265 (0.61)	301 (0.53)	366 (0.45)

Figure 4.1.1-1 Optical Design Matrix

4.1.1.2 Atmospheric Dispersion Corrector

A very critical consideration in the optical design and the camera design would be the requirement for an atmospheric dispersion corrector (ADC). An ADC is a set of optical elements that adjust as a function of the telescope pointing angle to correct the chromatic displacement of light on the detector as a function of the air mass the telescope is observing through. To address the issues of an ADC for the LSST, a working group was established to evaluate the science requirements, address engineering impact, and recommend if the baseline design should include an ADC. To make this evaluation the group addressed the following primary questions:

a) What is the air mass range the LSST will observe through in addressing each science mission?

   The first effort to address this question focused on the total area of sky available to the telescope. Figure 4.1.1-2 shows the results of the purely geometric evaluation of available sky as a function of allowable zenith angle, latitude location, and levels of sky exclusion around the galactic plane. The two families of curves are, solid = no galactic exclusion, and dashed = +/- 20 degree galactic exclusion. The quantitative values where the galactic exclusion data meets the 20,000 square degrees goal and 15,000 square degree requirements are indicated. Also, Table 4.1.1-1 shows the full set of values for limiting

zenith angle necessary to meet the sky coverage for sets of latitude placement of the observatory and 3 levels of galactic exclusion.

The fundamental assessment of available sky shows that there is sufficient sky that can be accessed from the current observatory locations of interest at 28 to 31 degrees latitude. The observing simulator (discussed in Section 4.1.3) was then used to determine if the LSST can sufficiently utilize the time and capture the sky given the overhead of slewing, filter changes, and readout. The results showed that the LSST can meet the 20,000 degree goal in five filters in the 10 year life of the study by limiting the zenith angle to 48 degrees (air mass =1.5). This assessment is based on proper summing of the sky only for the fields that meet other weak lensing criteria and filling in the cadence with observations to address the other LSST science missions.

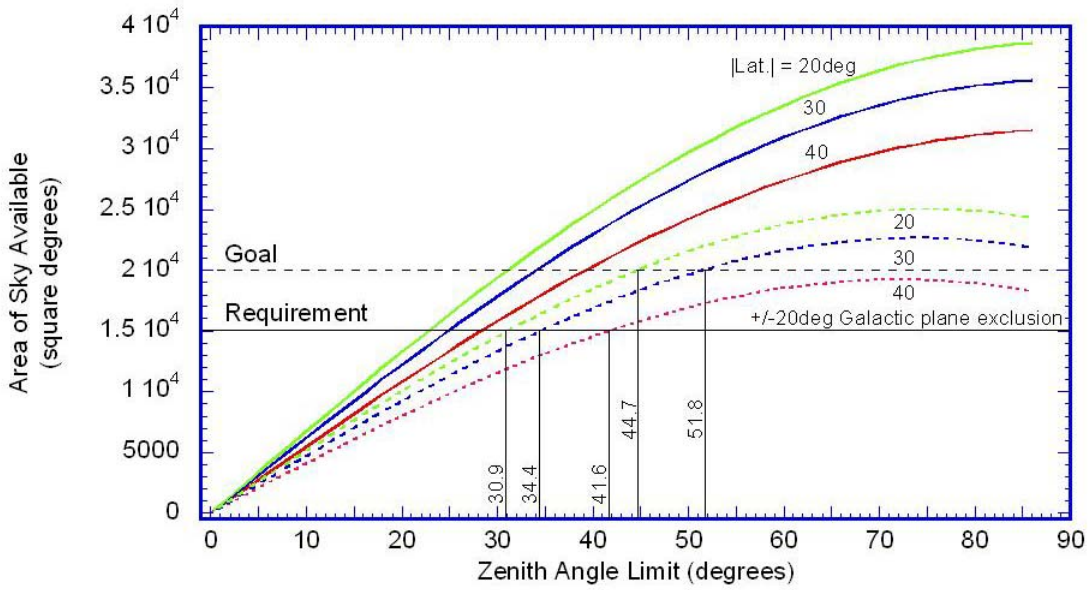


Figure 4.1.1-2 Sky Availability

Table 4.1.1-1 Necessary Zenith Angle to Observe to Reach Levels of Sky coverage versus Latitude Position of the Observatory and levels of Sky Exclusion Surrounding Galactic Plane.

Observatory Latitude (N/S)	**Galactic Latitude Exclusion**		
	0°	**±15°**	**±20°**
Meets Requirements			
20°	22.8°	28.1°	30.9°
30°	24.9°	30.8°	34.4°
40°	28.4°	36.1°	41.6°
Meets Goal			
20°	31.1°	39.3°	44.7°
30°	34.1°	43.7°	51.8°
40°	39.3°	53.7°	N/A

b) What is the impact with and without an ADC for the observation range of interest (Weak lensing, Optical Transients, Solar Science, Galactic Science)?

There are two fundamental conflicting results from use or omission of an ADC. First there is the impact on fundamental best image quality. Even with zero correction in the ADC the fundamental achievable image quality reduces with the addition of the element. Even calibrating our a substantial portion of induced error leaves a ? % impact on the image quality. However without the ADC there will be elongation of the image as chromatic dispersion is aligned with field angle on the sky. Narrow filter bands can limit this affect and a significant portion of the affect can be removed analytically however the demanding ellipticity limit to support the weak lensing program required careful consideration of the affect. Figure 4.1.1-3 below shows the ellipticity as a function of zenith angle for five different filters.

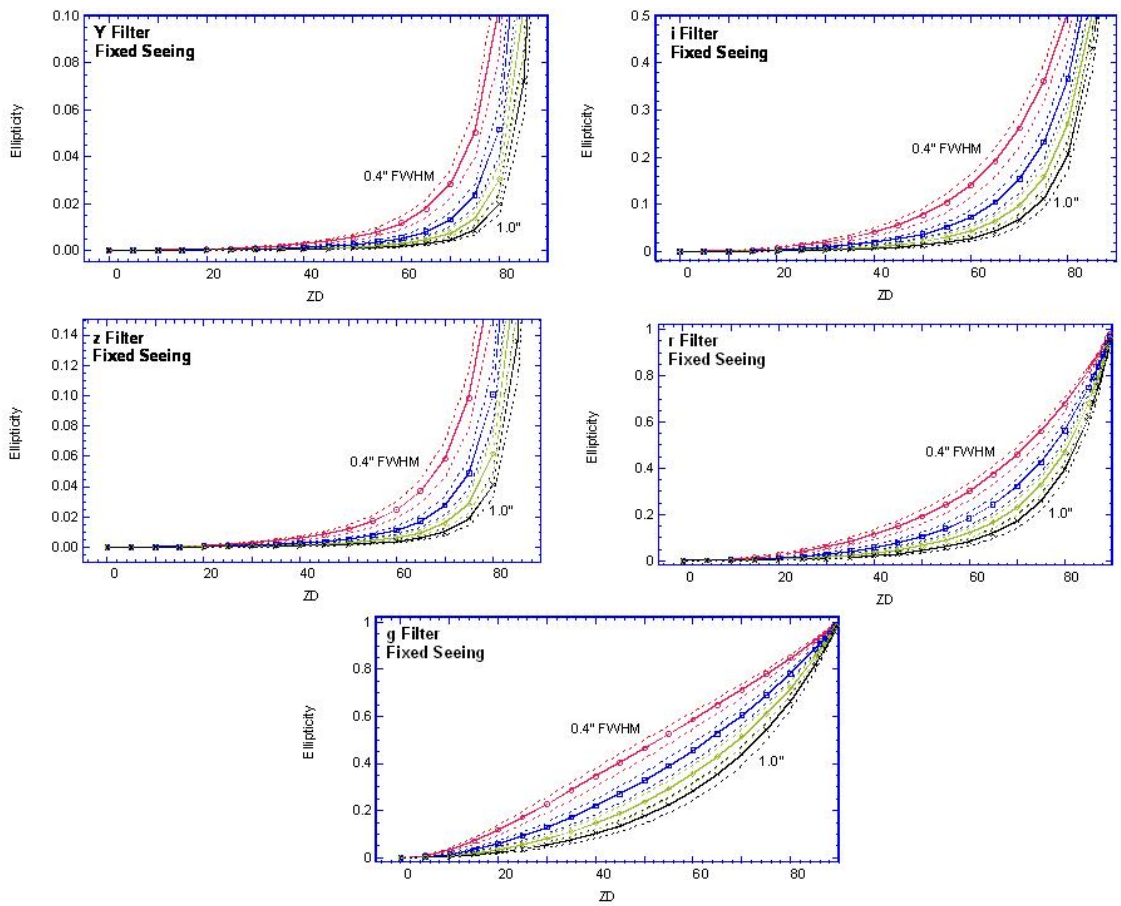


Figure 4.1.1-3 The dependence of zenith distance (ZD) on atmospheric dispersion induced ellipticity for each of the LSST filters griz Y in 0.4, 0.6, 0.8 and 1.0 arcsecond seeing (as measured at zenith). For each case three models were computed for altitudes of 1000m (upper dotted lines), 2000m (solid lines) and 3000m (lower dotted lines).

These curves define the input ellipticity contributed by atmospheric dispersion which will be present in the data. The analytic tools for compensating all images to extract all systematic ellipticity by using point source stars with the filed is an ever improving techniques as described in Section 2. The current state of these algorithms was tested to assess the range of correction that can be accommodated yielding an input ellipticity requirement of 0.1 as acceptable to achieve the final corrected goal. An error budget considering all sources of elliptical distortion was constructed to allocate this parameter determining that there was no need for an ADC to support the weak lensing science. In fact, the additional systematic errors that result from the additional element and the impact of the tolerance on its known position drive the desire to avoid an ADC.

Each science case and its observation technique and analytic method was investigated yielding similar result. There was only limited justification for and ADC and the technical implication of its inclusion in the design was very undesirable. The image differencing technique was the only area where dispersion correction had benefits for which the analytic tools only marginally compensated. These differencing tools are currently under development and even the current capability met requirements. Figure 4.1.1-4 shows that image differencing without an ADC can still be successful.

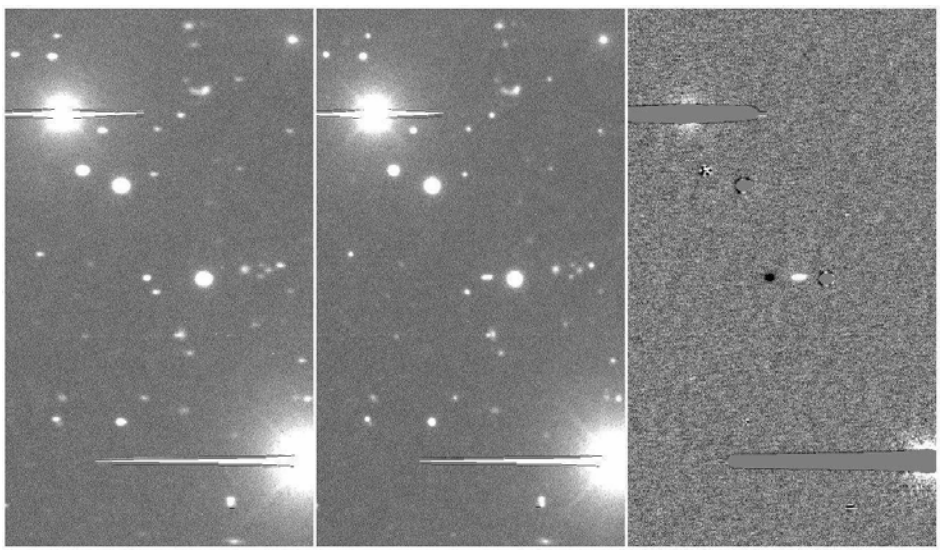


Figure 4.1.1-4 Successful image difference (right) showing no systematic effects between images with ADC tracking active (left) and ADC inactive by nulling the correction (middle). Note the asteroid moving between exposures as indicated by a positive-negative signature in the difference image.

c) What is the current state of the art for optics necessary to deploy an ADC?

Several techniques have been used to deploy an ADC. The concept developed for the LSST is a set of edged lenses that differentially correct the light path as a function of color by the amount of rotation difference in the two elements. Three aspects of the mechanical solution were investigated to assess feasibility. The first is the general optical design and the availability of appropriate glass types at the sizes desired. A review of glass types from providers worldwide yields the pair PSK3 and LLF1 as available and properly matched. For the ADC the glasses should have identical indices of refraction at the central wavelength but different dispersion values (differing change in the index of

refraction as function of wavelength). The second investigation was the fabrication of the lenses and packaging them into a functioning set of wedges. Figure 4.1.1-6 and Figure 4.1.1-7 below show the conceptual design for the wedge set and the mechanical concept for mounting the two optics that allow them to rotate with respect to one another.

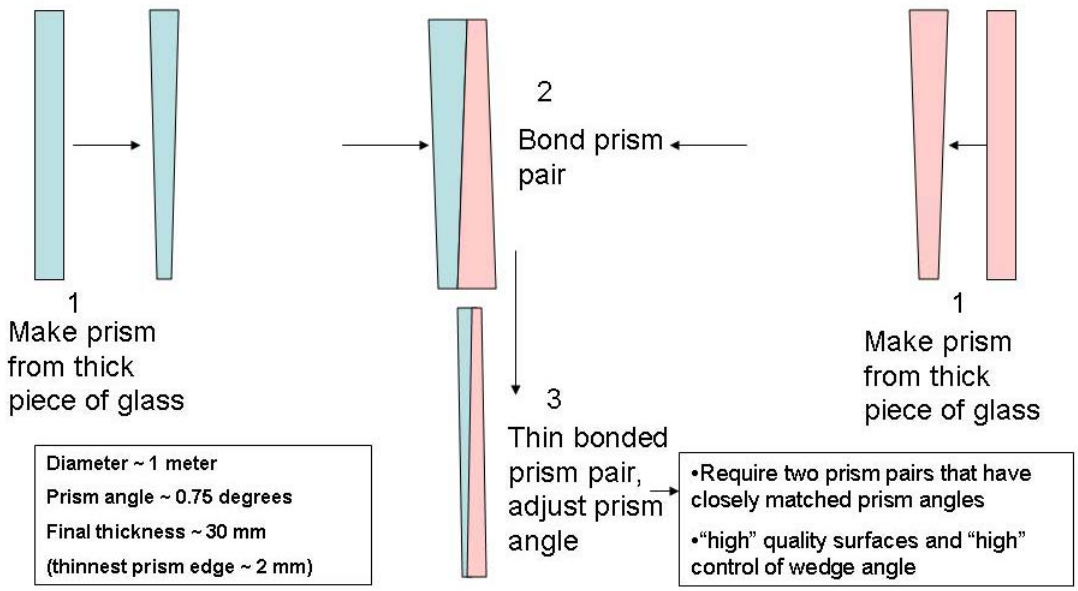


Figure 4.1.1-5 Fabrication process for the two wedges of the conceptual ADC

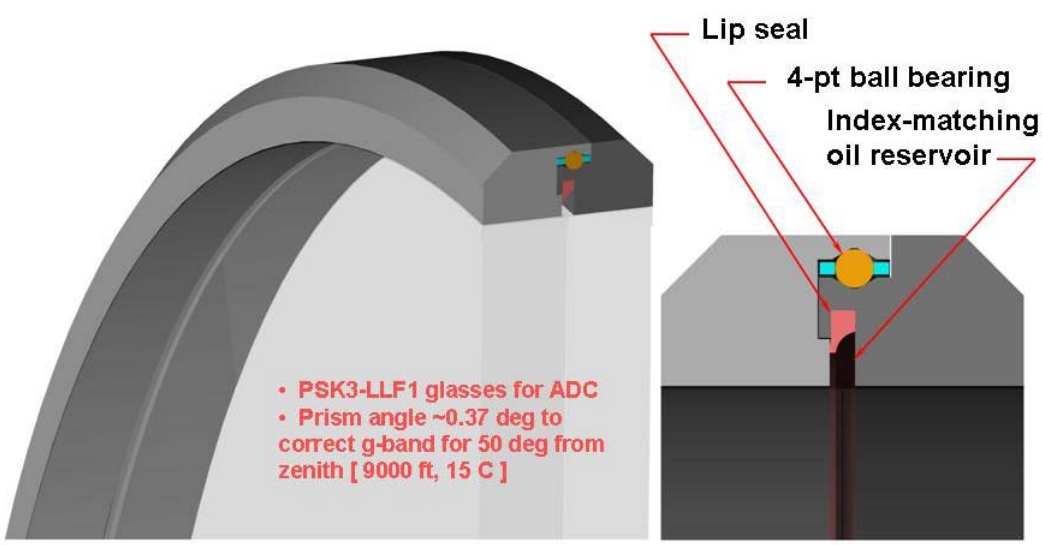


Figure 4.1.1-6 Conceptual Mechanical Design of the LSST ADC

A third investigation of the feasibility and implementation assessment was to determine the impact to the camera geometrical concept in supporting the ADC. A layout was developed that showed the ADC could be supported between Lens 2 and Lens 3 in front of the filters and shutter (as illustrated in Figure 4.1.1-7), and that there was still sufficient space for the currently conceived filter changing mechanism to move the filters in and out for change-out.

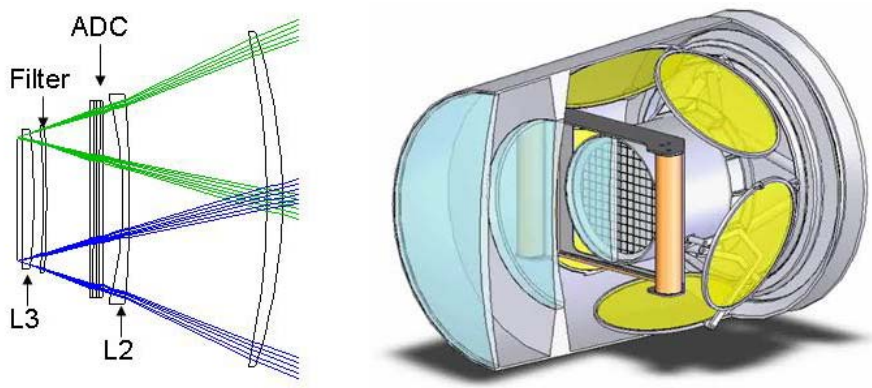


Figure 4.1.1-7 An illustration of the possible location of an ADC within the camera body.

d) Should the LSST Baseline include an ADC?

Following a probe of the science, the technology and the analytic tools for calibration and data manipulation it was decided that no ADC would be included in the baseline LSST design. The space remains available in the camera and a feasible concept exits but there is no justification for the complexity and cost of this device. The LSST filter band passes are sufficiently narrow and sufficient sky can be accessed within the 1.5 air mass limit that LSST does not need the ADC and sufficient qualitative impact on the system makes an ADC undesirable. Figure 4.1.1-8 below summarizes the ellipticity per filter as a function of zenith angle in the presence of real seeing values. Altitude variations are also plotted. The current LSST sites of interest lie with the 2000 to 3000 meter elevation.

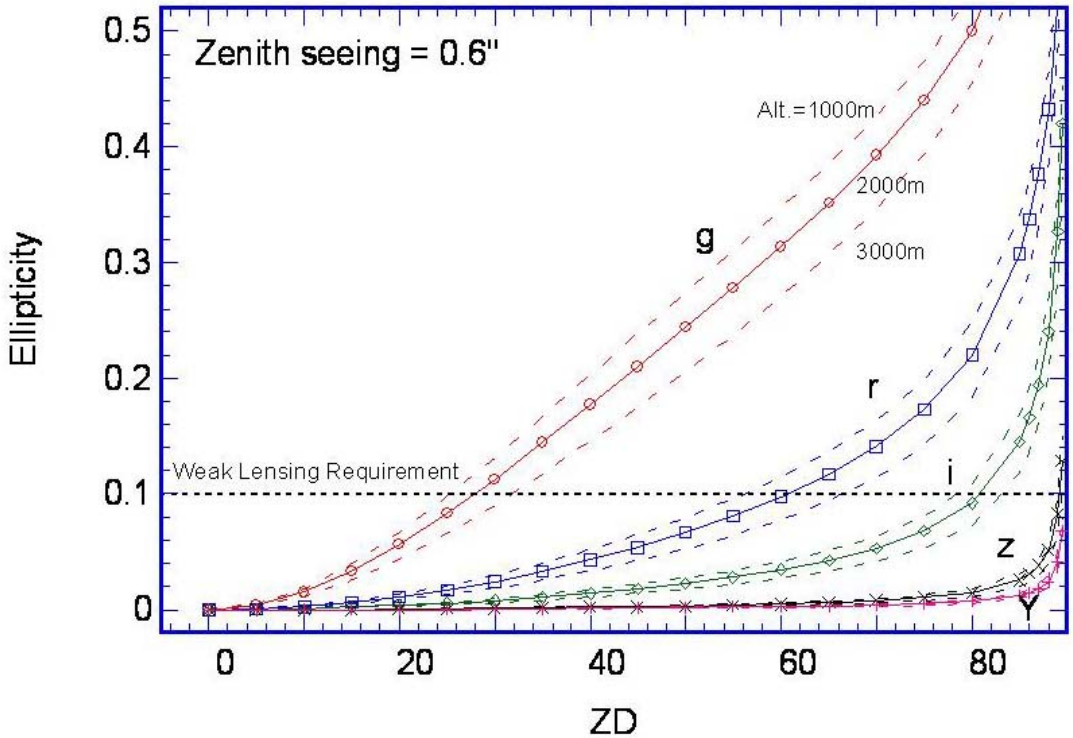


Figure 4.1.1-8 Zenith dependence of differential atmospheric dispersion induced ellipticity in grzY spectral bands (see text).

4.1.1.3 Conclusion

System Engineering is a high priority within the LSST project. As the project can staff up its central staff the systems issues can increasingly be addressed there. As the team develops the central issues impacting the overall system and science will continue to be addressed by specifically assembled and focused teams. In addition to the Optical Design and ADC groups, other working groups continue their development work. The current status of these efforts is represented in dedicated subsections that follow. The diverse LSST membership has shown the necessary level of cooperation to make working groups successful and will continue to do so throughout the project development.

4.1.2 Optical Simulation

4.1.2.1 Technical Overview

The optical simulations are a distinct subgroup within LSST project. This subgroup ultimately will have responsibility for modeling the sky, atmosphere and instrument. Products from this group have two main applications. The first is to demonstrate that the instrument baseline design will meet the science goals. The question here is whether a ground-based 8.4 m instrument with wide field-of-view and 15 second integration times can deliver sufficiently good image quality to allow weak-lensing science to high red shift. Therefore, the first product of the simulation group has been a science validation of the baseline LSST design from detailed astrophysics, atmosphere and instrument simulation. The second product is aiding the development of other LSST

subsystems. These simulations will ultimately include details down to the level of, e.g., diffusion in the CCD sensors and feedback noise in the instrument control system. This simulation therefore allows the Camera, Telescope and Data subsystem groups to refine and validate their designs. An increasingly important LSST subsystem is the "data pipeline", shorthand for all aspects of the data handling. LSST produces 10's of petabytes of data per year. These data are indexed in a database and made accessible to users around the world. Further, the real-time data pipelines include physics and astronomy analyses. Examples of these analyses include near Earth objects alerts, and reconstruction of seeing to allow exposure-to-exposure filter selection and pointing direction. At a low level, the pipeline subsystem needs large volumes of data in order to exercise the data handling and module-to-module interconnectivity. At a higher level, the pipeline needs high-fidelity data in order to allow development of real-time science algorithms and instrument operations models. The simulation group has overall responsibility for developing the simulation chain and generating simulated data sets.

The main technical components of the optical simulation are (1) generating images on the sky (galaxies, stars, etc); (2) propagating the image wave front through the atmosphere; (3) ray tracing the image through the LSST instrument optics model; (4) connecting the optical simulation to the "operations simulator" (this later simulates where to point the instrument, the choice of filters, when to make return visits to some direction, etc.); (5) programming to provide support and continuity to the project, librarian activity for the simulated data files, and creating the user interface and platform for the various simulation components; (6) validation of the simulation with existing instrument data. Approximately, each of the above corresponds to a significant part of a single person. Figure 4.1.2-1 shows the main way these various components fit together. In fact, the details of the interdependencies are more complicated.

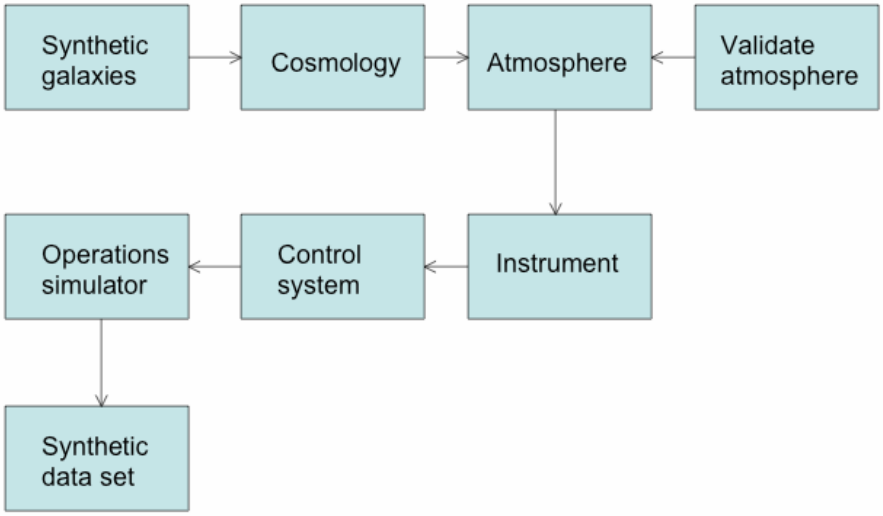


Figure 4.1.2-1 Components of the LSST simulation chain. The Operations Simulator is not directly part of Optical simulation.

Some of this kind of simulation has been done for earlier telescopes. However, in no case has such a comprehensive simulation chain been developed prior to instrument turn-on, and in no case has the target simulated data set been so extensive, covering the full range of physics over the full range of seeing conditions over all the availably sky. The nearest parallel is that of the large high-energy physics detectors, which include such simulations as a major component of the detector construction project. Such simulations ensure the instrument design is sensible in light

of the science goals, and the data pipelines and data analysis chains are immediately productive. There are three key technical challenges facing the LSST simulation project. The first challenge is that the simulated data sets are large, take a long time to generate and require a large amount of data storage. The second challenge is that while earlier telescope projects included some aspects of simulations in a labor-intensive way, the LSST simulations need a considerably higher level of automation. The third and perhaps the most open-ended challenge is that the largest uncorrected distortion of the images is likely due to the atmosphere. The effect of the atmosphere is not yet measured for the LSST field of view, exposure time and aperture. Although we can make a reasonable guess as to an appropriate atmospheric turbulence model, there will eventually have to be a comparison of the model with data for from an existing large-aperture telescope equipped with atmospheric monitoring instruments.

4.1.2.2 Simulation Project Structure

The simulation starts with generating images of the sky. One class of these images are synthetic, for example: (1) arrays of stars over the field of view (to study PSFs); (2) galaxies and stars (and transient or high proper motion objects) generated according to a phenomenological model; (3) galaxies and stars generated from N-body simulation. Right now, we are regularly generating star arrays to study the PSF; this capability is ready immediately. We have started generating images of galaxies and stars according to a phenomenological distribution, and then applying a cosmological model to the light propagation. This has been done before in support of the Deep Lens Survey, and its adaptation to LSST is relatively minor, consisting mostly of modifying the image-creation chain to make it less user-intensive. No one within the simulations group has yet placed transient or high proper-motion objects in images. We are also collaborating in the development of supercomputer studies of cosmological N-body simulations. These synthetic images will then be input to the simulation chain, with goal of comparing the reconstructed cosmology with the cosmology hypothesis.

A second class of input to the simulations are actual images. We have started with Hubble Deep Field images. We are also processing SuperMACHO data for the simulation pipeline. Cataloging these various input files will be the responsibility of a programmer/librarian.

The images are then propagated through the atmosphere according to some model. At present, the simulation group has two atmospheric modeling efforts. The first involves the Arroyo package (a Thirty Meter Telescope project package by Matthew Britton/CalTech). Arroyo was chosen because it has considerable flexibility, is well supported now and will likely to supported in the future, its output has been shown to be sensible when compared to data, and is able to directly use as input the processed atmospheric data from Cerro Tololo. This package has several options for the light propagation model through the atmosphere. The highest fidelity one is based on a Fourier transform of the wavefront phases. This includes refractive and diffractive atmospheric effects. The drawback of this approach is the long computing time and large memory requirements. Consider a atmospheric layer 6 km above ground, moving at 10's of km/hour, with turbulence cell size varying from a few centimeters to 10's of meters, and a three degree across field of view and 10 second exposure. This atmospheric layer alone (a "screen"), for full fidelity will require hundreds of gigabytes of storage and minutes of computing time. The LSST/Arroyo package has been ported to 32 bit processors, and we are in the process of parallelizing processing among the CPUs in a cluster. This package has processed arrays of stars and is just starting to process galaxy and star images. Figure 4.1.2-2 shows a sequence of 10 0.1 second images at the LSST image plane from a single point object, using a realistic Cerro Tololo atmosphere. Figure 4.1.2-3 shows the corresponding amplitude and phases from a single Figure 4.1.2-1 integration.

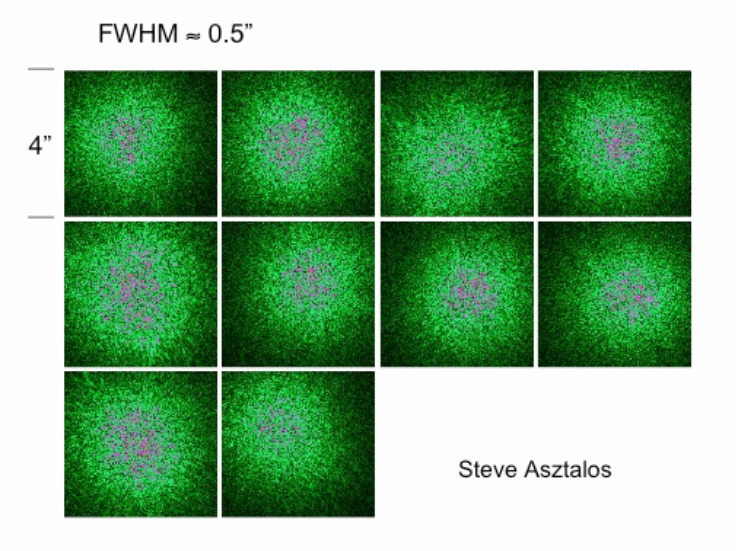


Figure 4.1.2-2 Arroryo predicted PSFs at the LSST image plane for a realistic Cerro Tololo atmosphere. Each image is a sequential 0.1 second integration. (Asztalos).

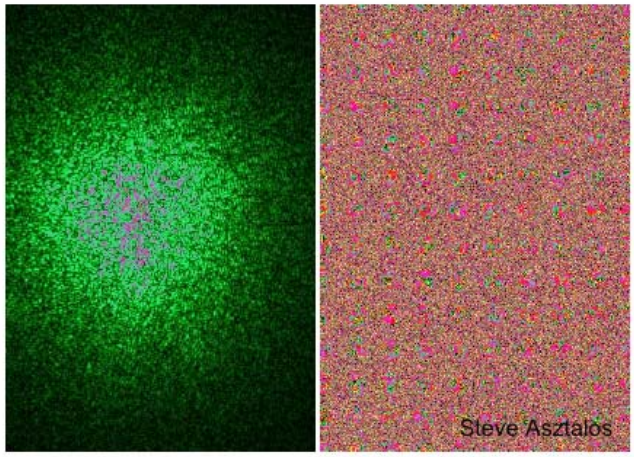


Figure 4.1.2-3 Amplitude (left) and phases (right) at the LSST image plane from a point object.

The second atmospheric modeling effort uses an ab initio code of a several layer Kolmagorov turbulence model and ray traces through the layers. By making a simplifying assumption that the diffractive part of the propagation can be ignored, this code runs extremely fast. Though, of course, the need for very large data storage for atmospheric screens is still required. Its fast execution has made this tool the most used by the simulation group so far. It has processed arrays of stars and images. The output still needs validation, and the assumption of no diffraction needs scrutiny. Right now, the Arroyo and ab initio packages are being cross compared. Likely within

a year, the two simulations approaches will be mature enough to be evaluated and there will be a down-select to one package.

The atmospheric simulation generates the wave front at the telescope image plane. This wave front is then ray-traced through an instrument model. This model is ab initio code developed specifically for LSST. This code executes very fast, tens of millions of rays per hour. It also includes considerable detail about the telescope model with hooks for more. Figure 4.1.2-4 shows a network of rays in the instrument model. At present, the three mirrors are in the model, as are very preliminary mirror perturbations from wind shake and feedback control. Figure 4.1.2-5 shows an array of point objects before and after applying mirror perturbations in the ray tracing. Because the code is not a black box, essentially any telescope parameter can be reasonably accommodated. For instance, a preliminary CCD model was incorporated in less than a day, as was a simple spider geometry. Figure 4.1.2-6 shows a Hubble Deep Field image transported through the preliminary atmosphere and instrument models. There is some discussion among the collaboration on whether the ray tracing should be built from a commercial package. This would have the advantage of being supported and already validated. At present, we plan on validating the Peterson/Jerigan code on simple geometries. However, in principle, the existing code is as conceptually valid as a commercial product. Therefore, we plan to retain for now the present code and will defer to later in the year the question of whether to change to a commercial product.

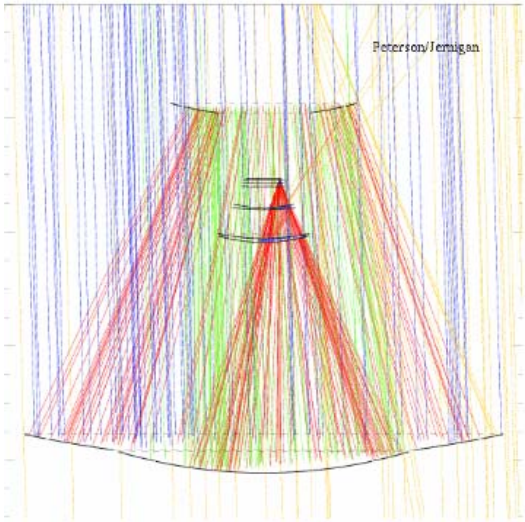


Figure 4.1.2-4 Ray tracing the LSST optics.

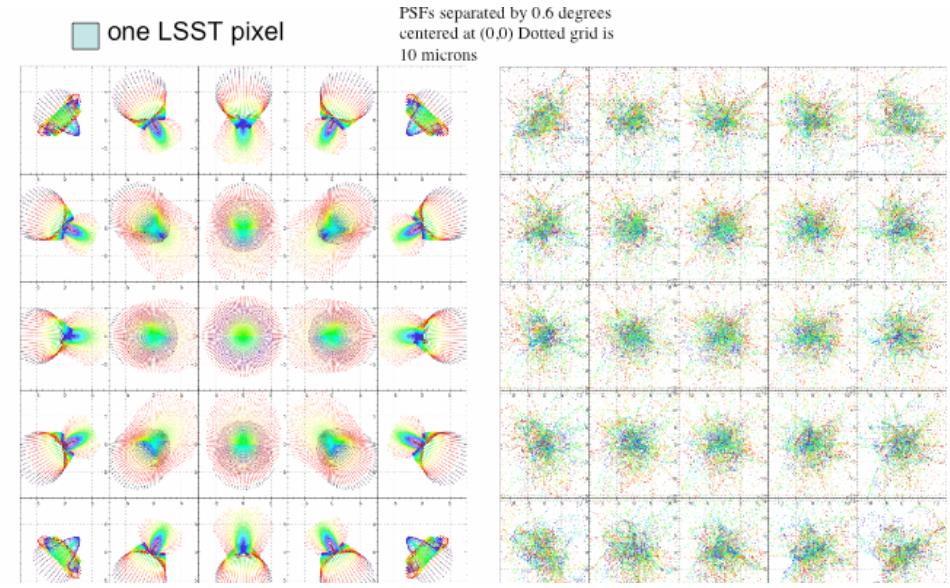


Figure 4.1.2-5 Array of points ray traced through LSST optics. Without mirror perturbations (left), with mirror perturbations (right).

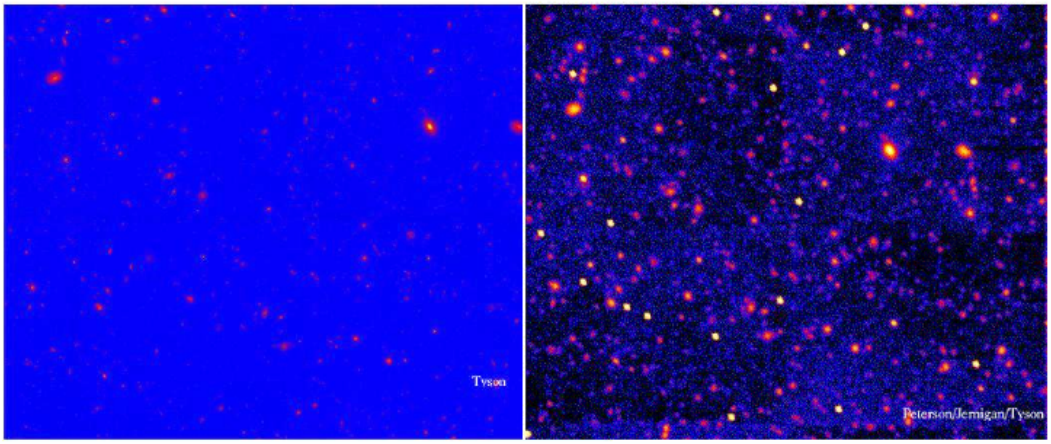


Figure 4.1.2-6 Hubble Deep Field N (left). HDF field propagated through atmosphere and ray traced (including mirror perturbations and wind shake).

4.1.3 Cadence Simulations

With its unprecedented combination of collecting area and field of view, the LSST will image large areas of the sky frequently and to great depth. The cadence of these observations, the order in which different fields of view are observed in each color and the frequency with which they are revisited, will determine just how much sky will be covered, to what depth, and with what temporal sampling. This in turn will determine how useful LSST data will be to different investigations.

CURRENT DESIGN

We have developed an operations simulator to answer the question, can the proposed telescope and site(s) deliver the required science. There are distinct science goals for LSST which encompass the need for large sky area and well as the need for various time samplings in the temporal domain. The weak lensing survey, the Solar System survey, the Galactic science case and the transient universe (and in particular supernovae) are all science programs which have different spatial and temporal sampling requirements. We are using the operations simulator to investigate observing cadences and strategies. We are also investigating different telescope parameters ranging from etendue to filter change time. The operations simulator is also being used to determine if proposed sites for LSST will allow the needed science goals to be achieved by simulating years of operation with historical weather and seeing data.

The current operations simulator is based upon an open source simulation language, SimPy. This provides the infrastructure for event based activities and time keeping. The design is highly modular, with separate science programs being described in separate python modules. There is a sophisticated telescope module with all motions parametrized for ease of testing different telescope capabilities, eg effect of acceleration capabilities of various motors on science output. We use the Krisciunas and Schaeffer (1991) sky brightness module and various SlaLib routines to track the sun, moon and planets.

All important parameters for the telescope, the site and the science programs are easily accessible in configuration files. Each science proposal ranks potential observations based upon the internal logic of the proposal. The current method of ranking potential observations between different science programs is currently a simple, linear weighting of the internal rank of an observation within a science proposal.

We have generated weather and seeing data for four sites which were considered possible LSST sites. The weather data is derived from satellite observations for cloud data and local seeing data, primarily from DIMMs. The DIMM coverages were not particularly complete for any of the sites, so continuous seeing data was generated by creating seeing data which matched the power spectrum of the available, real data.

When the simulator is run, details from simulated observations are stored in an open source database, MySQL. This database is also used to store the weather and seeing input data. For each observation, we store 34 attributes including sky conditions, filter, seeing, airmass, the time it took the telescope to move from the previous pointing, the date and the position of the camera with respect to the telescope and the sky.

While simulations are ongoing, early simulation runs have investigated the impact of the field of view on the success of multiple science programs undertaken simultaneously. These simulations used real seeing and weather data from CTIO, a site which is not under consideration. The first three figures show the sky coverage for a *single* simulation with a 3.5 degree FOV where three science goals are being sought: 1) Weak lensing (WL) with a minimum of 15, 15, 15,25, 25 visits per field in g, r, i, z, and y, 2) Near Earth Asteroid survey where this survey is limited to +/- 10 degrees of the ecliptic and a complete search sequence has 3 sets of 2 visits per night in a lunation, and the nightly visits are separated by 30 minutes and the 3 sets are separated by 5 nights each, and 3) Super Nova survey where a super nova sequence requires a visit roughly every 3 days for 60 days and sampling in all filters. The first figure shows total visits per field, the second shows visits which could be used for the WL survey, and the third shows fields which had completed super nova or NEA sequences.

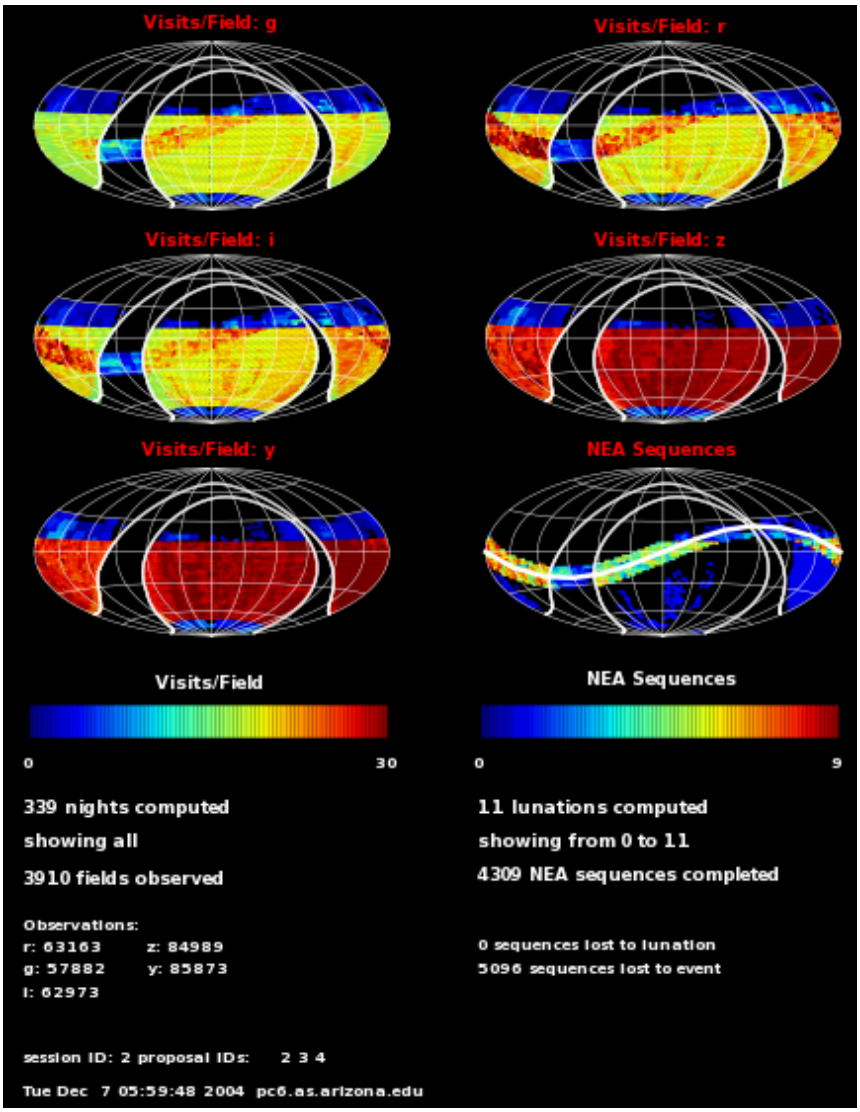


Figure 4.1.3-1 Three programs coverage of the sky from Cerro Pachon using real CTIO seeing corrected to Cerro Pachon.

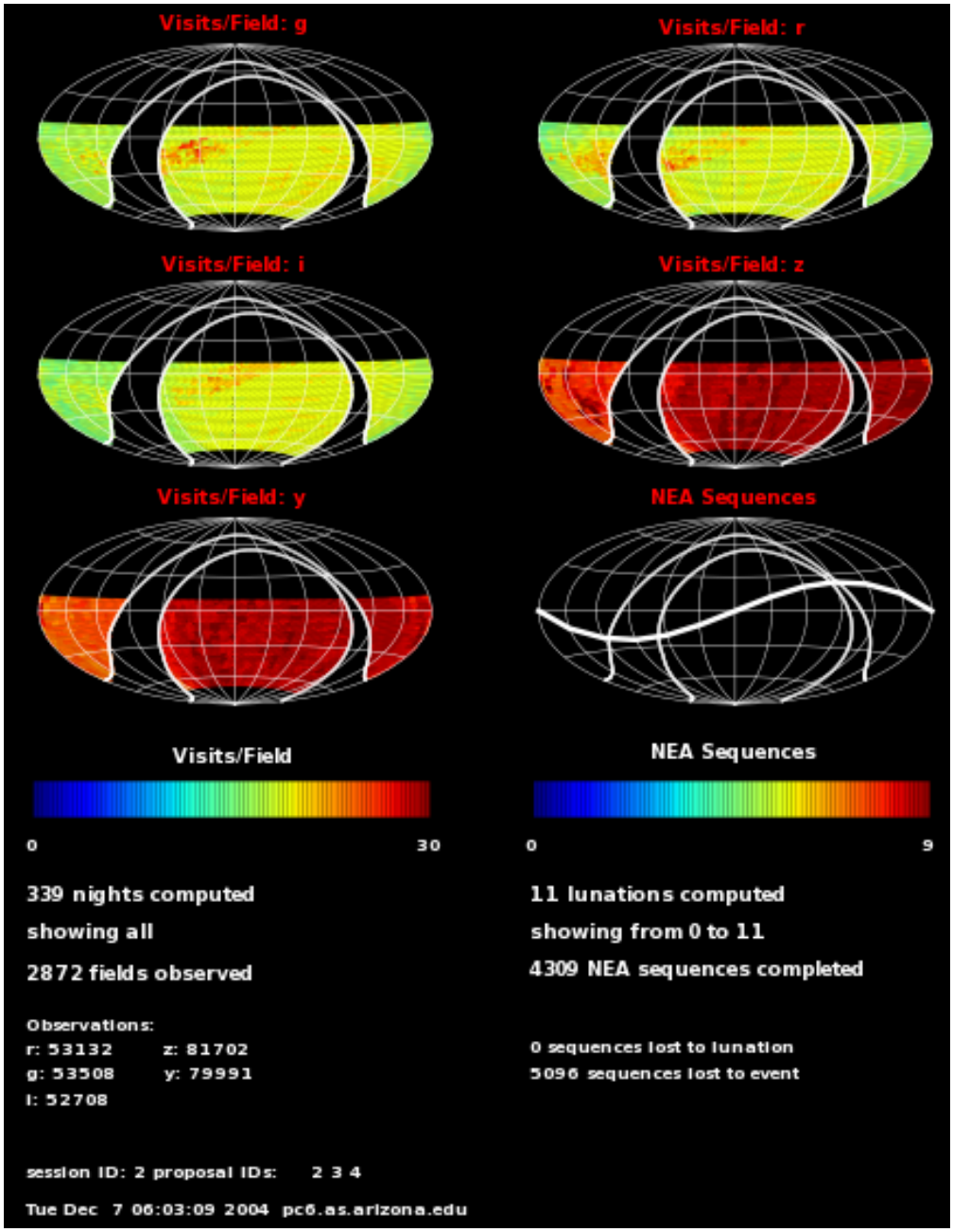


Figure 4.1.3-2 Weak lensing appropriate sky coverage for 339 days of CTIO weather from Cerro Pacho with SN/NEA.

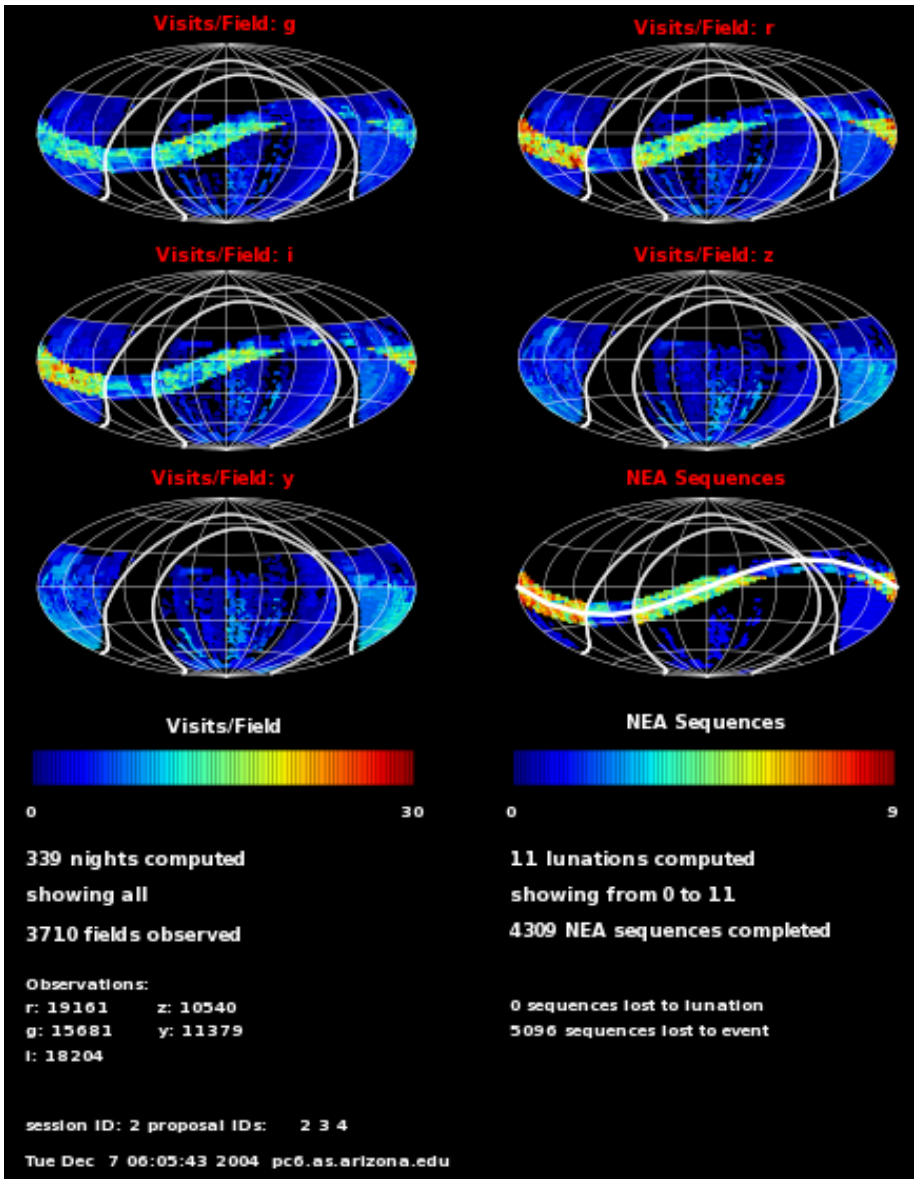


Figure 4.1.3-3 Transient fields with completed sequences observed (NEA and SN programs) for 339 days CTIO weather from Cerro Pachon with WL/SN/NEA programs running.

The next figures show the sampling achieved when running simulations for 3.0 and 3.5 degree FOV focal planes performing simultaneous weak lensing (WL), supernova (SN) and near earth asteroid (NEA) surveys using minute-by-minute, CTIO seeing and cloud data. The plots are histograms of the number of visits per field in each filter where that field has a minimum of the required visits in each filter. The visit set of 10, 10, 10, 15, 15 per year in g, r, i, z, y, is the minimum needed for the WL science. *A 2.5 degree FOV is not plotted because it does not achieve the needed sampling for any field for multiple science goals with the current weighting criteria.* There is surprisingly little confict between various cadence and visit requirements when the telescope has a large FOV.

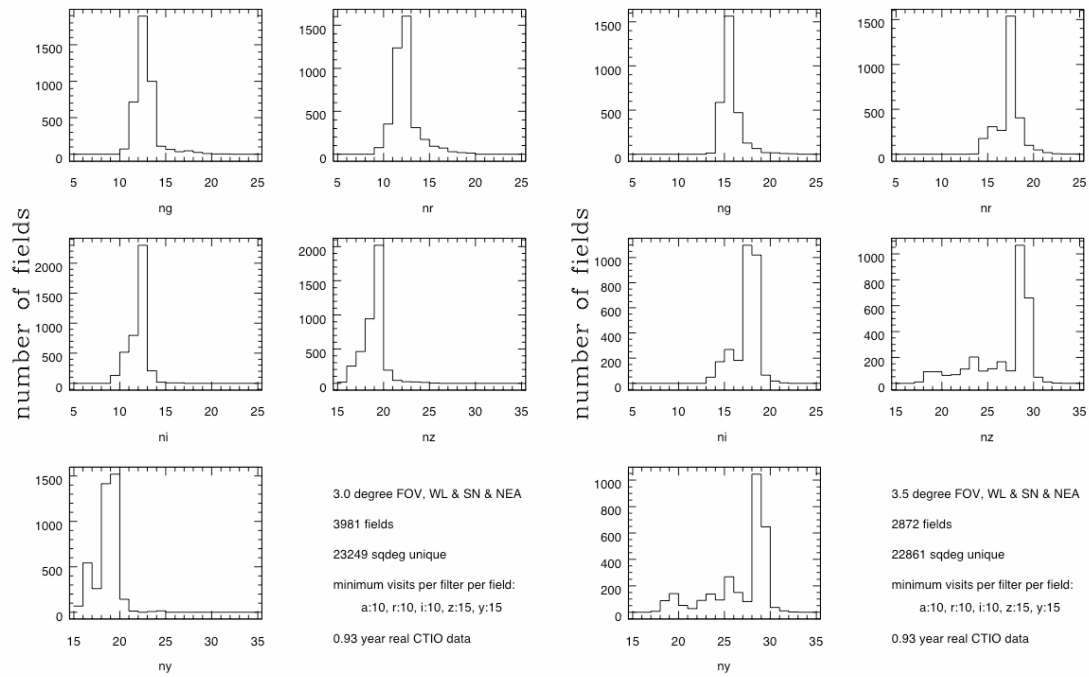


Figure 4.1.3-4 Distribution of visits per field in five filters for simulations using a 3.0 degree FOV (left panel) and 3.5 degree FOV (right panel). For multiple science programs, the 3.0 FOV barely meets (does not completely cover the accessible area) the WL visit requirement, while the 3.5 degree FOV covers the available area with visits to spare.

We are continuing to search for the optimum instrument and telescope parameters as well as develop more sophisticated science programs, The current implementation executes a year's simulation in 5-17 hours depending on the number science programs and their particular parameters. We currently have about two years of weather and seeing data for the three sites which LSST still has under active consideration.

4.1.4 End-to-End Simulations and Integrated Modeling

End-to-end simulations will serve two key purposes in the development of the LSST. First, we will use the top down end-to-end simulations discussed above to test and ensure that the LSST is able to achieve its science missions given reasonable parameterization of the system performance. Second, a bottom up simulator will be used as a design tool to guide the development of the conceptual and detail design of the LSST. Using the tools and techniques for integrated modeling this detailed bottom up end-to-end simulator will serve multiple purposes, including supporting the rationale for design trades and decisions, defining the interfaces and interactions between major LSST subsystems, post construction performance monitoring and analysis and refinement of the constituent input physics as discoveries are made in the LSST data.

4.1.4.1 Description of the Bottom-up Engineering-Science Simulations

When compared to the HEP community's modeling tools (*e.g.* CERN's GEANT4) the integrated modeling tools used for ground based telescope facilities are in its early stages of development. However, significant progress has been made recently in modeling large optical systems, largely driven by new technology space systems [ref] and the 30m and large ground based telescope designs.

Our approach and overall design to the detailed engineering-science simulation is illustrated in the functional block diagram shown in Figure 4.1.4-1 and based on the earlier work above. At the core of this simulator are three integrated models for each of the LSST major subsystems: the telescope, the camera and the data processing system. The inputs to the integrated models that represent the LSST realization are collected from the engineering and design descriptions (box on the left) and parsed to generate the specific "LSST Model" by the configuration and model generating tools. The model pre-processor is responsible for keeping all the various modeling elements synchronized with the current LSST design and configuration.

Within each integrated model are specific modules built up from standard engineering tools (indicated by the internal blue boxes) that most familiar to the people doing the design and engineering work. These include commercial software tools such as MSC NASTRAN for finite element analysis, OSLO or CODE V for optical ray tracing and where necessary special project specific developed tools. Information between each of these modules is communicated through and interface node (shown as a red dot) and is facilitated using a common backbone architecture (*e.g.* MATLAB/SIMULINK).

The LSST Model takes for science input an image representing a section of the night sky equal to the LSST field of view. The scene in the input image is defined by the intersection of the constituent input physics (indicated on the right side of Figure 4.1.4-1) and the place on the sky as generated by the Operation/Cadence Scheduler. The image is handed to the integrated telescope model which modifies it according to perturbations from the atmosphere, the telescope optics and control system and filter spectral pass band. This is the delivered image to the integrated focal plane array (FPA) model. Separate analysis of this image can be used to provided design and trade feedback as indicated by the pink dashed line in Figure 4.1.4-1.

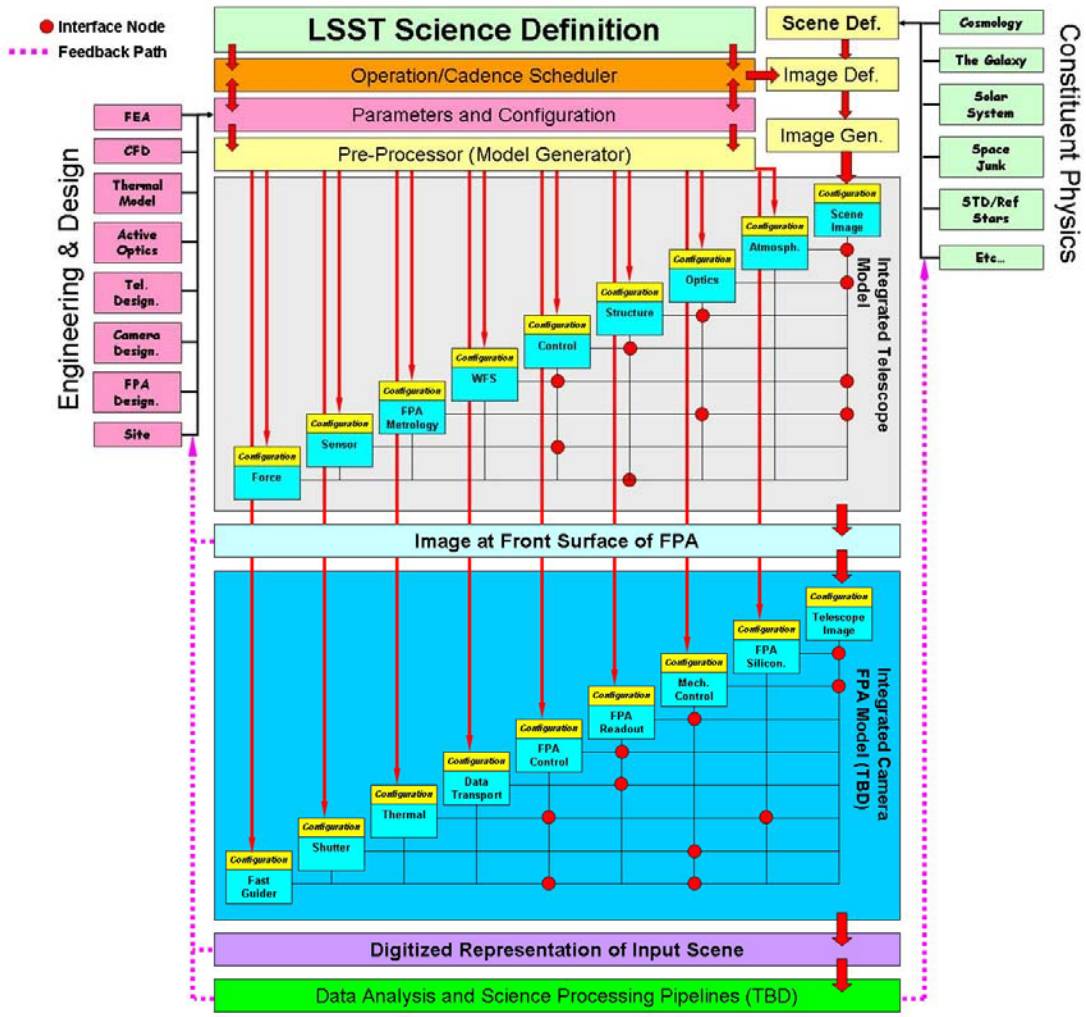


Figure 4.1.4-1: The function block diagram for the LSST end-to-end simulator.

The integrated FPA Model modifies the telescope delivered image to represent the processes internal to the "inner dewar". These include the effects of refraction in silicon, electron-hole pair production and diffusion, array readout, analog and digital signal conditioning, etc. The output of the integrated FPA model is a digital representation of the of the input night sky and constitutes what would typically be referred to as synthetic LSST data. Here again, these images can be analyzed separately to provide design and engineering feedback.

Ultimately, this process is repeated as driven by the Operation/Cadence Scheduler (see Section 4.1.3 for a detailed description) to generate a time series of synthetic images as would be produced by the LSST. The image time series image stack is handed off to the Data Analysis and Science Processing pipelines where the true test of the LSST is done. The Data Analysis and Processing Pipelines will test the effectiveness of algorithms in their ability to recover the relevant physical parameters that originally were used to define the LSST Model science input. This is perhaps the most essential and critical of all the feedback paths show in Figure 4.1.4-1 in demonstrating the effectiveness of the LSST system end-to-end.

Later, as the LSST project continues towards construction, commissioning and science operations this simulator will evolve in its fidelity and serve as an essential tool for validating and monitoring performance.

4.1.5 Optical Design

4.1.5.1 Telescope Design:

The baseline optical design for the LSST is a modified Paul-Baker 3-mirror telescope that includes a 5m tertiary mirror (M3) coplanar with an 8.4m primary mirror (M1). After a first reflection on M1, the optical beam converges on the 3.4m convex secondary mirror M2. From M2, the reflected beam diverges toward M3, and is then focused toward a 3-element correcting camera located in front of M2 on the optical axis (see Figure 3.1.5-1). The current design employs three aspheric mirrors, and two of the refractive elements in the camera have aspheric surfaces. The three-mirror telescope system delivers, without the camera corrector optics, a spherical wavefront on axis that will greatly help in initial assembly and alignment. The exact optical prescription is given in Table 4.1.5-1.

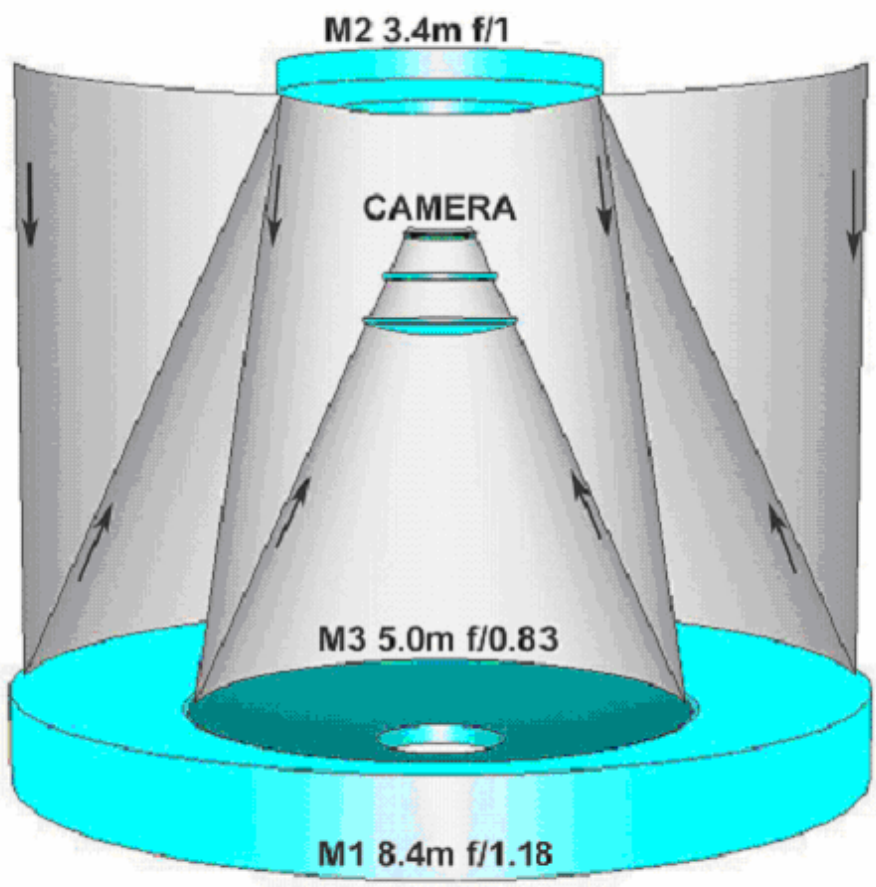


Figure 4.1.5-1: LSST baseline optical design

This optical design delivers a flat focal plane with a 3.5-degree diameter field of view (FOV) and an image spot size better than 0.2 arcsec FWHM for 5 spectral bands covering a wide wavelength bandpass (from 400 to 1030nm). In addition, the LSST is a very fast telescope (f/1.234) with a plate scale of 50microns/arcsec and a detector diameter of 0.64m. The 10-micron pixel size detectors match the superb performance in image quality.

Table 4.1.5-1 LSST baseline optical prescription for r band filter: all units are mm except as noted

Surface	Radius of curvature	Center thickness	Outer optical clear aperture semi-diameter	Inner optical clear aperture semi-diameter	Name	Aspheric departure from best-fit parabola (BFP) or over annulus
1		1345.5				
2		4810.7	4327	2412		
3	−19835	−6156.2	4180	2558	Primary	0.111
4	−6788	6390	1700	900	Secondary	0.017
5	−8344.5	−3631.2	2508	527	Tertiary	0.403
6	−2824	−82.2	775		L1	
7	−5021	−412.6	775		L1	
8		0				
9	–	−30	509		L2	
10	−2529	-357.5	509		L2	
11		0				
12	−5624	-17.7	380		Filter	
13	−5597	−43.3	380		Filter	
14		0				
15	−3169	−60	352		L3	
16	13360	−28.5	352		L3	
Image			317		Focal Plane	

Surface	Conic Constant	y^4	y^6	y^8	y^{10}
Primary	−1.215		1.38E-24		
Secondary	−0.222		-1.27E-20	-9.68E-28	
Tertiary	0.155		-4.50E-22	-8.15E-30	
L2	-1.57		1.65E-18		
L3	-0.962				

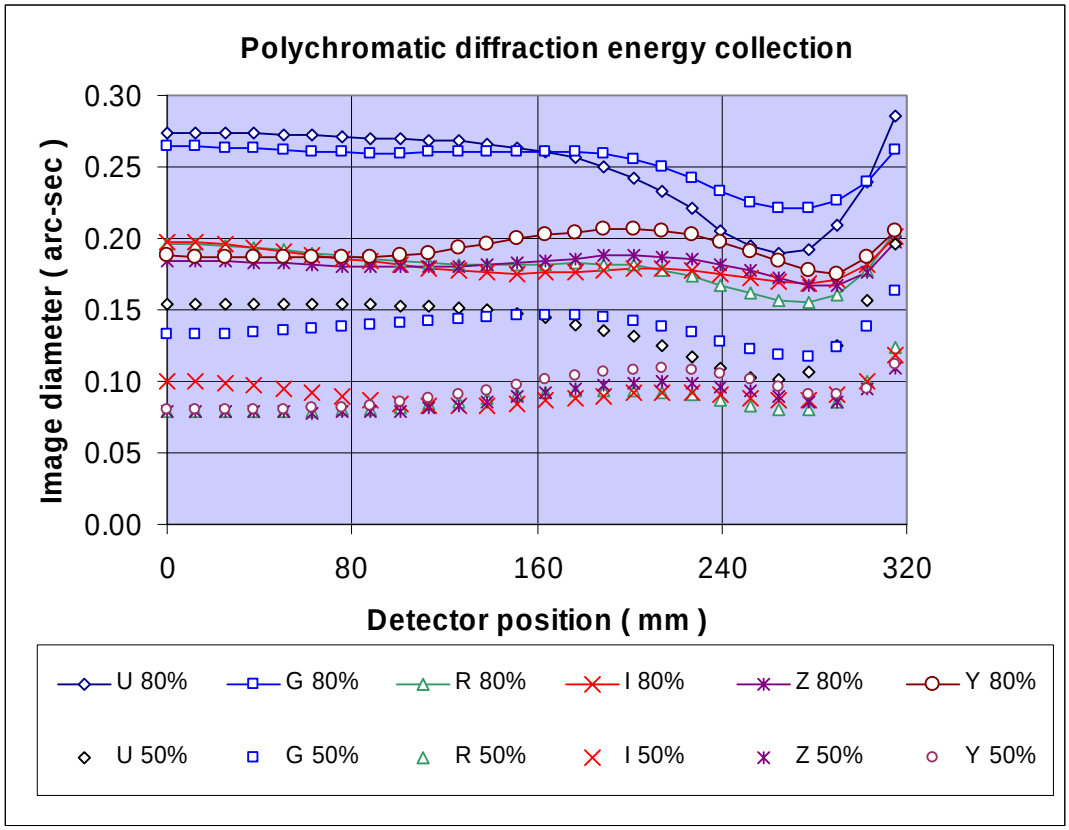


Figure 4.1.5-2 The polychromatic image sizes are less than 0.2 arcsec across the field for 50% energy collection. Except for the g-band, the polychromatic image sizes are less than 0.24 arcsec across the field for 80% energy collection. Filter bands: u: 320nm – 400nm g: 400nm-560nm, r: 540 nm-720 nm, i: 685 nm-870 nm, y: 840 nm-950 nm, z: 945 nm-1030 nm.

With this optical design, the camera can be attached directly to the top end of the telescope, limiting the number of spiders intercepting the optical beam. Previous designs had the camera supported closer to the primary mirror, forcing additional support veins. The camera can also be inserted in and out of the telescope through the M2 central hole, a significant operational enhancement over previous longer designs. The detector plane is located 1.73 m from the vertex of the secondary mirror. It procures also a very compact design. The camera assembly from the detector plane to the vertex of L1 is only 1.03 m long. And the telescope length is 6.39m, which is shorter than the primary mirror diameter of 8.4m.

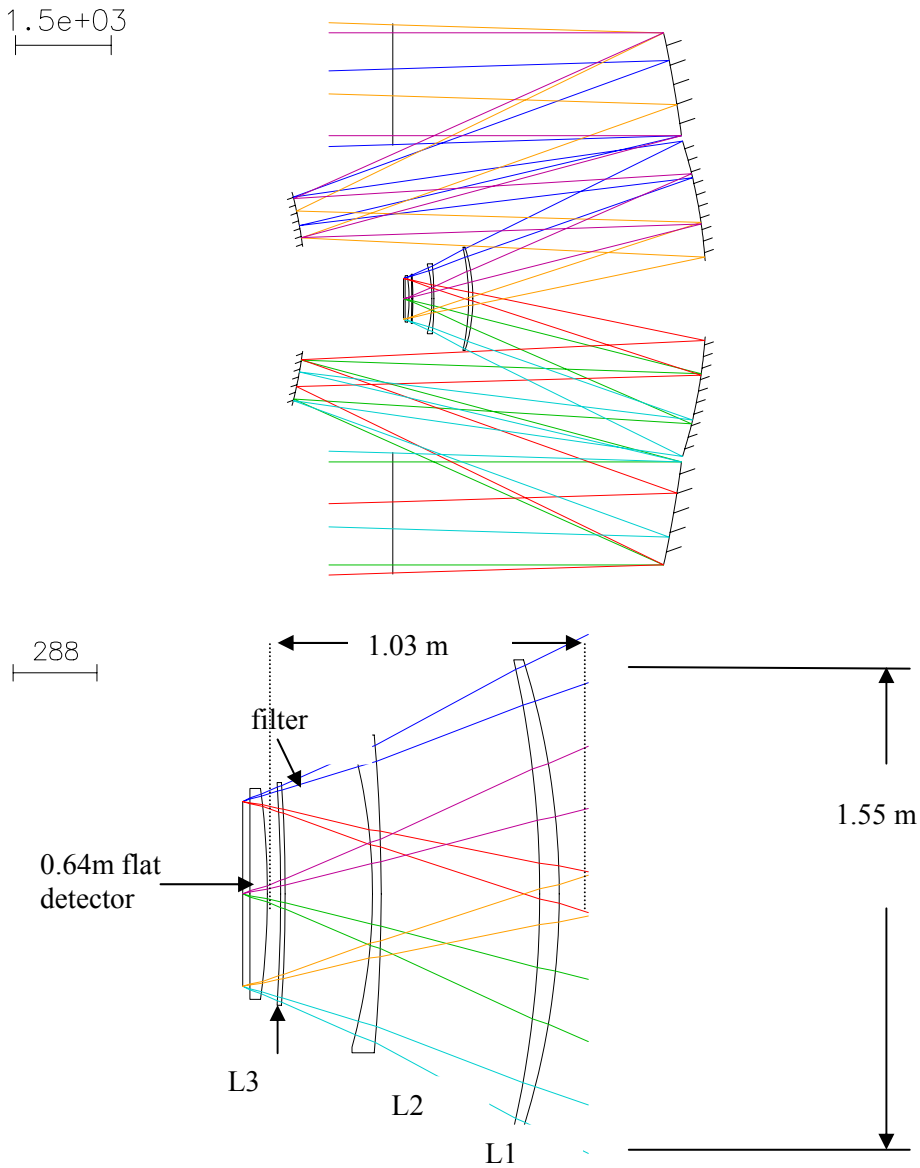


Figure 4.1.5-3 LSST optical layout *(top)* and camera corrector optical layout *(bottom)*

The combination of large diameter optics with fast f/# and a large detector at the focal plane makes the LSST a unique facility achieving an etendue of 318 m^2deg^2, a factor of >50 beyond current existing facilities. This etendue value takes into account the large central obscuration of 5.1m in diameter and the variable vignetting toward the edge of the field (11.2% at the edge).

The wide range of wavelengths specified for the LSST requires some adjustments for operating at different spectral bands in order to preserve the high image quality. First, filters with different passbands need to be inserted to change the spectral range. Each filter has a unique central thickness to compensate for chromatic difference in aberrations. Thicknesses range from 26.3 mm in the U band to 13.5 mm in the Y band. Furthermore, each filter has a slightly different second radius of curvature to further correct for chromatic aberration. The central i-band filter is equi-meniscus, with radii of curvature of 5624 mm convex and concave. The second radius of other filters varies from 5507 mm in the U band to 5597 mm in the i band. Second, the entire camera

assembly is axially refocused, ranging from 1.78 mm away from the tertiary in the U band to 0.58 mm toward the tertiary in the Y band. The negative lens L2 is kept at a fixed position for all wavelengths.

4.1.6 Observatory Controls

The LSST Control System (LCS) is the ensemble of computer hardware and software, tailored to efficiently and safely perform astronomical observations individually or through automated scheduling. It also provides support for engineering, set-up, maintenance, and creates a dynamic environment for development and evolution of control applications. This system insures safe, effective on-sky time, low down-time, and high data pipeline throughput. In addition to it control functions, the LCS captures, organizes and stores system wide state information, to make it available for monitoring, evaluation and calibration processes. This telemetry and observatory conditions information will be distributed to support both local and remote operations in an efficient manner. The LSC is also developed around a robust architecture to maximize the control efficiency and to support the highly robot nature of the LSST System.

4.1.6.1 Overview of LSST Control

The LSST Control System (LCS) is the ensemble of computer hardware and software, tailored to efficiently and safely perform astronomical observations individually or through automated scheduling. It also provides support for engineering, set-up and maintenance, and creates a dynamic environment for development and evolution of control applications. The LCS is also developed around a distributed architecture to maximize the control efficiency and to support the highly robotic nature of the LSST System.

The LSST Control system architecture is based on the four fundamental steps performed through an observation lifecycle:

- Creating and executing observations
- Configuring the telescope
- Operating the camera
- Handling data

The LCS is designed to support these steps and the associated ancillary requirements through its five principal subsystems:

- The Observatory Control System (OCS)
- The Telescope Control System (TCS)
- The Camera Control System (CCS)
- The Data Management Control System (DMCS)
- The Calibration Control System (CACS)

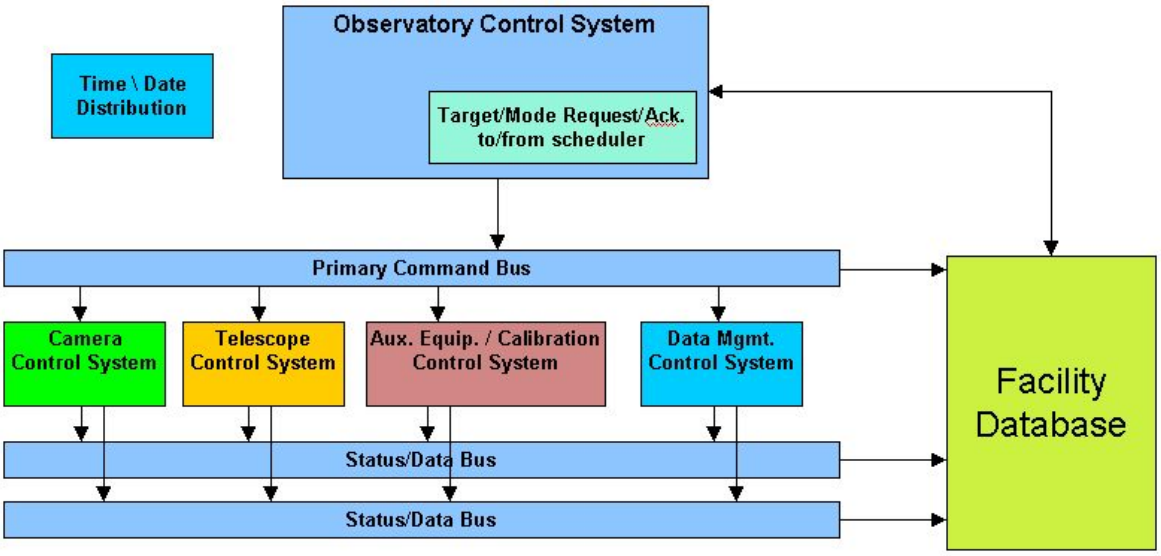


Figure 4.1.6-1 The following diagram gives an overview of the LCS system architecture.

Several basic principles are taken into consideration for this architecture:

- The baseline design concept distributes control to the local subsystem level. Time critical loops will be closed at the local level for robust accurate timing, safe operation and organized control and development.
- The principal systems will follow a "Master / Slave" strategy. The OCS will act as the Master of the system, accepting observation commands from an automated scheduler or operator and orchestrating the performance through interactions with the subsystems.
- Communication in the LSC is organized in bus structures. There will be a Primary Command Control Bus and status and data buses organized by content and or performance characteristics.
- The system will have extensive logging capabilities of all aspects of the LSST operation. The Observatory Telemetry System (OTS) will eavesdrop on control and data buses, logging status to archive observing conditions and support operational maintenance and reliability efforts.
- Coordination of the LCS elements will be achieved by the exchange of messages through the system's interfaces.

The Observatory Control System (OCS) is the master control system that schedules, coordinates, commands and monitors the observatory. The OCS is responsible for high level observatory operations including user interfaces, scheduling, resource allocation and system monitoring and maintenance. The OCS orchestrates and controls all aspects of the observatory for all observations (science, calibration, and engineering) and all operation modes. Through the OCS the system can be started, monitored, adjusted during operations and stopped, both locally and remotely. The OCS provides the means to support safe observatory operations day and night.

The Telescope Control System (TCS) is the central coordination facility for the delivery of high quality field images to the camera. It is responsible for the precise pointing and tracking

calculations necessary to observe a certain field. The TCS does not itself operate any mechanical component; rather it delegates this responsibility to the various telescope subsystems and manages them according to the observation requests.

The Camera Control System (CCS) commands the elements that compose the camera systems and delivers the pixel images to the DMCS. It receives commands from the OCS regarding filters, exposure time and triggers. The CCS sends the science data to the Data Management system. The WFS images and guider images are sent to the TCS. These streams of information are sent through dedicated data buses.

The Data Management Control System (DMCS) is in charge of the data handling, data real time processing and the data production/mining tasks. It gives feedback status regarding the quality of the science data to the OCS "in real time" so the scheduler is able to reassess its priorities dynamically and operators can assess observatory performance.

The Calibration Control System (CACS) handles the calibration and auxiliary equipment (cameras, weather, etc), needed to setup the telescope and camera to match the observing conditions. This system will also provide rigid body alignment feedback during operation to limit the misalignments and reduce the load on the higher resolution WFS.

In addition to its control functions, the LCS captures, organizes and stores system-wide state information to make it available for monitoring, evaluation and calibration processes. This telemetry and observatory condition information will be distributed to support both local and remote operations in an efficient manner.

During operations the observation control originates from a sophisticated LSST observatory scheduler, a scientific monitor engaged in real time at any remote operation center, and an operator on the summit. There will be a limited summit staff to ensure safe operation and to troubleshoot as required.

4.1.6.2 Control System Communications

The Control System is founded on a communications layer that provides an efficient and reliable way of exchanging messages between the different systems and components of the LSST complex. The communications layer will be built on industry standard software known as *middleware*. Successful development of distributed application systems relies on the middleware separating application-specific functionality from the logic complexities inherent in a distributed infrastructure.

In real-time distributed applications like LSST, messages are often encapsulated as one of the following types of data flow:

- **Signals.** Rapidly generated and time-critical data. In most instances, it is more important to get the next issue than to retry a dropped issue.
- **Events.** Asynchronously generated, time-critical messages which must be delivered reliably.
- **Commands.** Sequential instructions which must be received in order.
- **Status.** Persistent data about state or goals. Its timeliness differs from one application to the next.

- **Requests.** Two-way request-reply transactions for a specific service or data.

The LSST communications layer, implemented on top of an industry standard middleware, will provide the hooks for applications to seamlessly participate in the message flow.

4.1.7 Calibration Strategy

In order to reach the level of photometric and astrometric accuracy set in the science requirement document, the LSST demands a global calibration process, from the observing strategy all the way to the pipeline outputs.

In the core of the observing strategy is included the need to minimize the systematic errors, which translate into the strategy for acquisition of the data. The reduction of absolute calibration error will first come naturally from the averaging of multiple exposures of same fields taken on different nights. To extract the most of this averaging process, the systematic errors will be beaten down by ensuring that in the acquisition sequence:

- the sky orientation is variable between exposures of a same field in order to image the field in different locations of the focal plane
- the fields are observed at different airmasses
- the center sky-coordinates are offset between exposures of the same field
- the overlap between adjacent fields is large enough for adequate comparison between images

Tracking all these parameters (the list will be completed during the R&D phase) during the operational phase will be a task embedded in the observatory control system.

In a second step, calibration data will be recorded and analyzed to produce models and transfer functions of the LSST observations. The goal is to identify and to generate all the calibration information required to restore the intensity map of each science field as a function of accurate celestial position, before absorption and distortion by the atmosphere and the instrument system.

The traditional approach is to extract part of this information directly from sky observations using celestial standards or fields. For the astrometric calibration, observations of sufficiently densely populated astrometric fields will provide a reference frame between all the detectors of the large mosaic. This model would incorporate the relatively low distortion of the LSST optical design (reaching 0.1% at the edge of the field), Each chip will also be modeled independently with low order polynomials to generate a distortion map, using other observations of moderate galactic latitude fields. Such approach has been used successfully by other projects using large mosaic. Moreover, laboratory measurements of the focal plane, done during the characterization of the instrument, will be included in this modeling process as a way to constraint the variables of the model. A-priori knowledge of the distortion will improve the quality of the fit. During operation, thanks to the large field of view, each exposure will contain enough celestial standards to compute an accurate astrometric solution for each image and to compare these results with the models. This tool could be useful to monitor the stability of the telescope. Eventually, LSST will be able to use its own catalogs of fainter objects as reference by defining new standards for each possible science field.

For the photometric calibration, multiple physical processes occur in the atmosphere and in the instrument (see Figure 4.1.7-1) that would affect the photometric accuracy. A large set of information would be required to correct these processes as their temporal and spatial variations could be non-negligible.

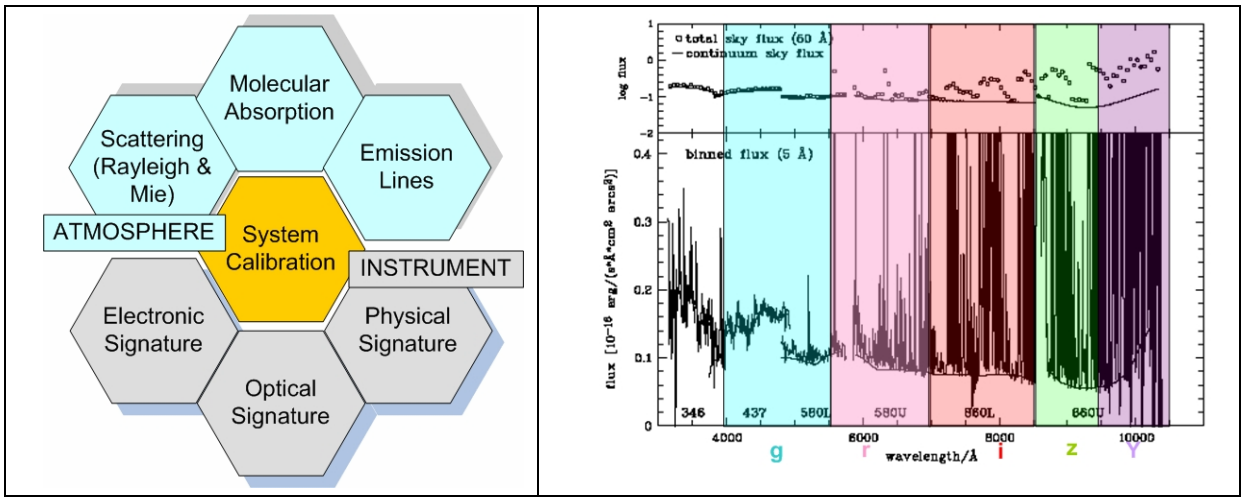


Figure 4.1.7-1 On the right side, sky emission spectrum with plot of grizY filters. Top: total sky emission flux across 50A bins. Bottom: the sky emission spectra sampled into 5A bins (from UVES instrument, VLT). On the left side, schematic view of physical processes for atmospheric and instrumental effects.

For the relative calibration, all of the measured magnitudes are placed onto a magnitude system corrected for instrumental effects, but ignoring the magnitude zeropoint. The characterization of the instrumental signature includes:

- The electronic signature evaluated through cross-talk measurements, bias frame imaging linearity measurements and mapping of bad pixels,
- The optical signature captured with dome flats and illumination correction map to take into account the different light path between sky and dome measurements, as well as determination of fringing for longer wavelength bands,
- The physical signature including measurements of mechanism effects (for instance shutter timing) and of scattered light in the focal plane.

The overlap areas will be used to get a relative photometric calibration between all the fields.

For the absolute calibration, the calibration of known standard stars is transferred onto other celestial sources after correction of atmospheric effects. LSST science requirements demand an accurate monitoring of the extinction across the broad band filters for precise estimation of the zeropoints. We plan to take advantage of the multiple images of each field to improve the accuracy of the absolute calibration by using different observations of standard stars taken on different nights. Different scenarios will be studied during the R&D phase to estimate this possible gain. Color-color diagrams could be used also as a method to check the accuracy of the absolute calibration as was demonstrated by the SDSS project.

The initial all-sky calibration may require to extend the dynamic range to brighter sources to exploit the whole potential of existing reference catalogs, depending on the progresses of existing survey projects. If necessary, observations through an r filter+Neutral density (or shorter exposure time) could be implemented to avoid saturation on brighter sources.

Last but not least, the calibration strategy is included in the overall data management structure as it would be impossible to do these tasks in a non-automatic mode and keep up with the LSST data rate. A schematic view of the calibration process is given in figure xxx. In that process, all the calibration data is archived in the data management system and processed to produce calibration data products. The calibration data products are used as input in the main image

processing pipeline that produces calibrated images. Reference catalogs generated by LSST (after so many years of observations) or initially from previous surveys are injected in the detection pipeline to produce calibrated catalogs of celestial sources.

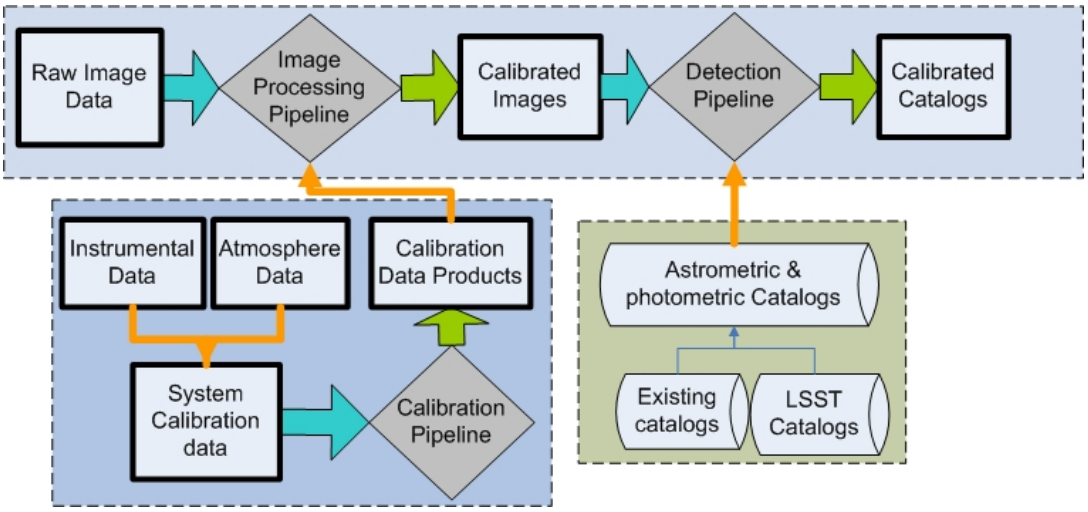


Figure 4.1.7-2 Calibration process schematic.

4.2 Telescope/Site

4.2.1 Overview and Requirements

4.2.1.1 Scope and Definition

The Telescope and Site task includes the buildings and facilities for both operations and maintenance of all the LSST summit and base assets. The telescope is the light-collecting module that supports the optical system and camera assembly and positions them to acquire and track fields in the sky to demanding precision. The telescope has active position control for these critical elements as well as systems to mitigate the impact of natural operating conditions on the image collection process. This includes light baffling systems, protection from wind-induced alignment and optical figure perturbations, active controls of the thermal environment for both alignment and image quality stability, and the general facility and dome for protection from the elements and the support of all operating and maintenance functions.

The LSST telescope and facility are designed to be highly robotic with minimum human intervention. During operations, the control of observations originates from a sophisticated LSST observatory scheduler; a scientific monitor may be engaged in real time at any remote operation center, and an operator will be present on the summit. In addition, there will be a limited summit staff to ensure safe operation and to troubleshoot as required.

Figure 4.2.1-1 Rendering of the LSST Telescope and Summit Facilities.

4.2.1.2 Telescope and Site Requirements

The requirements for the Telescope and Site are naturally divided into technical performance requirements and operational support. The technical requirements derived from the flowdown and error budgeting described in Sections 2 and 3 are summarized in Table 4.2.1-1 and impact most significantly the actual telescope and the telescope enclosure. The operational constraints imposed by both the science requirements and the hardware itself define the scope and performance of the additional assets at the summit and base. The following sections are aligned with the major sections of the work breakdown structure and provide the scope and definition of the baseline design for the Telescope and Site portion of the LSST.

Table 4.2.1-1 Telescope and Site Requirements

Main Telescope and Site Technical Requirements	
Tech-01	At all elevations, the absolute telescope pointing requirement is 1.5” RMS with a goal of 1” RMS.
Tech-02	For an offset of 3.5 deg, the pointing error must not be more than 0.5” RMS with a goal of 0.1” RMS.
Tech-03	For smaller offset, the pointing error must not be more than 0.1” RMS.
Tech-04	Tracking must be better than 0.02” RMS over a 1 min interval, and better than 0.4” RMS in 1 hour.

Tech-05	The telescope will be capable of doing a 3.5 deg offset in any direction in 5 sec or less (including slewing and settling time).
Tech-06	The telescope will be designed with an alt-az mount that will be capable of ± 270 degrees rotation in azimuth, and will be capable of moving from zenith to horizon in elevation (0 deg to 90 deg).
Tech-07	The lowest resonant frequency associated with each of the axes will be above 10 Hz. Care shall be taken to prevent the coupling of vibration modes. The natural frequencies of all structures and components shall be sufficiently greater than the lowest resonance frequency to prevent coupling of vibration modes.
Tech-08	The telescope effective collective aperture will be at least equal to 6.5 m, and the telescope will provide a field of view of 3.5 deg in diameter.
Tech-09	The telescope intrinsic image quality FWHM will be better than 0.3 arcsec in the R band at zenith.
Tech-10	The telescope will be actively controlled to maintain its image quality at all elevations.
Tech-11	The telescope optical efficiency will be maximized from the U band to the Y band.
Tech-12	The telescope design will be optimized to have negligible stray light and scattered light.
Tech-13	The spider design will be optimized to reduce the vignetting of the entrance beam.
Tech-14	The site clear sky statistics should be approximately 70%.
Tech-15	The median site seeing should be better than or equal to 0.7 arcsec at 500 nm.
Tech-16	The site should have negligible light pollution and contrails, and a low extinction.

Main Telescope and Site Operational Requirements	
Oper-01	A telescope control system (TCS) will provide control of all operations of the telescope.
Oper-02	The telescope will be designed for low maintenance and safe and easy handling of equipment.
Oper-03	Telescope downtime will be on average less than 3 days/month.
Oper-04	The telescope minimum operational lifetime will be 10 years. The design requirement is 30 years.
Oper-05	A coating facility and all other ground support equipment will be readily accessible on site.
Oper-06	A calibration system will be available to provide the necessary level of photometric and astrometric calibration of the system.
Oper-07	The telescope will be designed to reduce and to control thermal effects during operation.
Oper-08	The telescope will include a camera rotator with a range of rotation equal to or larger than 90 degrees to reduce image systematics.
Oper-09	A telemetry system will be incorporated into the telescope design for status monitoring and rapid alerts for troubleshooting. This system will also monitor the site environmental conditions.
Oper-10	A high bandwidth data communication system based on fiber optics will be available on site.

4.2.1.3 Site Location

The LSST Project is considering three observatory locations for siting the telescope. The three locations are listed in Table 4.2.1-2.

Table 4.2.1-2 LSST Sites (in order of north latitude to south latitude)

Site	Latitude	Longitude	Elevation (m)
San Pedro Mártir	N31.04	W115.45	2800
Las Campanas	S29.05	W70.68	2200-2600
El Peñón	S30.2	W70.8	2600

These locations were selected through an evaluation process established by the LSST Project. An independent committee was formed to guide the site evaluation and recommend the down selection to the LSST management. This approach narrowed the options in two steps. The first considered all worldwide observatory locations and recommended four to focus on; the second removed one additional option, leaving the present list of finalists. At each stage a large data set was compiled to measure the potential sites against the LSST requirements for both performance and operation. The process continues with the third stage of data collection now in progress to lead to the final site determination. The final phase of evaluation is described in detail in Section 5.2.1

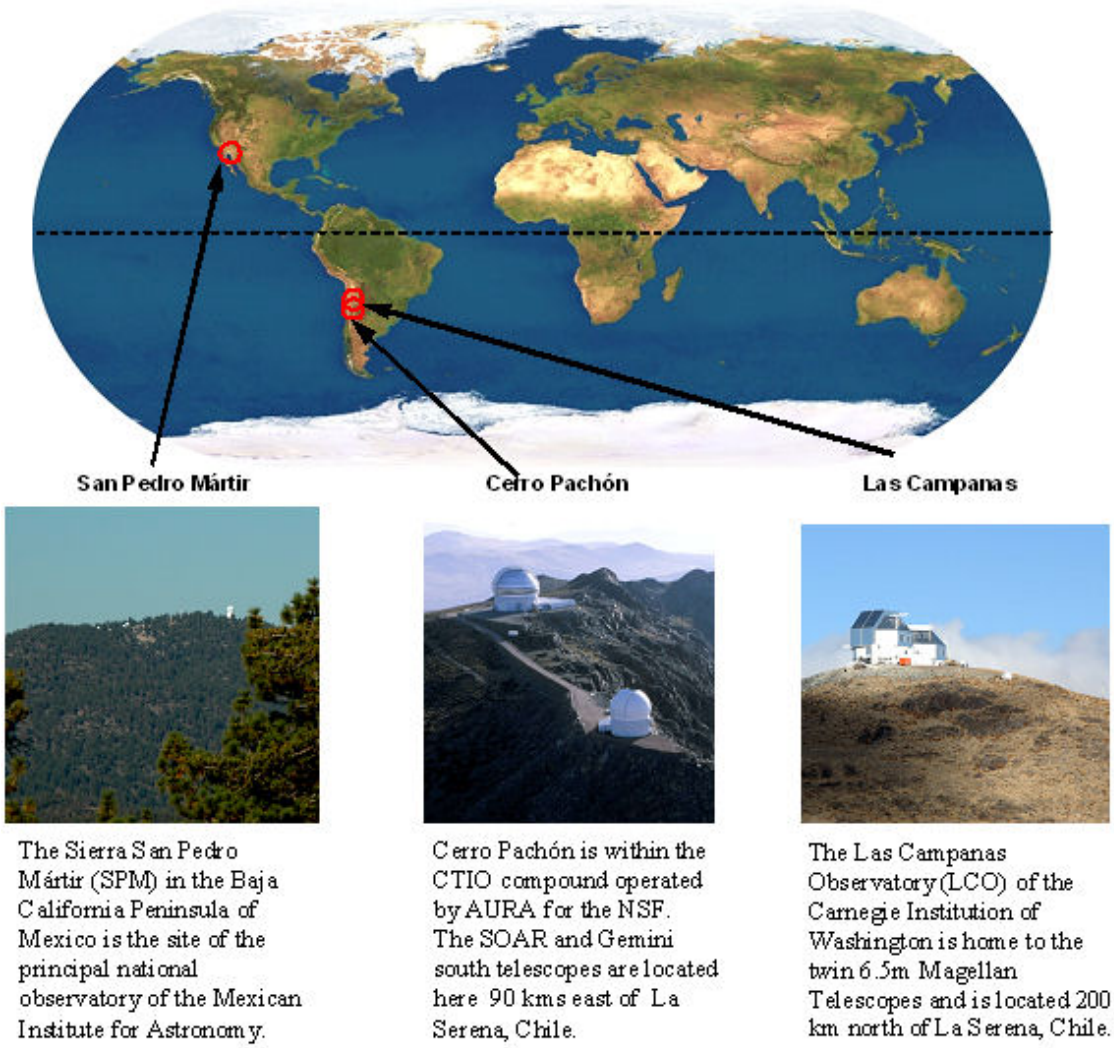


Figure 4.2.1-2 Candidate LSST site locations

4.2.2 Facilities

4.2.2.1 Scope and Requirements

The LSST has a significant requirement for several specialized facilities both on the summit and in the nearest support town. Included are: the Telescope Enclosure (Dome & Lower Enclosure), the Summit Support Facility, the Summit Dorm & Dining Facility, and the Base Support Facility. The preliminary requirements identified for each of these facilities, along with the requirements for the site itself, are described in the following sections.

The requirements for the baseline LSST support facilities have been based on those established for existing large nighttime observatories. Additionally, the survey mission of LSST mandates continuous operation, an essentially robotic telescope control program, a specialized observing staff with few or no visiting observers, and a critical requirement for rapid data reduction and transmission.

4.2.2.1.1 Enclosure (Dome)

The basic requirement of the dome is to protect the telescope and camera from weather conditions both during observing and when not in operation.

- **Dimensional Requirements** - The clear interior area of the dome will allow the telescope to fully rotate in both azimuth and elevation without requiring co-rotation or retraction of the dome. This is intended to facilitate maintenance of the telescope-related systems and is in keeping with the relatively compact geometry of the LSST mount and optical design.
- **Aperture** - The observing aperture of the dome must provide a 9 m diameter unobstructed observing aperture from 15 deg above horizon to 10 deg past zenith. The lower border of this aperture will be formed by an appropriate windscreen.
- **Tracking** - The requirements for speed and accuracy of the dome rotation and shutter drive systems will be dictated by the telescope acquisition and tracking specifications. This will require the dome to match the unusually fast slew-and-settle time for the telescope, a very significant requirement for traditional dome implementations and a major use of the summit power.
- **Telescope Maintenance Facilities** - The dome will provide adequate space and appropriate facilities for engineering and maintenance work on the telescope and camera. This will include suitable lifting equipment for the primary mirror assembly, the top end assembly and the camera.
- **Basic Shape Determinants** - The dome will be designed for a minimum amount of enclosed air volume and exposed exterior surface. This will allow for faster and more effective passive flushing during observing and will reduce the amount of air conditioning required to maintain and pre-cool the telescope environment during the day. The shape of the dome will take into account the maintenance and climatic conditions of the selected site. Flat surfaces that collect snow and ice and joints that require complex seals or flashing will be avoided as much as possible.
- **Environmental Control** - A fundamental driver in the design of the dome will be minimizing local seeing and maintaining a beneficial thermal environment around the telescope. Operable ventilation openings and/or other strategies will be employed for that purpose. The dome will also allow for effective pre-conditioning of the telescope and its immediate environment to the anticipated ambient air temperature at evening start-up.

4.2.2.1.2 Lower Enclosure

The lower enclosure will be a fixed structure that surrounds the telescope pier and supports the dome at an appropriate height. The anticipated height of the lower enclosure (to the base ring of the dome) is ~15m. The structural design of the lower enclosure and its foundation must provide for adequate vibration isolation from the telescope pier. The lower enclosure will also include any supplemental building space necessary for functional and operational requirements in the immediate vicinity of the telescope. The tentative space requirements are listed in Table 4.2.2-1.

Table 4.2.2-1 Space requirements for lower enclosure and attached building

Space Description	**Area**		**Height**	
	ft^2	m^2	ft	m
Observing Floor (~32 m diam.)	8650	804	up to dome	
Utility Distribution Area (under obs floor)	6000	558	9	2.7
Engineering Control Area	400	37	9	2.7
Base of Pier (~12 m O.D.)	1220	113	20	6.1
High-bay Receiving/ Mirror Prep	2400	223	50	15.3
Mirror Coating Area	2400	223	50	15.3
Storage	1080	100	20	6.1
Platform Lift (~9 m x 9 m)	900	84	shaft	
Elevator (~1.5 m x 2 m)	80	7	shaft	
Stairs	180	17	shaft	
Mech. Equipment Space	800	74	10	3.1
Machine and Service Rooms	250	23	9	2.7
Restroom	60	6	8	2.4
Total Required Space (net)	**24,420**	**2,271**		

4.2.2.1.3 Summit Support Facility

The basic function of summit support facility is to provide suitable space for the operation and maintenance of the LSST telescope, camera, data processing system and related equipment. As the physical and operational requirements of all these systems are more clearly defined, the information will flow to the facility requirements and be incorporated into the facility design process. The summit support facility is expected to be a separate structure within 100 m of the enclosure. Its location (based on prevailing wind and topography) and design (low profile, utilizing materials with low thermal inertia) will minimize potential thermal disturbance in the light path of the telescope. The summit support facility will house the telescope control room, the summit data processing facility and other necessary on-site support spaces (as listed in Table 4.2.2-2). The size and layout of the building will be governed by the functional requirements of the individual spaces.

Table 4.2.2-2 Space requirements for summit support facility (detached)

Space Description	**Area**		**Height**	
	ft2	**m2**	**ft**	**m**
Control Room	650	60	9	2.7
Offices (4)	520	48	8	2.4
Workstations	650	60	8	2.4
Kitchen/Break Area	190	18	8	2.4
Restrooms (2)	140	13	8	2.4
Data Processing Room	600	56	10	3.1
Camera/Instrument Shop	2000	186	12	3.7
Electronics Shop	600	57	10	3.1

Machine Shop	640	60	12	3.7
Receiving	1000	93	12	3.7
Interior Mechanical Equipment	650	60	9	2.7
Exterior Mechanical Equip.	750	70	N/A	
Generator Enclosure	1000	93	N/A	
Total Required Space (net)	**9,390**	**873**		

4.2.2.1.4 Summit Dorm & Dining Facility

The remote nature of the three potential sites under consideration requires facilities for on-site sleeping quarters and meal service. The necessary accommodations would likely be combined with similar facilities for existing observatories, but construction of a separate LSST Dorm & Dining Facility is potentially necessary. Table 4.2.2-3 describes the basic space requirements to support 20 day staff and 10 overnight staff.

Table 4.2.2-3 Space requirements for summit dorm/dining facility

Space Description	**Area**		**Height**	
	ft2	m2	ft	m
Dorm Rooms (w/ bathroom) (240 ft^2 /rm.)	2400	223	8	2.4
Dining Room	1020	95	8	2.4
Kitchen/serving & food storage	1000	93	8	2.4
Lounge	500	47	8	2.4
Utility/Mechanical	200	19	9	2.7
Total Required Space (net)	**5,120**	**476**		

4.2.2.1.5 Base Support Facility

The administrative offices, intermediate data processing capacity, and the other remote support functions of LSST are expected to be located in the nearest appropriate community at the base of the mountain where the telescope is sited. The nature of this facility is to some extent dependent on the nature of the surrounding community. It may be possible to take advantage of common resources at an existing support complex, or an entirely stand-alone facility may be necessary. A preliminary assessment of the basic functional space requirements for an LSST base facility has been projected. (Table 4.2.2-4).

Table 4.2.2-4 Space requirements for base support facility

Space Description	**Area**		**Height**	
	ft2	**m2**	**ft**	**m**
Reception Lobby	300	28	9	2.7
Administrative Offices (12)	1680	156	8	2.4
Visitor Offices (2)	280	26	8	2.4
Shared Office Space (~20 people)	2000	186	8	2.4
Conference Room (30 person cap.)	600	56	9	2.7
Break Room	200	19	8	2.4
Data Processing and Archive	600	56	9	2.7
Data Technician Work Space	120	11		
Auxiliary Observatory Control Room	600	56		
Camera/Optics Lab	800	74	12	3.7
Other Shops?	600	56	10	3.1

Storage/Receiving	650	60	10	3.1
Restrooms (2)	160	15	8	2.4
Mechanical Equip.	800	74	N/A	
Total Required Space (net)	**9,390**	**873**		

4.2.2.1.6 Observatory Site

The scientific (seeing and weather) criteria for site selection are described elsewhere. The basic technical criteria for a suitable site for LSST are as follows:

- **Accessibility** - A range of vehicles, from standard passenger vehicles to large construction cranes and flatbed trucks, must be able to reach the facility, both during construction and in long-term operation.
- **Dimensions** - The topography of the site has to accommodate the creation of suitable level area(s) (~100,000 sq. ft. total) for the required summit facilities and their associated service areas. A larger area would allow a more advantageous site layout to minimize local seeing effects and would facilitate construction staging.
- **Structural Characteristics** - The soil/rock of the site must have sufficient bearing capacity to support the loads imposed by the telescope pier and the building foundations while also allowing for adequate isolation between the two. Stiffer natural substrates that increase the lowest resonant frequency of the telescope support system are very advantageous. The lateral force factors (seismic and wind) inherent to the site must be of a magnitude that can be safely mitigated without prohibitively expensive structural measures.
- **Utility Infrastructure** - Sufficient electrical power, data connection, and domestic water/sewer service must be achievable at the site. Existing services that can be extended to LSST and connection to local utility company lines are advantageous.

4.2.2.2 Conceptual Design

4.2.2.2.1 Dome Conceptual Design

Early dome design concepts have focused on minimizing the enclosed volume for economy and reducing dome seeing effects while still serving the given functional requirements. This has led to an initial proposed baseline of a simple hemispherical shape with a series of ventilation openings around the equator. The most space-intensive functional requirement is the need to handle the primary mirror assembly (PMA) and the top end assembly (TEA), which includes the camera. In the proposed maintenance scheme for LSST, both of these large assemblies will have to be removed from the telescope mount on a periodic basis and transported to adjacent facilities for routine maintenance. This will involve a high capacity lift and/or crane that is integrated into the dome structure and environment. The approximate range for the dome diameter, based on this requirement, is from 28 m, allowing for only a small overhead crane and an external lift for transport of the major assemblies, up to ~40 m diameter, allowing for the incorporation within the dome of a large (80 ton capacity) crane or platform lift. (Figure 4.2.2-1)

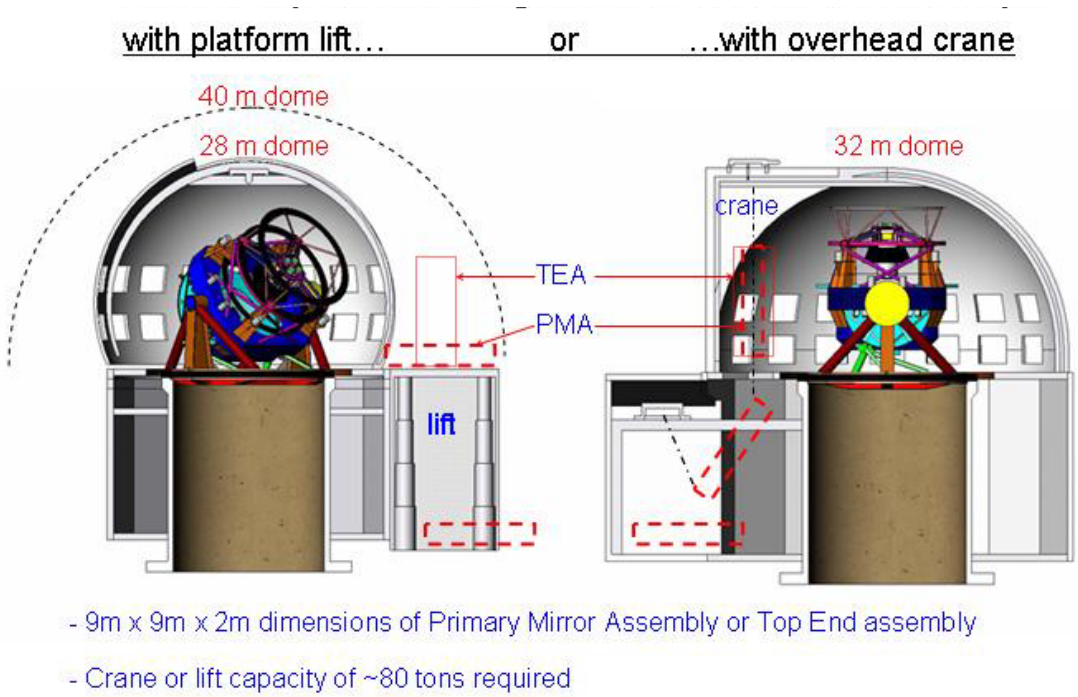


Figure 4.2.2-1 Mirror & Top End Handling as driver of dome size and shape

In recent years observatory domes have been built with a variety of structural configurations other than the traditional hemispherical shape. Some of these have been intended to reduce cost and others have been devised to address special functional requirements. As the LSST design moves forward, a range of options for the basic size and shape of the dome will be explored, including cylindrical, rectangular (co-rotating), and roll-off enclosures. The basic criteria used to evaluate alternative options will be dome seeing, functional requirements and cost.

4.2.2.2.2 Lower Enclosure Conceptual Design

The principle structural elements included in the lower enclosure are the telescope pier and the cylindrical support structure for the dome. The lower enclosure design will also include any additional building volume needed to incorporate a platform lift, crane, mirror coating facility or other functional necessities.

The structural criteria for the telescope pier will be dictated by the required stiffness of the overall telescope structure and by the soil/rock conditions of the selected site. For the purposes of the lower enclosure design, the pier is assumed to be a concrete cylinder ~12 m in diameter, ~15 m high, with telescope utility wrap enclosed inside. The specific dimensions are subject to change, but the general size, shape and material of the pier are relatively predictable based on previous observatory construction.

The cylindrical dome support structure will also be similar to that of existing observatories. The most likely configuration, and the one initially assumed for LSST, is a circular pattern of braced steel columns supporting a fixed ring beam at the top with panels of insulated metal siding between. This dome support structure would be isolated from the telescope pier to mitigate the impact of wind buffeting and facility vibration on the telescope. In addition to the dynamic loads of the dome, the lower enclosure will support a telescope-level floor and any necessary intermediate floors.

The functional spaces in the lower enclosure are subject to a wider range of design possibilities, with the main factor being the potential incorporation of the mirror coating facility. For reasons of economy, the large-scale spaces required for this facility will be partially accommodated by the inherent enclosed volume of the lower enclosure. Another factor in favor of this option is that there is a significant operational and risk-management benefit in not requiring exterior transport of the mirror assembly. The base of the telescope enclosure would likely require an extended structure to accommodate the mirror coating chamber and related spaces. The location of the PMA transport lift or crane and the diameter of the dome will impact the size and shape of any necessary extension for the mirror coating area. This could have an undesirable effect on natural air flow near the aperture of the adjacent dome, so the mirror coating plant may instead be located in the separate summit support facility. The other anticipated functional areas of the lower enclosure (listed in Table 4.2.2-1) can likely be housed in the enclosed space inherently created by the height and diameter of the structure.

4.2.2.2.3 Summit Support Facility Conceptual Design

This facility will accommodate all the functional areas that are required on the summit but do not require direct proximity to the telescope enclosure. The spaces listed in Table 4.2.2-2 have been arranged into a conceptual building floor plan shown in

. The purpose of this plan is to allow for early exploration of site layouts and to initiate discussion with the LSST team about the design of this critical operational area. The most important and functionally demanding areas are considered to be the control room, the data processing room and the camera/optics shop.

4.2.2.2.4 Summit Dorm/Dining Facility Conceptual Design

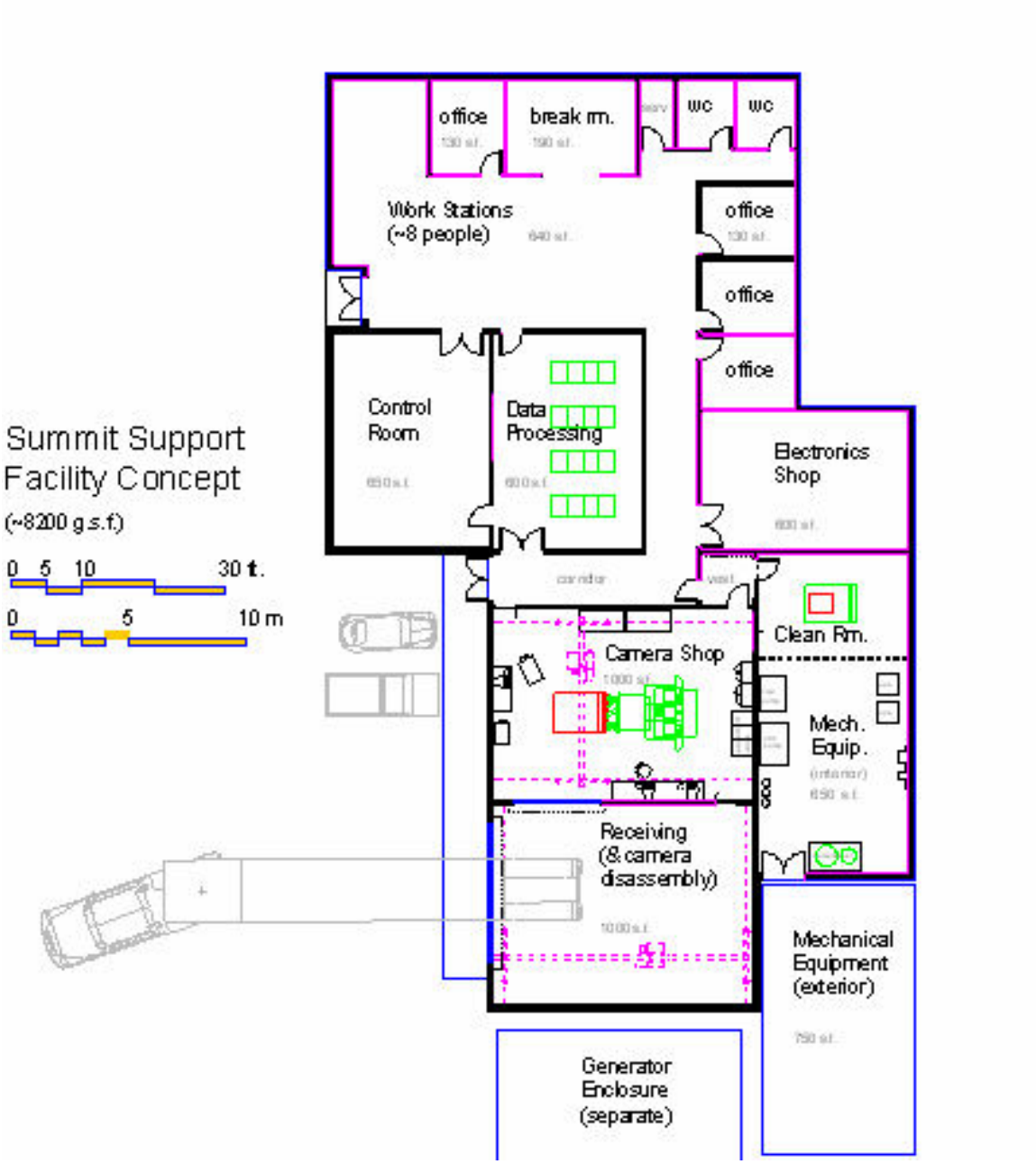


Figure 4.2.2-2 Summit Support Facility Conceptual Plan

No design work has been initiated. The summit Dorm/Dining facility is very dependant on the specific site and the existing assets available for use by LSST.

4.2.2.2.5 Base Support Facility

No building design work has as yet been initiated. Early exploration of base facility locations for the three remaining candidate sites indicates that proximity to the existing base facilities of other observatories will likely be possible and advantageous. For the San Pedro Mártir site the probable location would be in the city of Ensenada adjacent to the headquarters for the Mexican National Astronomical Observatory. This complex, located in an academic campus area in Ensenada, houses the operational support functions of the SPM observatory as well as offices for staff scientists and engineers. For either of the Chilean sites, the likely location for an LSST base support facility would be in the city of La Serena adjacent to the base facilities of the other AURA-operated observatories - Gemini, SOAR and CTIO. The design of a base facility for LSST in either Mexico or Chile will take into account the community context, the characteristics of the available site and potential sharing of support spaces with neighboring facilities.

4.2.2.2.6 Site Conceptual Design

During the site down-selection process, feasibility studies of all the potential sites were conducted. That work included tentative layout of the LSST summit facilities at specific proposed locations on San Pedro Mártir, Cerro Pachón, and Las Campanas.

4.2.2.2.6.1 San Pedro Mártir

The proposed site for LSST at SPM is a currently unutilized area within the existing telescope compound, which is also identified as a possible location for a proposed new large Mexican telescope. If the final recommended location for LSST is San Pedro Mártir, it is anticipated that the Mexican National Observatory would make this entire site available for the LSST facilities. This site has the advantage of being close to the existing utility infrastructure and paved access road. The layout shown in Figure 4.2.2-3 shows a 32 m diameter LSST enclosure at the high end of the ridge with the support facility in an adjacent site area that is lower and closer to the road. Other potential sites at SPM have been identified, but they are much more remote or undesirable for other reasons.

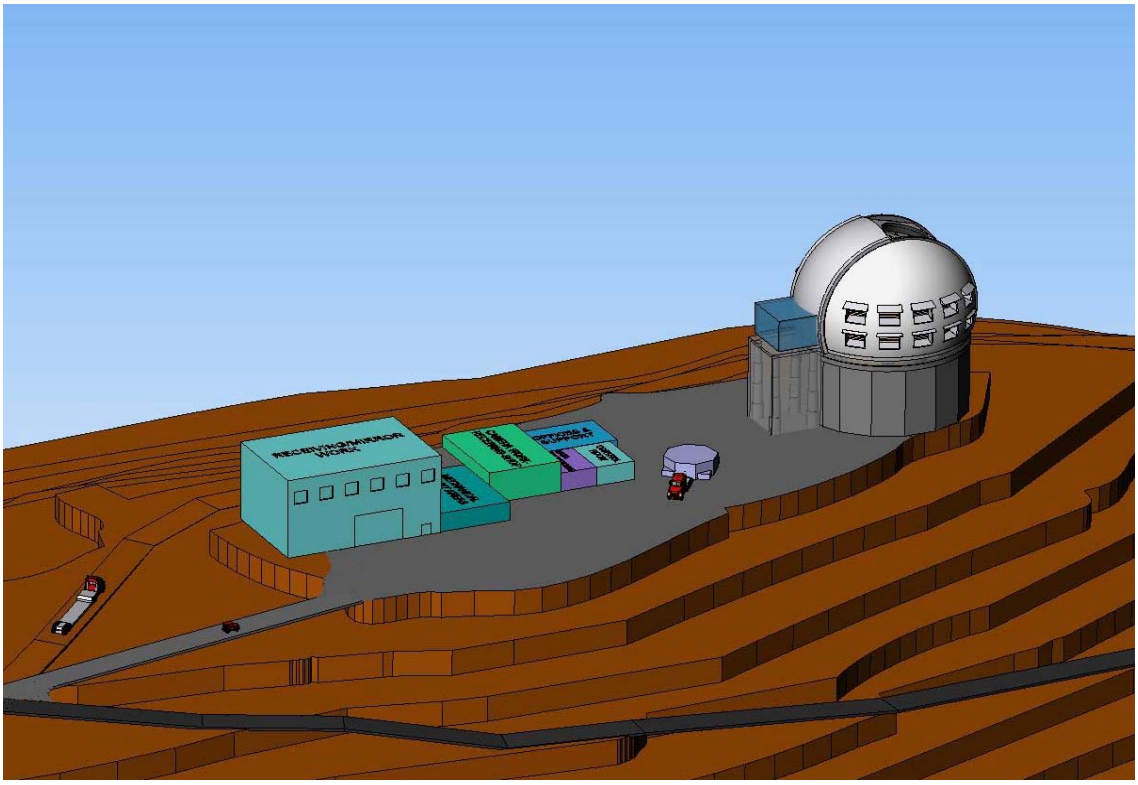


Figure 4.2.2-3 Potential Site for LSST at San Pedro Mártir Observatory

4.2.2.2.6.2 Cerro Pachón

The primary proposed site for LSST at Cerro Pachón is a peak named El Peñón, located ~1.5 km southwest of the existing Gemini Observatory. There are other locations in the vicinity which may also be appropriate, but the hilly topography and steep drop-offs that are characteristic of Cerro Pachón will impact the choice and the arrangement of the telescope enclosure and facilities. The layout shown in Figure 4.2.2-4 depicts the smallest envisioned LSST telescope enclosure (28 m diameter) and a minimal utility structure located at the peak. The rest of the summit support spaces, including the mirror coating facility, are shown in an adjacent saddle area. Significant cut and fill would be required to create adequate level platforms for all major structures. The common support areas for the Gemini and SOAR observatories (utility yard, dorm/dining facility) are close enough to the El Peñón site to be potentially shared by LSST.

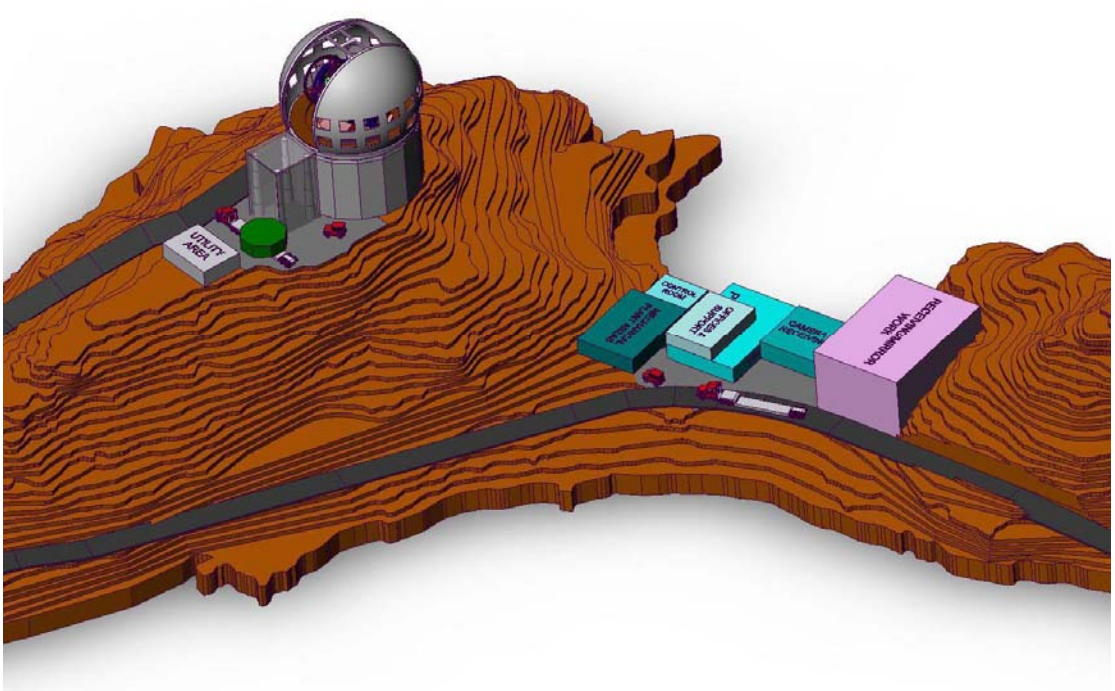


Figure 4.2.2-4 Potential site for LSST at Cerro Pachón

4.2.2.2.6.3 Las Campanas

There are multiple potential sites for a new large telescope along the Las Campanas ridge. The Magellan Telescope, the DuPont Telescope and others are currently located there and several of the largest unoccupied sites are being evaluated as candidate locations for the proposed Giant Magellan Telescope. The site that so far has been identified as likely to be available and appropriate for LSST is the current location of the 1 m Swope telescope. This small facility would be removed and some additional cutting of the hill would be required to provide an adequate platform for LSST. Figure 4.2.2-5 depicts the 32 m LSST enclosure and the summit support facility at that proposed location.

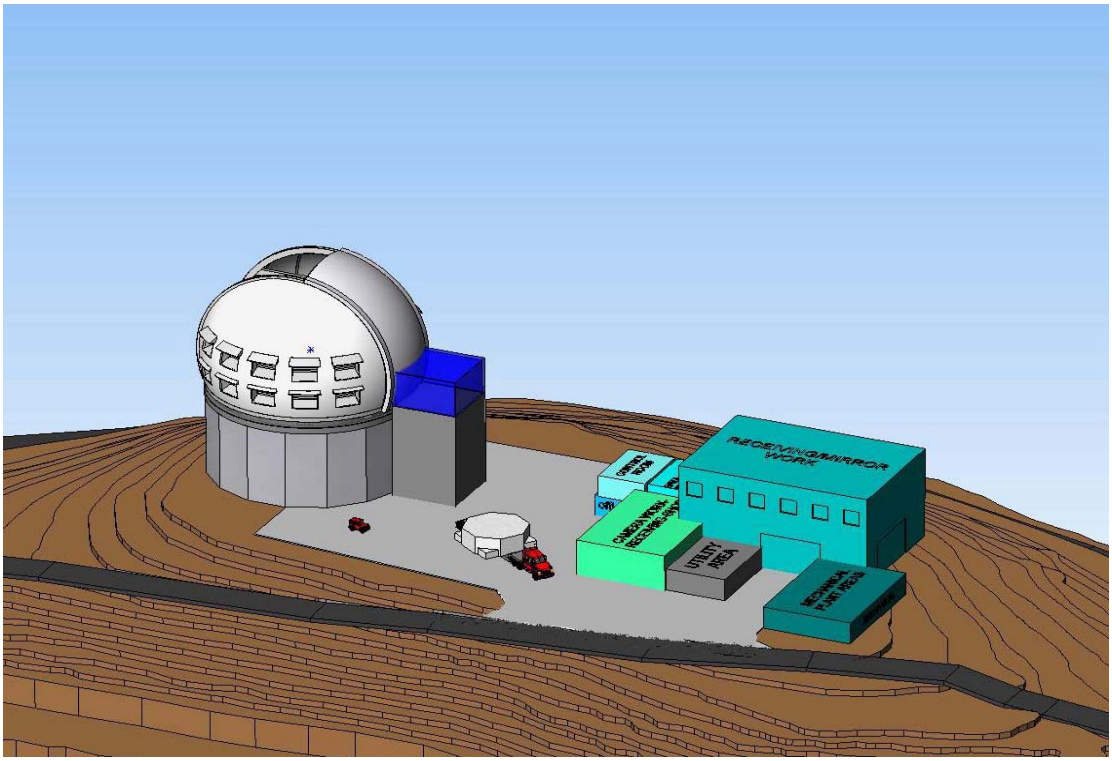


Figure 4.2.2-5 Potential site for LSST at Las Campanas

4.2.3 Telescope Mount

4.2.3.1 Overall Description

Optically the LSST telescope has a unique 3-mirror system. The primary mirror circumscribes the tertiary mirror such that both surfaces can be made into a single monolithic substrate (see Figure 4.2.4-1). The camera assembly is also circumscribed within the secondary mirror assembly, forming a convenient package at the telescope top end.

Although the LSST optical design is unique, it can be supported by a conventional telescope structural arrangement, Figure 4.2.3-1. A stiff mirror cell is used to support the primary and tertiary mirrors, and the top end assembly supports both the secondary mirror assembly and the camera assembly. Both the elevation axis and the azimuth axis are expected to utilize hydrostatic bearings, which are common on large telescopes.

The LSSTs structural arrangement facilitates maintainability. The primary/tertiary mirror cell is connected to the rest of the elevation assembly at four flange locations. This facilitates convenient removal and re-installation of the mirror cell for recoating and any significant maintenance needs. The top end assembly is also only attached at four flange locations to facilitate removal. The hydrostatic bearing surfaces are enclosed to reduce contamination and susceptibility to damage.

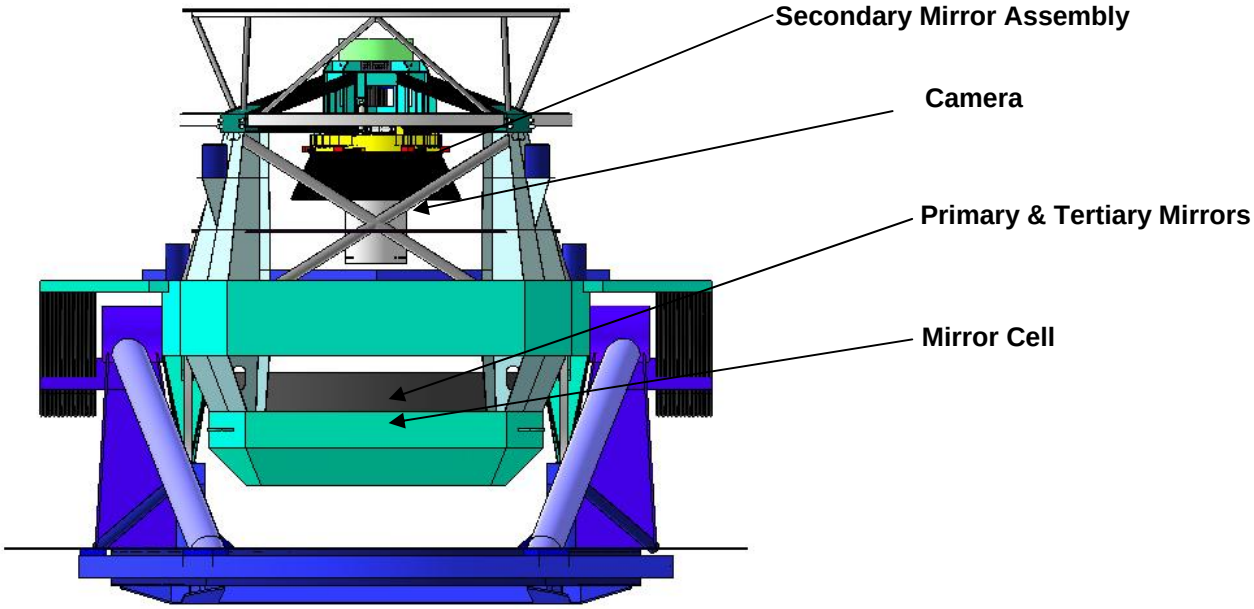


Figure 4.2.3-1 Front View of Telescope Assembly

The LSST design also incorporates many essential auxiliary components (Figure 4.2.3-2.). Among these are the baffle system, balancing system, damping system, mirror cover, cable wraps and motor drives. The mirror cell is a 2 m deep sandwich with access to the complex systems required for mounting and thermal control of the primary and tertiary mirrors.

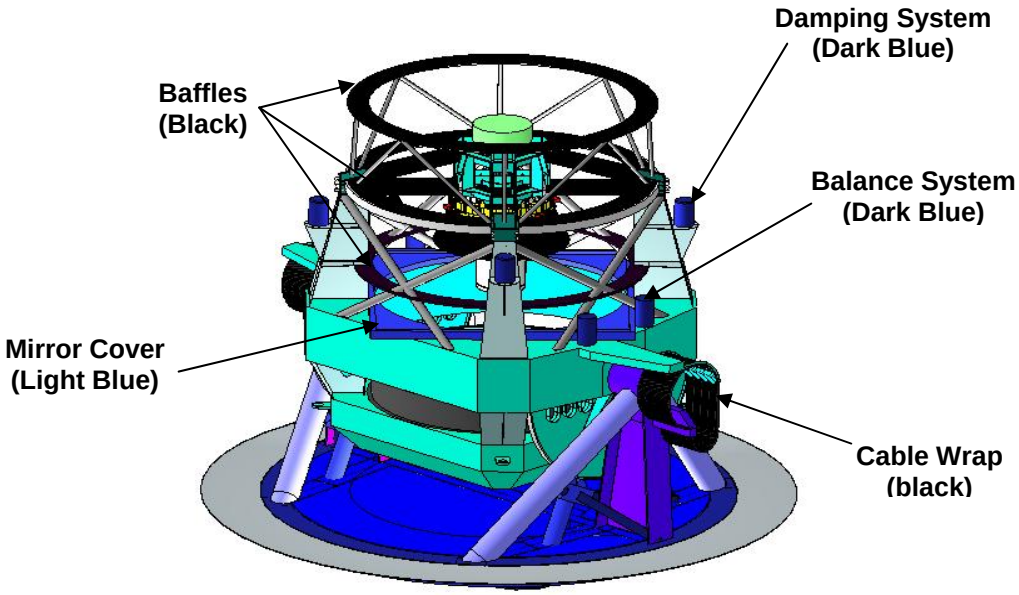
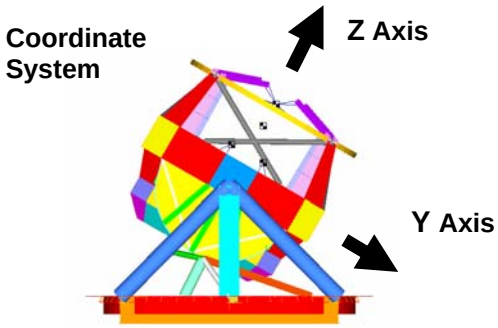


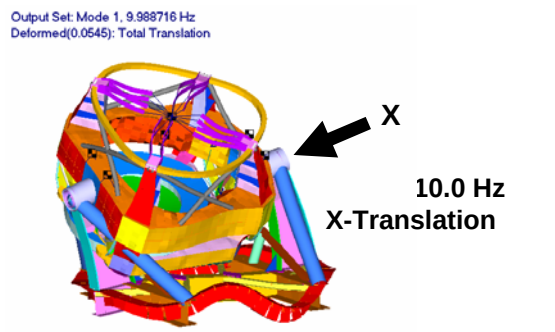
Figure 4.2.3-2 Auxiliary Telescope Components

4.2.3.2 Mount Analysis

Preliminary analysis determined that the lowest natural frequencies of the telescope assembly should be 10 Hz or greater to meet the slew and settling requirements. The telescope mount assembly was designed and analyzed with FEA, with the goal of meeting this 10 Hz requirement (Figure 4.2.3-3).

LSST Short Tube Design		
Vibration Characteristics		
No Parasitic Mass Added		
#	Hertz	Mode
1	10.0	Elevation Assembly X Translation
2	10.6	Elevation Assembly X Rot, Y Trans
3	13.7	PM Y Trans
4	15.7	PM Y Trans, TEA Z Trans
5	16.3	PM Y Rotation
6	17.5	Azimuth Ring Vibration
7	17.9	Azimuth Ring Vibration
8	18.6	Azimuth Ring Vibration
9	18.6	Azimuth Ring Vibration
10	18.6	Azimuth Ring Vibration





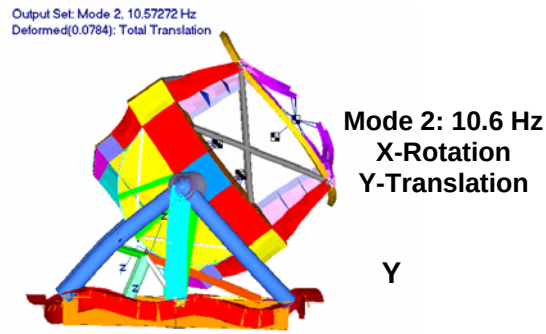


Figure 4.2.3-3 Mount Vibration Characteristics

The basic conceptual structure meets the 10 Hz requirement, however with the addition of 20% parasitic mass, a technique to model a more realistic case predicts first frequency at 9.1 Hz. The impact of reduced stiffness is a deficit in settling time. As the design develops, additional attention will be given to the stiffness and approaches to damping will be considered in order to achieve the 10 Hz requirement frequency for the design with the parasitic mass.

4.2.3.3 Top End Assembly

The top end assembly (Figure 4.2.3-4) supports the mass of the secondary mirror assembly and camera assembly through the use of 16 hollow rectangular spiders (Figure 4.2.3-5). These hollow spiders are structurally efficient, and the interior provides a convenient location to route the many cables required by the camera and the secondary mirror. These spiders have exterior dimensions of 300 mm x 50 mm and interior dimensions of 210 mm x 36 mm.

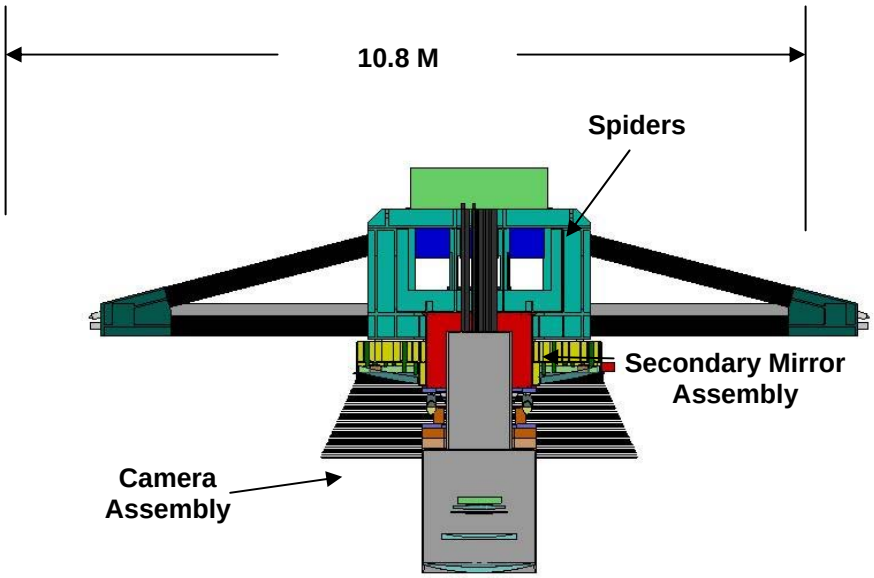


Figure 4.2.3-4 Section view of top end assembly.

The spiders are arranged to minimize the image degradation. All the spiders are arranged in axially aligned pairs. Consequently, the focal plane only sees eight spiders. The eight spider pairs are in a parallel/perpendicular arrangement, which only produces 2 diffraction spikes.

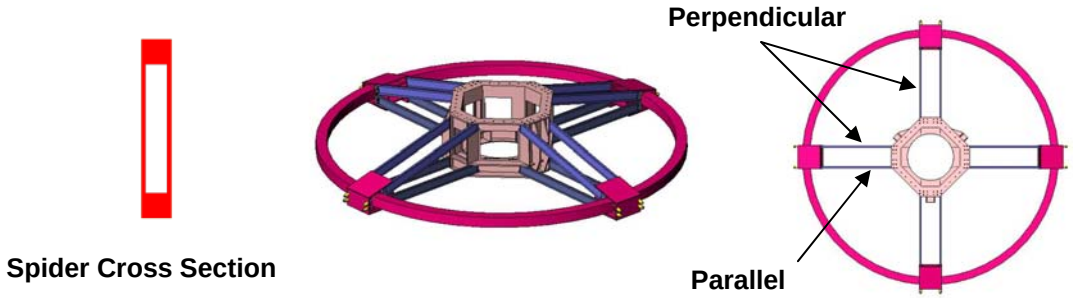


Figure 4.2.3-5 Top end assembly spider arrangement.

4.2.3.3.1 Instrument Assembly

The instrument assembly includes the camera, rotator, hexapod, cable wrap, integrating structure and electronics assemblies. The rotator is located between the hexapod and the camera to provide rotation about the optical axis during tracking. The hexapod resides between the rotator and integrating structure, and is used to provide alignment and positioning. The electronics assemblies mount to the interior of the integrating structure. The cable wrap resides on the top of the integrating structure as shown below.

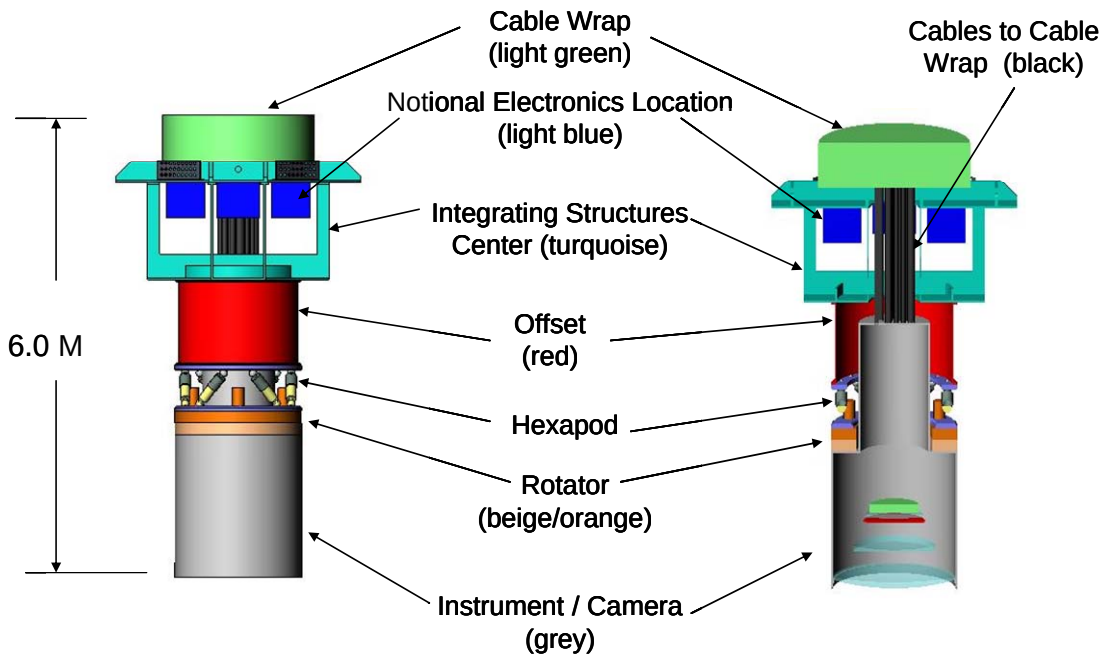


Figure 4.2.3-6 Instrument assembly

The entire instrument assembly can be installed and removed as a single unit. This allows the entire instrument assembly to be put together and tested before integration into the telescope. It also provides for the removal for service and repairs. This installation feature requires that all cabling for the camera be routed from the camera's top surface, through the hexapod and the cable wrap and to the integrating structures top surface.

4.2.3.3.2 Secondary Mirror Assembly

The baseline design for the secondary mirror assembly (Figure 4.2.3-7) is a 100 mm thick glass meniscus supported by 120 axial actuators and 6 tangent actuators and a structural cell for support. The entire secondary mirror assembly is attached to the top end spider spindle by 6 positioning actuators. The mounting system includes an interface plate to allow removal of the secondary mirror assembly without disconnecting the position actuators (shown previously in Figure 4.2.3-1). The secondary mirror assembly also incorporates a large baffle (not shown). The design of the secondary mirror system is discussed in more detail in Section 4.2.4.2.

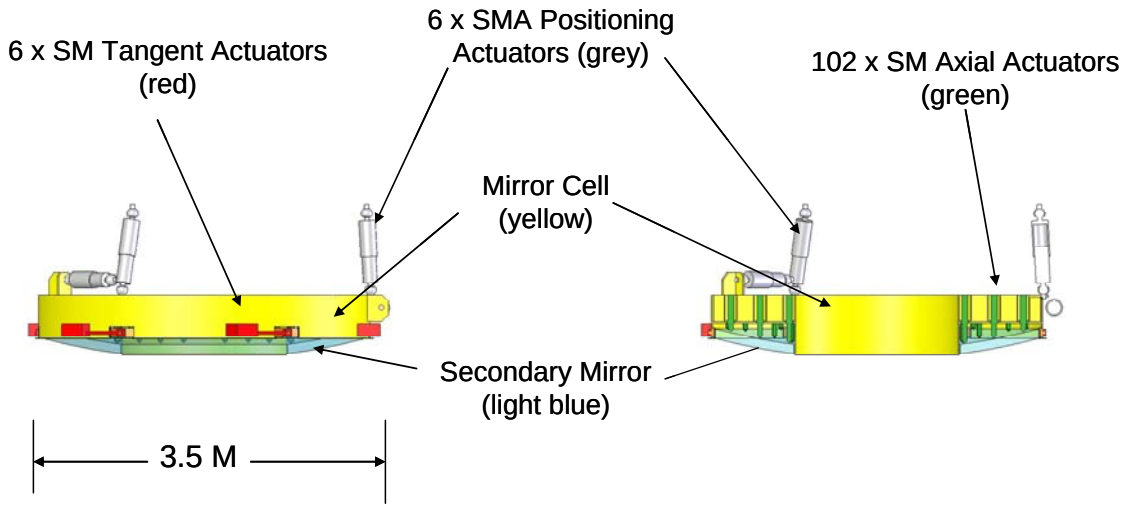


Figure 4.2.3-7 Secondary mirror assembly

4.2.4 Optical System

The LSST optical system consists of three reflective elements and three refractive lenses that collect and focus the light through science filters onto the detector. The telescope provides the three mirrors for the system and the three lenses and filters are incorporated into the camera. The physical size and precision requirements for the mirror system are challenging at each surface. The optical system quality is specified so the telescope only degrades the best quartile seeing by 10%. The fast f/1.2336 LSST system demands very tight tolerances on rigid body positioning of the mirrors and camera with respect to one another. Table 4.2.4-1 below summarizes the significant parameters that describe the three mirrors and Figure 4.2.4-1 and Figure 4.2.4-2 present the three-mirror design and optical ray trace over the +/-3.5 degree field of view.

Table 4.2.4-1 Parameters for LSST Optical System

	Outer Diameter	Inner diameter	Radius of curvature	Shape	Conic Constant	Aspheric departure
M1	8.36m	5.116m	19.835m	Concave	-1.215	111μm BFP
M2	3.4m	1.8m	6.8m	Convex	-0.222	17μm BFS
M3	5.0m	1.0m	8.4m	Concave	0.155	403μm BFS

	50% EE contribution to error budget (FWHM @ 500nm)	RMS Surface Error	R0 for mirror fabrication structure function at Zenith	R0 for mirror support structure function at Zenith
M1	0.11”	< 30nm	92cm	212cm
M2	0.07”	< 30nm	145cm	225cm
M3	0.07”	< 30nm	151cm	225cm

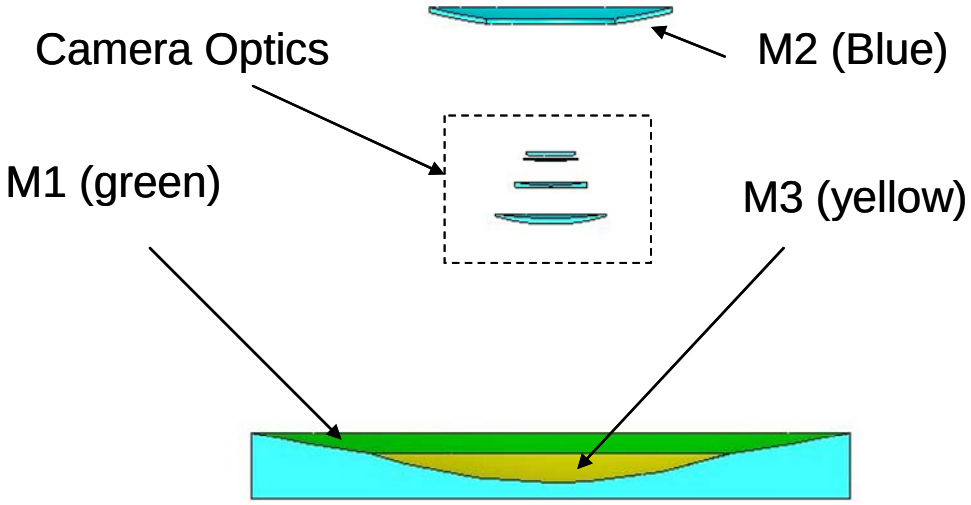


Figure 4.2.4-1 Telescope Optical Design

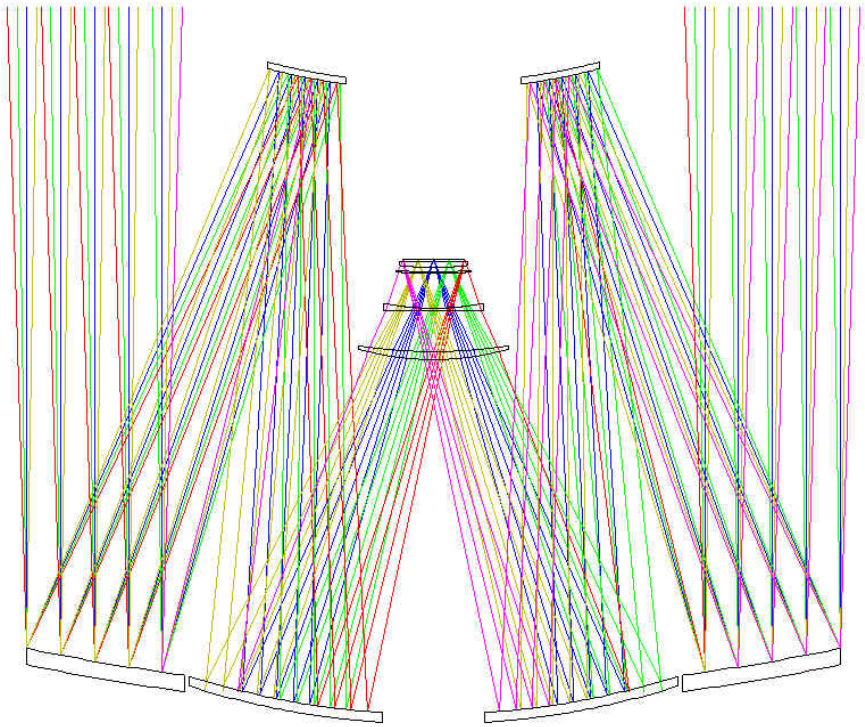


Figure 4.2.4-2 Telescope Ray Trace

4.2.4.1 Primary/Tertiary Mirror

Since the optical surfaces of the primary and tertiary mirrors are nearly continuous, the two mirrors will be fabricated from a single monolithic substrate, as shown below in Figure 4.2.4-3. Any misalignment of the two optical surfaces will be permanent so proper positioning of these surfaces during figuring is critical. Producing the two mirrors as a single monolith shifts the positioning burden from operation to fabrication and from a continuous operational alignment to a one-time effort. This monolithic design is a significant departure from previous telescope experience but the design has undergone extensive evaluations. The more conventional separate mirror design would have required independent edge sensing, and an additional feedback loop, to maintain the relative position of the optics. Installation and removal of the individual optics would also be problematic.

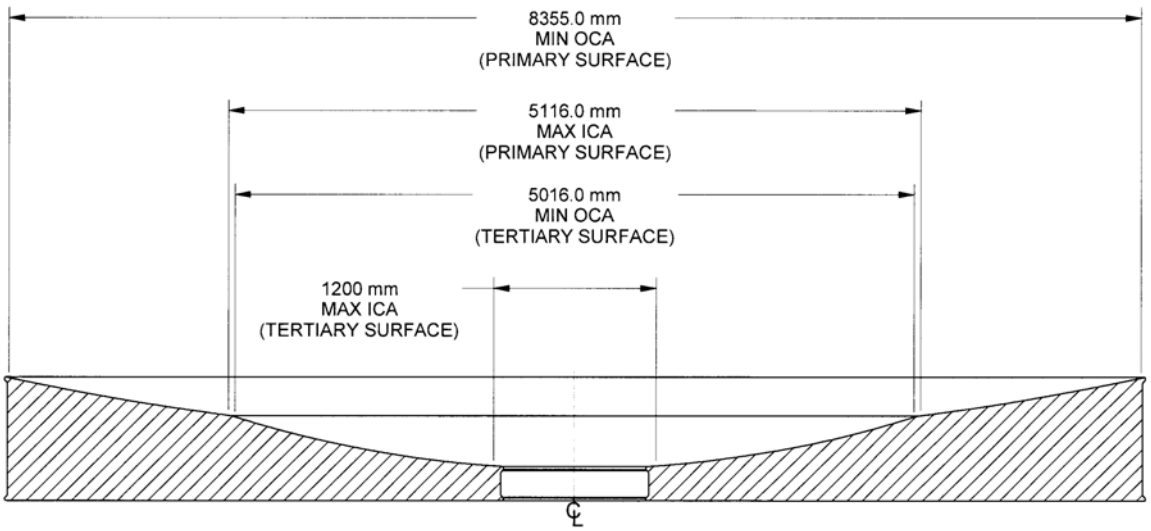


Figure 4.2.4-3 Primary/tertiary monolithic mirror optical surfaces

The LSST primary/tertiary mirror has already been ordered from the University of Arizona. Through a generous private donation this long lead item has been procured from the U of A Steward Observatory Mirror lab. The lab will produce the LSST primary/tertiary mirror (Figure 4.2.4-4) using its borosilicate casting process, which produced mirrors for the Large Binocular Telescope at the same 8.4 m outer diameter, the two Magellan Telescope mirrors at 6.5 m in diameter, the MMT, WIYN and others. The final design of the mirror will be completed during the initial contract phase but significant initial evaluation has been performed to verify that the borosilicate mirror is appropriate for the LSST. These studies focused on two major issues, thermal control and mirror support. The relatively high thermal expansion coefficient of the glass requires that strict thermal control systems be used in operation. Even with these systems operational, the mirror is shape controlled for active figure correction.

The LSST mirror differs from its more traditional counterparts in its unique multi curvature shape. However, it has been shown to be manufacturable and controllable with similar processes and support systems as used in previous borosilicate mirrors.

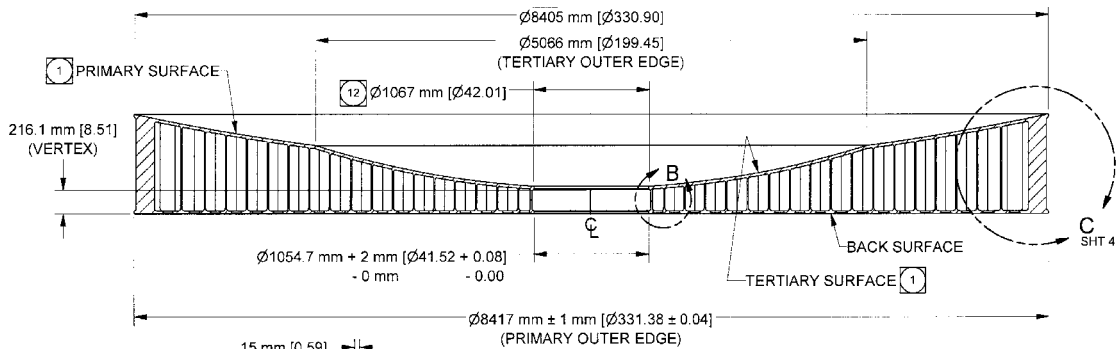


Figure 4.2.4-4 Primary/tertiary monolithic mirror design

The baseline primary/tertiary mirror design uses the standard hex cell pattern so the existing load-spreader designs can be used without modification. The actuators to support the mirror are single and dual axis type to manage both the axial and lateral loads of the 17,600 kg mirror. Table 4.2.4-2 shows the force repeatability of the existing actuators, which were used to model the LSST Mirror. The results of the primary/tertiary structural analysis demonstrates that the mirror can meet the required surface error requirements for both thermal distortion and gravity induced distortions

Table 4.2.4-2 Actuator error forces measured on the baseline LBT actuators. Each value is the rms of the actuator error forces. The error force for an actuator is the maximum measure on that actuator.

Error Force	Dual Axis		Single Axis	
	Repeat	Non-repeat	Repeat	Non-repeat
F_x	1.33	0.10	0.57	0.09
F_y	1.04	0.15	0.65	0.09
F_z	0.99	0.17	0.37	0.10
M_x	35.94	1.27	14.94	0.43
M_y	29.34	0.81	15.81	0.57
M_z	23.08	0.60	0.95	0.39

4.2.4.2 Secondary Mirror

The LSST Secondary mirror is a 3.5 m diameter convex asphere that must be supported and positioned in the top end assembly. To allow the assertion of the camera, the secondary mirror must have a large central hole, D= 1.8M.

The baseline for the secondary mirror is a 100mm thick meniscus, Figure 4.2.4-5. This configuration meets all of the requirements including gravity induced error, mass budget and the design envelope. This type of design was utilized on the two most recent optics of the same class. Both the 4.3 M primary mirrors of the SOAR and DCT telescopes used this type of construction.

The 100mm thick baseline meniscus design is supported by 102 axial supports and 6 tangent actuator supports, Figure 4.2.4-6. All the supports will be active to provide figure control. Most likely all the actuators for these supports will be electromechanical. Hydraulic supports would produce a significant leak risk. It is common for hydraulic mirror support systems to leak. For the LSST the secondary mirror is located above the primary/tertiary mirror. Since hydraulic fluid attacks metal, any leaks from the secondary mirror supports would damage the primary tertiary

optical coating. Pneumatic actuators are an option; however, they are generally less accurate and more flexible than electromechanical actuators.

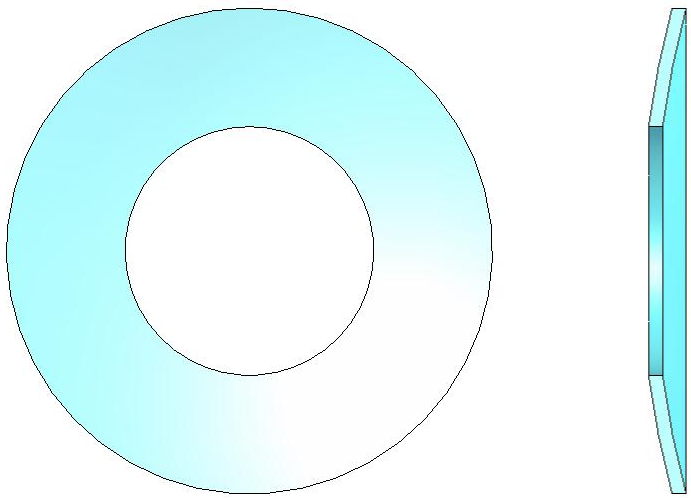


Figure 4.2.4-5 Secondary Mirror Baseline Design, 100 mm Thick Solid Meniscus

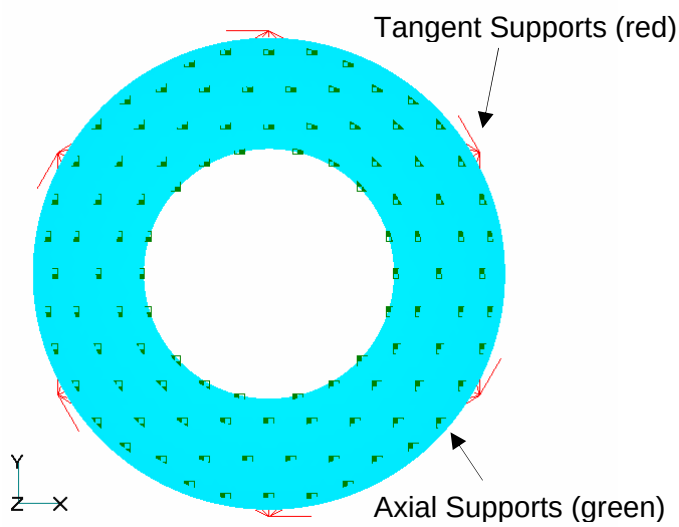


Figure 4.2.4-6 Secondary Mirror, Active Support Locations

To evaluate the cost effectiveness of light weighting the secondary mirror, a machined light weighted version of the 100mm thick meniscus is also under evaluation, Figure 4.2.4-7. Utilizing this structured meniscus design would reduce the mass of the secondary mirror assembly by 50%. This design would also have superior thermal properties. Meeting the gravity induced surface

error requirements would require the addition of another 6 tangent supports. Otherwise the supports of the structured meniscus would be nearly identical with the solid meniscus. Although this lightweight design has generally better characteristics, the increase in performance would come at a significant increase in cost. Producing a structured mirror by machining is very costly. Consequently a solid meniscus which meets all requirements is the baseline approach until system design indicates weight is an issue. Since the mirror is convex, the spin casting methods used for the primary/tertiary cannot be utilized. These methods require a balance between centrifugal and gravitational forces.

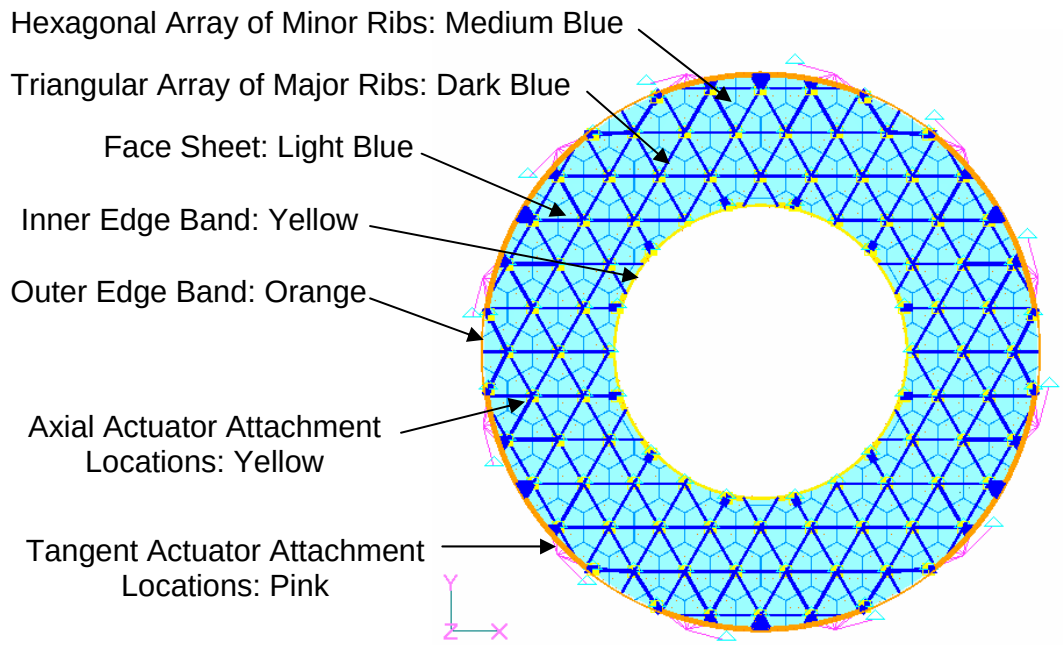


Figure 4.2.4-7 Secondary Mirror Design, Structured Meniscus Option

4.2.4.2.1 Secondary Mirror Testing

A particular challenge for the secondary mirror is the fabrication of the optical surface. The size, accuracy and convex shape of the surface offer challenges to the state of the art in optical metrology. A feasibility study has been conducted at the Optical Sciences Center to investigate two aspects of the optical metrology, profilometry and sub-aperture optical testing. The convex optical surface reviewed in this study possessed a few hundred waves aspheric departure. The profilometry study consisted of setting up and demonstrating performance of a 4 meter swing arm profilometer. For the optical testing a metrology set-up was designed and analyzed, and a small scale demonstration was performed.

The swing arm profilometer is a mature device developed at the University of Arizona, and now implemented in several optical shops. Figure 4.2.4-8 shows the measurement of a 1.7-m convex secondary mirror. Also shown are the measurements taken on that machine and then verified by an independent interferometric test. The accuracy of the surface profile is about 0.1 μm rms.

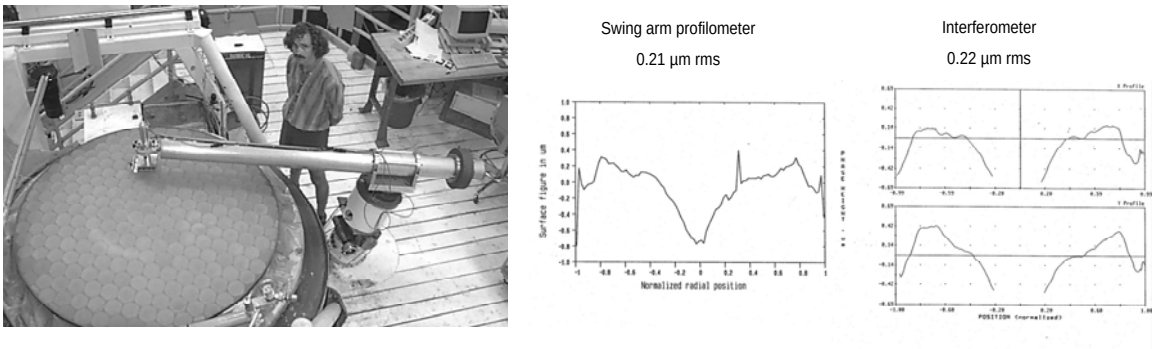


Figure 4.2.4-8 The swing arm profilometer has been used to measure giant secondary mirrors, such as the 1.7-m mirror shown here. The accuracy of this instrument is demonstrated by comparing data from the profilometer with an interferometric measurement.

Two new capabilities were added to the profilometer to meet requirements for LSST: rotational scanning was added to allow the creation of full surface maps, and the system was scaled up to allow measurement of 4-m optics.

The software for combining the scans was demonstrated using a simulation that included 0.1 µm rms noise. As shown in Figure 4.2.4-9, the data are successfully combined to produce an accurate map of the surface.

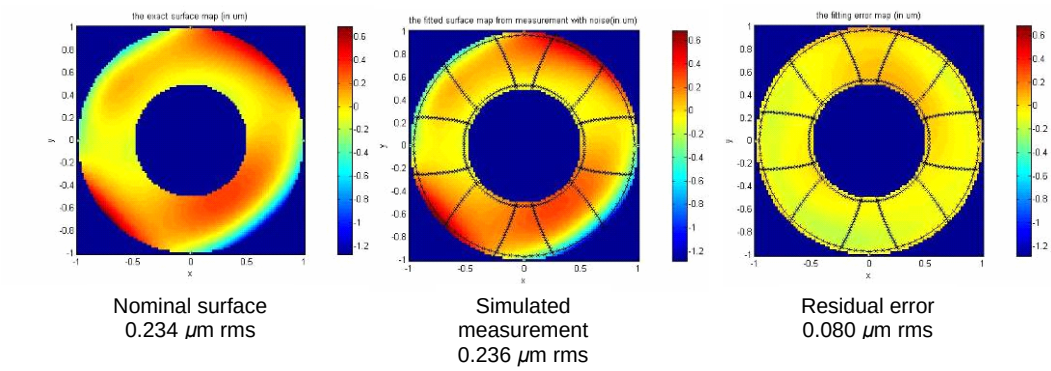


Figure 4.2.4-9 Demonstration of software for combining multiple swingarm scans to create a surface map. The figure shows (*left*) a nominal surface with simulated errors, (*center*) a simulation of a measurement of the surface using a combination of 6 swingarm scans and 2 circumferential scans, with 0.1 um rms simulated measurement noise added at each point, and (*right*) the computed residual error.

The scale-up of the profilometer to 4 meters was completed, including the precise rotation of the mirror for creating surface maps. The system was demonstrated to meet LSST requirements by measuring a 1.8-m diameter mirror, shown in Figure 4.2.4-10. These results demonstrate the technology necessary for efficient shaping of the LSST secondary mirror.

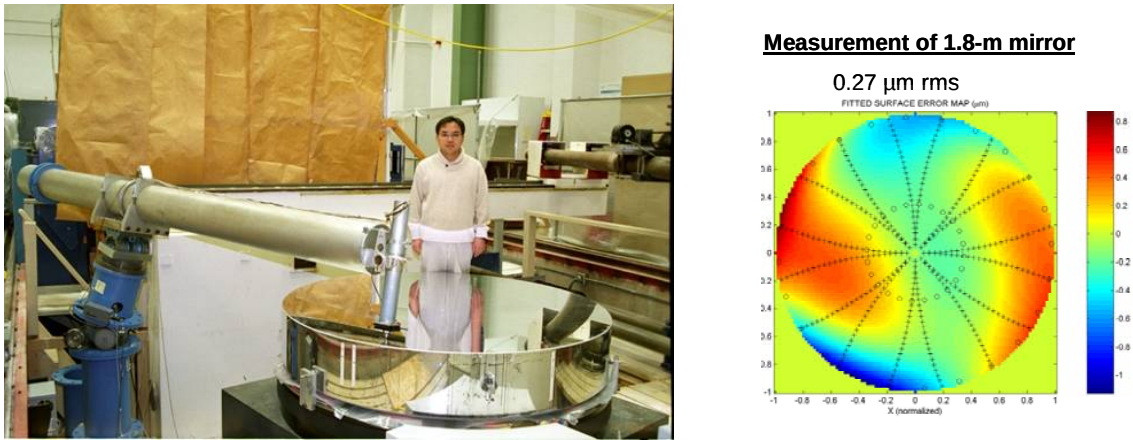


Figure 4.2.4-10 Demonstration of 4-m profilometer, measuring a 1.8-m mirror

The polished surface of the LSST secondary mirror can be measured interferometrically using a subaperture test plate. The convex surface would conventionally be measured in a Hindle test or equivalent. This would require test optics as large as the secondary itself making it difficult and prohibitively expensive to measure. The concept chosen for the baseline LSST approach is to perform a series of sub aperture tests and stitch the results together for a full accounting of the optical surface. Figure 4.2.4-11 shows the set-up design for such a test. It utilizes a special phase shifting polarization interferometer that is insensitive to vibrations and a test plate that has a calibrated matching aspheric surface to the secondary mirror. A small scale experiment using this set-up was performed yielding results shown in Table 4.2.4-3.

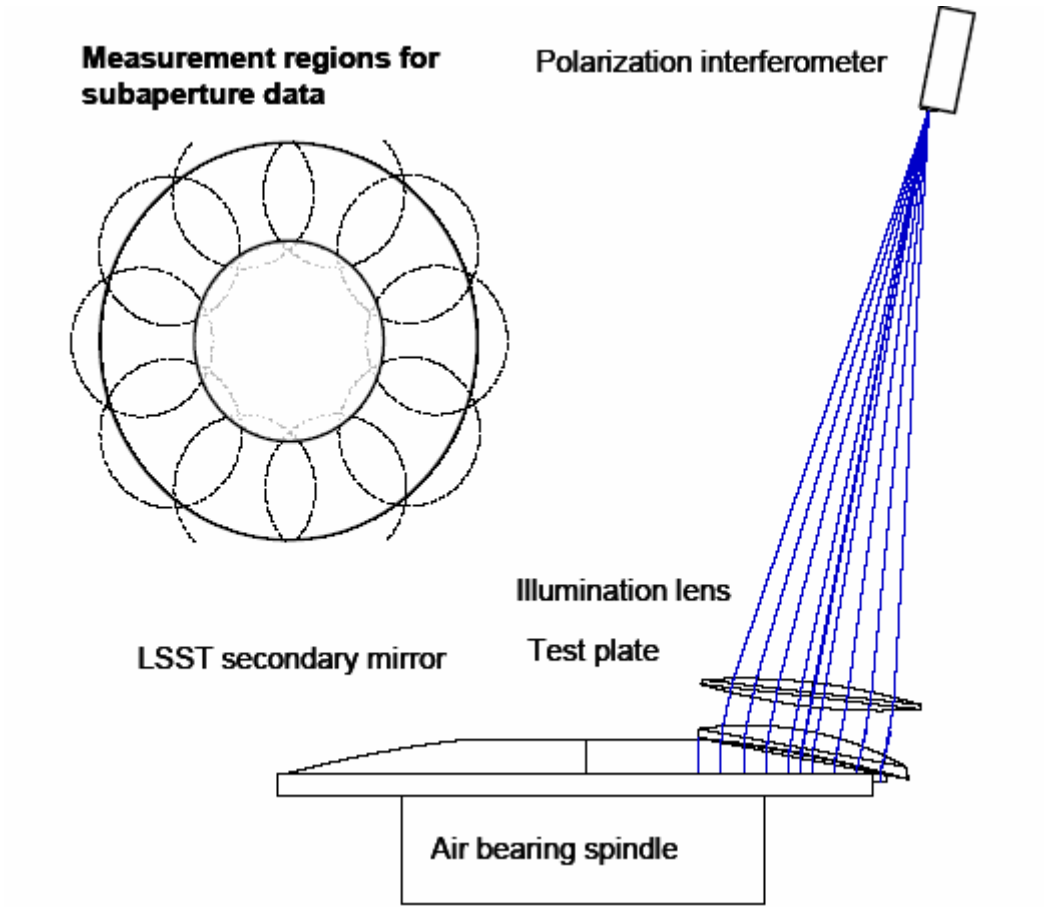


Figure 4.2.4-11 Sub aperture Interferometric test layout

The use of a large Fizeau reference with the polarization interferometer was demonstrated by measuring a 12-in flat. The configuration for the demonstration is shown in Figure 4.2.4-12 and the data that proves this concept is shown in Figure 4.2.4-13.

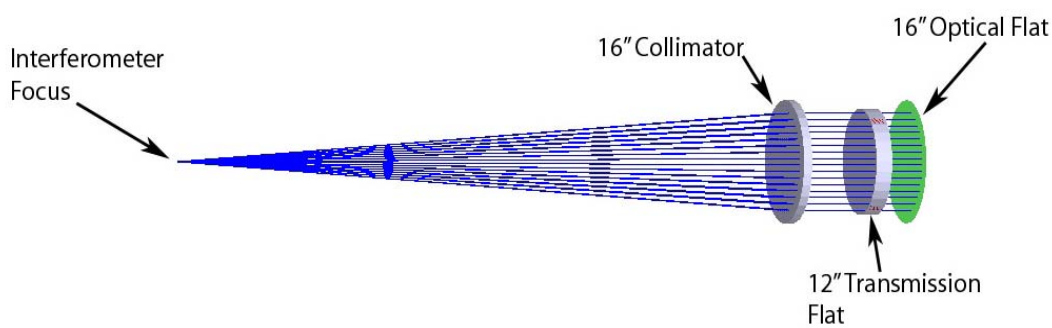


Figure 4.2.4-12 Layout for demonstration of vibration insensitive Fizeau test.

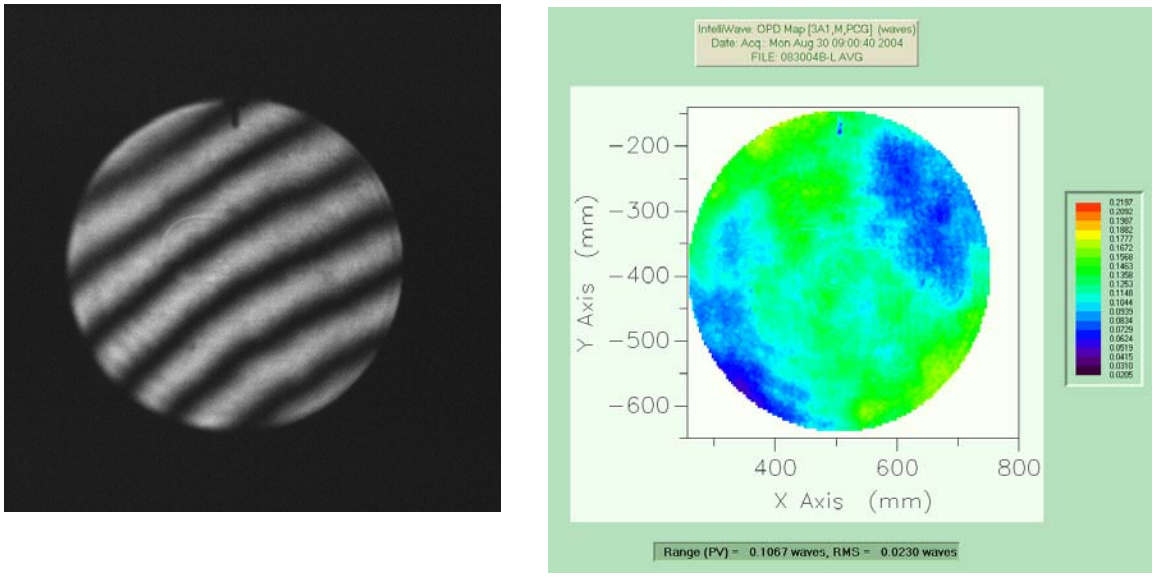


Figure 4.2.4-13 Interferogram and reduced surface data provided by the vibration insensitive Fizeau interferometer.

An error budget for the interferometric test was developed, and is summarized in Table 4.2.4-3. This shows that the surface measurements will meet the LSST requirements.

Table 4.2.4-3 Summary of the error budget for LSST interferometric surface test.

a- Measurement of aspheric test plate with CGH null lens: 5 nm rms

b- CGH distortion, assuming 1000 µm aspheric departure, 0.1 µm CGH distortion : 3.5 nm rms
3 nm rms residual from CGH substrate errors, after correcting for them (typically ~30 nm rms low order)

c- fixed errors in the interferometer : 2 nm rms

d- Test plate figure stability : 5 nm rms
Budget for change in support forces. Ref surface to be measured off line as supported.

e- Limitation backing out test plate surface: 4 nm rms.
Allow ref surface to have slope errors of 20 nm/cm. This couples with 2 mm knowledge uncertainty for mapping test plate data onto LSST mirror measurement.

f- Illumination : 5 nm rms
The imperfect illumination optics cause a second order effect. Allowing 1 mrad slope errors in this system couples as a cosine with the 10 gap to cause 5 nm rms in the surface measurement.

g- Test alignment/stitching : 4 nm rms
Budget for overall error from stitched subapertures. Stitching has been demonstrated to this level by several groups

Total: 10.5 nm rms (RSS)

The study has demonstrated with a combination of hardware and analysis that the measurement of the LSST secondary mirror is feasible.

Since this study was completed, subsequent expanded telescope optical design optimization has reduced the overall secondary mirror aspheric departure to less than 20 waves over its usable aperture. This technical achievement should further diminish the overall risk of fabrication and metrology for this mirror. Additional studies with credible optical fabrication vendors will be undertaken to examine the predicted cost and schedule to manufacture the secondary mirror system.

4.2.4.3 Alignment Sensitivity

The alignment of the primary, secondary, and tertiary mirrors, as well as the camera assembly with its three refractive lenses, is particularly critical to the image quality of the fast f/1.2236 LSST optical system. Sensitivity studies have been performed to quantify the impact of each potential rigid body alignment error to determine the impact on achieved optical quality.

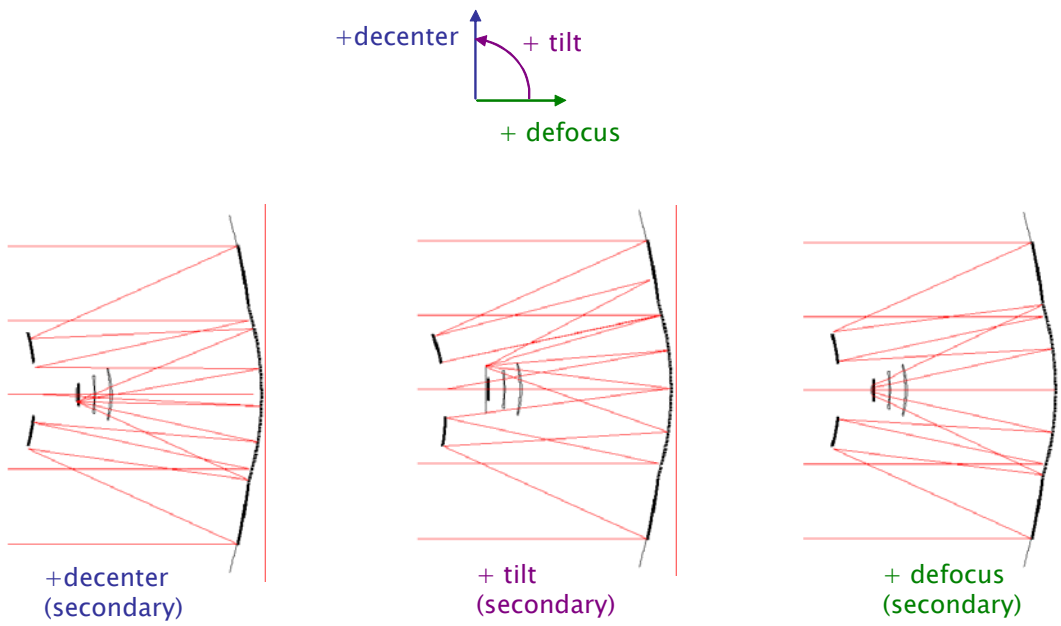


Figure 4.2.4-14 Rigid body motions and sign conventions

The perturbations applied to each element were lateral decenter, axial decenter (defocus), and tilt about an axis perpendicular to the optical axis (Figure 4.2.4-14). Nine field points, distributed across the entire field of view, were used to assess image quality, boresight error (for tilt and decenter), and, for defocus, the resulting plate scale variation (). Image quality was assessed as the growth in the blur diameter (in microns) for both the 50% and 80% diffraction encircled energy spots. Boresight or line of sight error is reported as the tilt (in milliradians) of the telescope assembly required to re-center the on-axis ray on the focal plane. Positive field angles introduce an image shift in the positive direction, so the apparent field angle is *opposite* the reported boresight error. Plate scale error was assessed as the change in the effective focal length of the telescope.

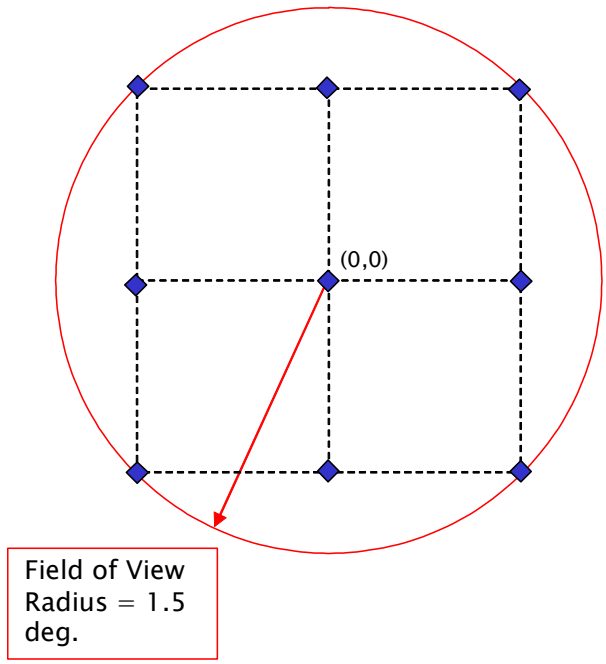


Figure 4.2.4-15 Field points used in sensitivity analysis. All points were equally weighted.

In calculating sensitivity coefficients, the spot size growth is assumed to be purely quadratic or linear over the range of motion evaluated. The coefficients are based on average system performance over all bands and all field angles. The range values associated with the coefficients are based on those used in the calculation. It may be that in actual practice the range over which the coefficients are valid extends beyond these limits.

As an example of this analysis, the variation of the spot diameter for 80% encircled energy versus decenter of the components is shown on Figure 4.2.4-16. The degradation in imaging performance is quadratic for small decenter. In other words, the spot size growth is proportional to the square of the perturbation. The sensitivity coefficients are given in Table 4.2.4-4 The decenter sensitivity is highest for the primary mirror. We present also the sensitivity results for defocus (Figure 4.2.4-17). The sensitivity coefficients are quite high, meaning that small perturbations lead to rapid change. It also shows that moving any one major component is as effective as any other for focus compensation.

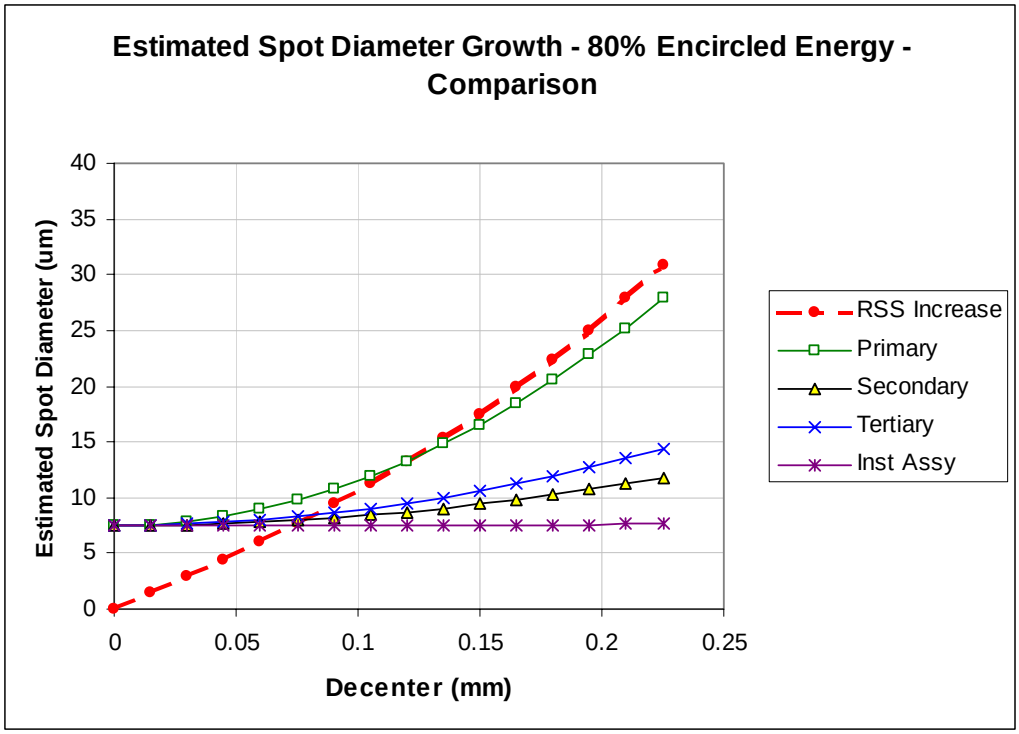


Figure 4.2.4-16 80% EE Spot Diameter versus decenter

Table 4.2.4-4 Sensitivity Coefficients

Element	Quadratic Coefficient	Linear Coefficient	Constant
Primary Mirror	401.89	0.000	7.5
Secondary Mirror	84.39	0.000	7.5
Tertiary Mirror	136.67	0.000	7.5
Instrument Assembly	2.42	0.000	7.6

Table 4.2.4-5 Ranges for valid quadratic approximation

Element	Decenter	Defocus	Units
Primary Mirror	0.2500	0.0300	mm (+/-)
Secondary Mirror	0.5000	0.0300	mm (+/-)
Tertiary Mirror	0.5000	0.0300	mm (+/-)
Inst Assy	1.0000	0.0300	mm (+/-)

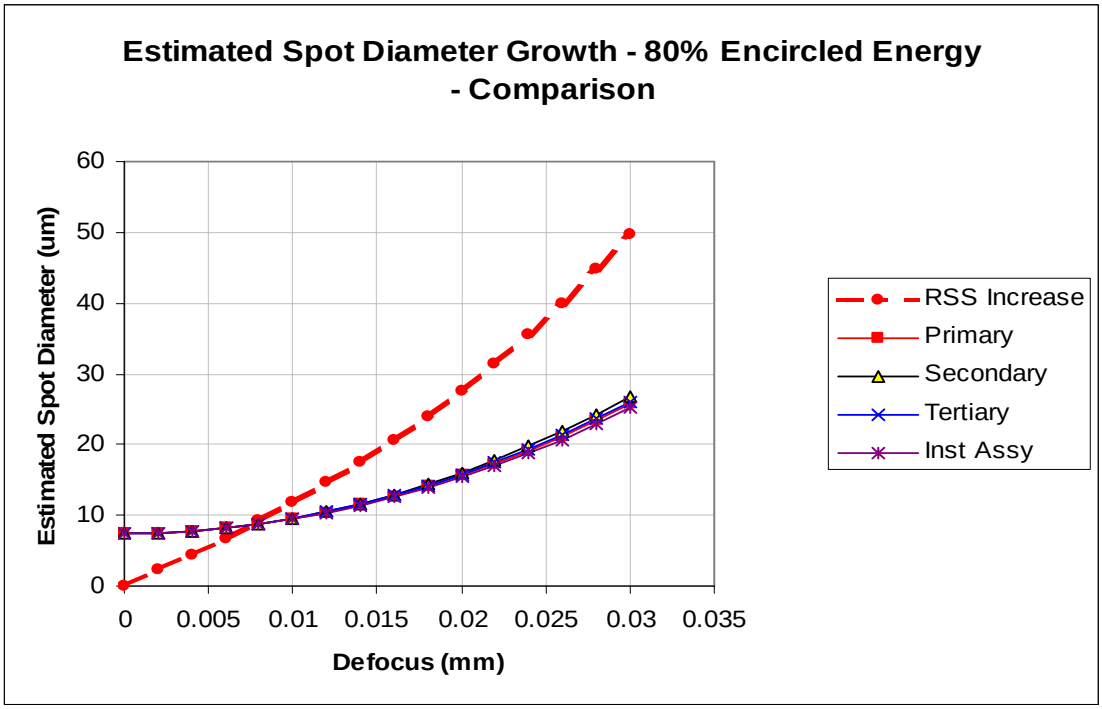


Figure 4.2.4-17 80% EE Spot Diameter versus defocus

4.2.4.4 Stray Light

Photon Engineering Inc. conducted a preliminary analysis to evaluate baffle concepts and to identify possible baffle locations and their impact on fractional throughput. Of concern is the identification of specular illumination paths that reach the image surface. To complete this task, an optical system model was imported into FRED, Photon Engineering's commercially available optical engineering software. FRED is a generalized non-sequential raytrace analysis program that allows for the creation and analysis of optical and opto-mechanical system models.

The FRED model of the Short Tube configuration is shown in Figure 4.2.4-18. Of particular note is the interface zone between the primary and tertiary mirrors. This zone is an annular ring 4 cm wide that separates the two mirror surfaces. Any rays striking the interface zone are halted. The secondary baffle has been sized and shaped to approximately match the interface zone radius and the convergent cone of light reflecting from the primary mirror. The large size blocks direct illumination of the tertiary mirror from illumination by objects within the field of view.

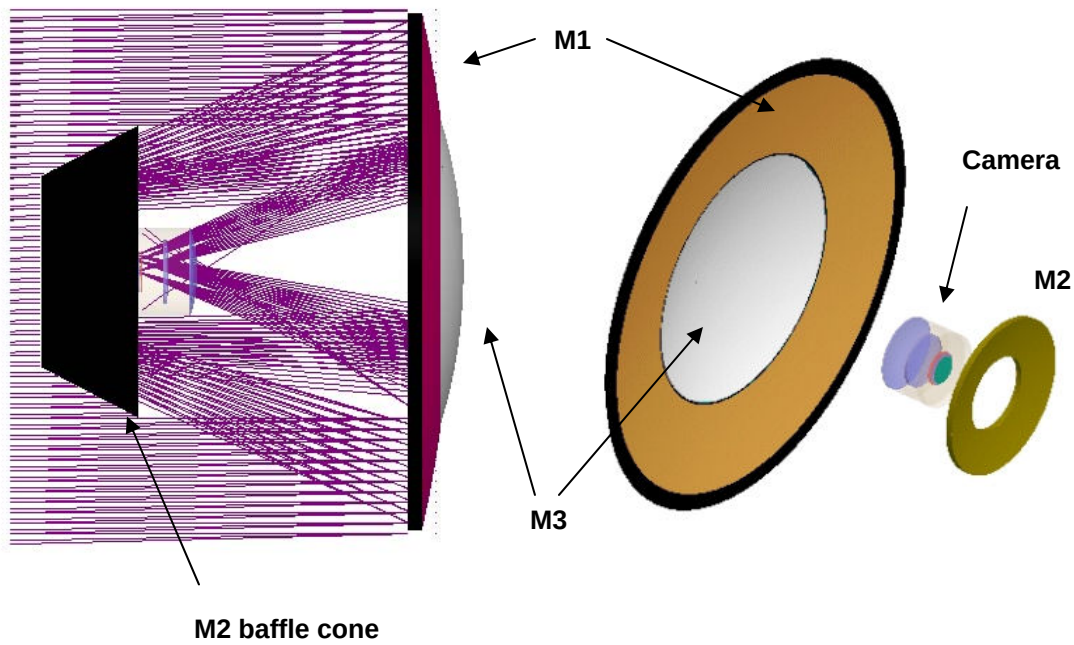


Figure 4.2.4-18 LSST model in FRED

An efficient way to identify potential stray light paths is to launch rays from the detector surface backwards through the optical system. Doing so provides a look at what the detector 'sees.' A spurious specular path from outside of the field of view is identified with this technique, as shown in Figure 4.2.4-19. Light incident on the telescope with an object space angle between 12 to 20 degrees (blue colored beam) strikes the primary mirror and enters directly into the camera with no reflection on the secondary and tertiary mirrors.

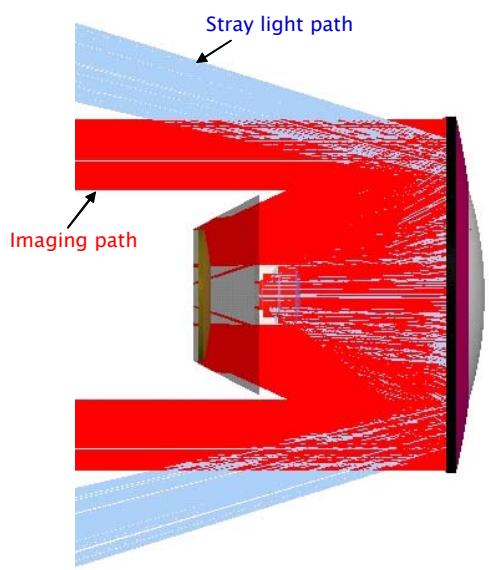


Figure 4.2.4-19 Stray light path (blue rays) and imaging path (red rays)

Mitigation of this specular stray light path is possible with the addition of a series of large annular baffles extending outward from the primary mirror. Figure 4.2.4-20 shows a rendition of the concept. The forward baffles are equally spaced at intervals of 3.5 m. The outermost baffle is 10.2 m from the vertex of the primary. Its outer diameter is 10.2 m and its inner diameter is 8.8 m. The innermost baffle is 3.2 m from the vertex of the primary mirror. Note also a baffle ring surrounding the primary mirror itself, which blocks ray paths that overfill the mirror surface. A configuration of baffles in this arrangement does have a number of advantages over a full-length telescope tube: they are relatively small and lightweight, they do not impede airflow over the mirror surfaces, and they could be designed to permit easy access to the instrument assembly.

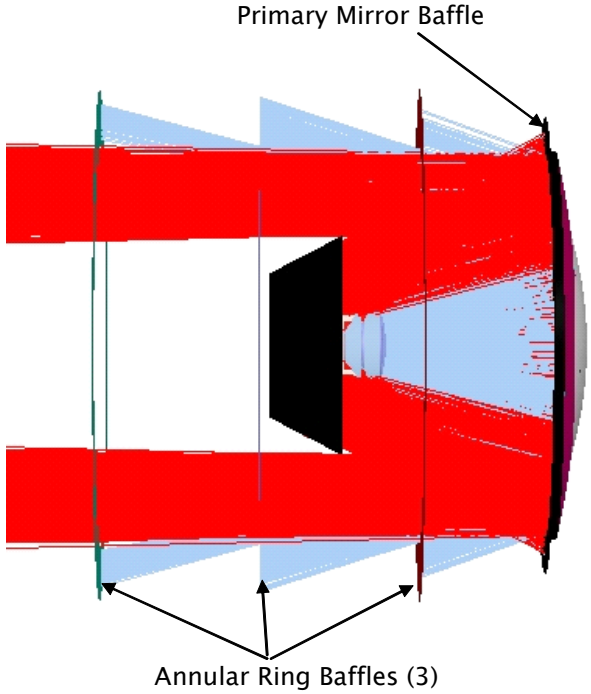


Figure 4.2.4-20 Annular ring baffles to block undesirable specular ray paths.

4.2.4.5 Mirror Coatings

The reflective surface quality of the three large LSST mirrors is important to the overall throughput and final signal to noise ratio achievable in the images. The system throughput is proportional to the surface reflectivity to the third power. Combined with the broad spectral range of the LSST this will present significant challenges for the coatings, and particularly at the blue end of the response. In addition to the initial, fresh coating quality, the durability and routine cleaning of the surfaces are also very important issues. The telescope down time associated with coating a large mirror can be roughly 1 week per mirror, removal from the telescope, cleaning, coating, installing and realigning included. The facility and operations plan for LSST will aim to reduce this down time and maximize the effort so all three large mirrors are addressed in an efficient parallel manor but there will clearly be a premium on coating longevity to limit the number of recoatings required.

Figure 4.2.4-21 below shows the throughput as a function of wavelength for three reflections of the identified coating materials. The LLNL coating described is a wide band durable silver coating developed at Lawrence Livermore National Labs with good results at 1 meter class size. The protected Ag curve is a silver coating with a dielectric protection that has recently been demonstrated at the Gemini Observatory on their 8 m mirror. The “fresh” and “aged” bare curves

are values for aluminum. Also shown in the chart are the 5 filter bands in the baseline LSST design.

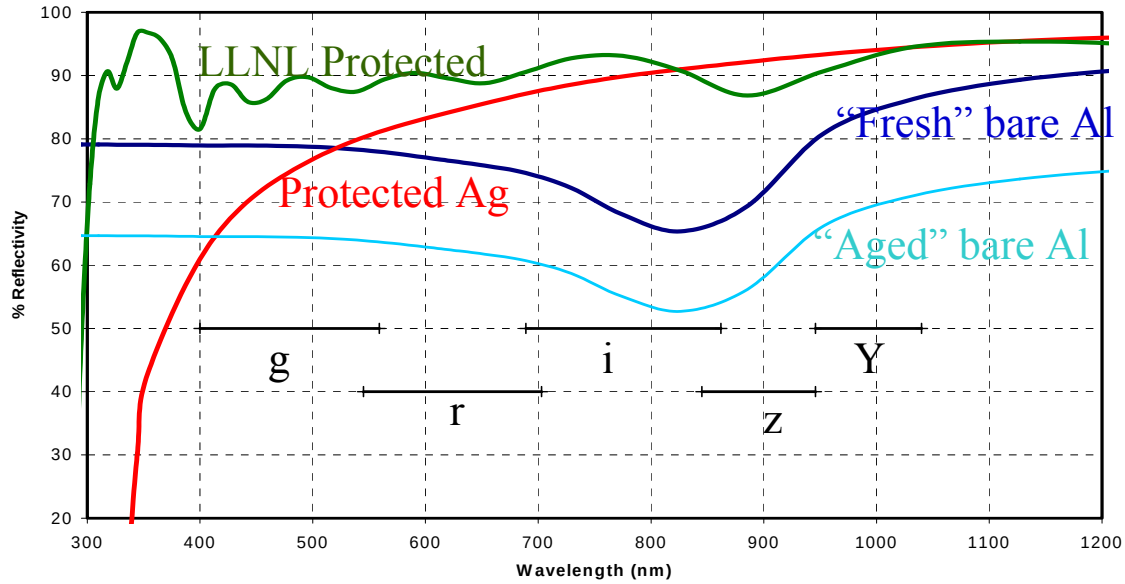


Figure 4.2.4-21 Reflectivity curves after three reflections

The difference in the net throughput of these coatings clearly shows the challenge and trades to be considered for the silver and aluminum approaches already demonstrated at 8 m sizes. It also shows the significant advantage to the LLNL coating if it can be scaled for similar results at an 8 m size. The development of LSST coatings will address the issues of coating longevity and process scaling to 8 m sizes. LSST currently has a long term evaluation program in place with 14 sample trays with three protected coatings on four different substrate materials distributed at 6 observatories to investigate longevity and durability. The scaling issue will be pursued during the development phase of the telescope effort.

Achieving the signal to noise and calibrations necessary to meet the LSST requirements will require an effective mirror cleaning approach and strict operation plans to routinely maintain the surface quality. The LSST will include in-situ hardware for the cleaning process. The exact nature of the process is the subject of the development program and will include investigations into dry processes such as CO_2 and UV excimer laser methods with obvious on-telescope benefits and wet processes that are traditionally most effective but pose many on-telescope issues. A wet process capability will be included in the design of the system if it is expected to be necessary only for sparse but routine cleanings

4.2.5 Alignment and Active Optics System

4.2.5.1 Introduction and Approach

The purpose of the Active Optics System (AOS) is to establish and maintain the high optical image quality in the LSST to its specification. The AOS measures position errors of the optical elements and the optical wavefront error at the focal surface and converts these errors into corrective action on compensated degrees of freedom. The approach we have adopted is an extension to LSST's large wide field of view and three mirror optical system of the methodologies used in existing modern ground based telescopes.

Most modern large ground based telescope telescopes use some form of active optics to maintain their optical performance by compensating for surface and alignment errors. These errors are due to fabrication, flexure of the optical support structure (OSS) as its orientation is changed and thermal gradients within the OSS. Active optics systems operate on relatively long time scales and do not attempt to correct for distortions caused by the Earth's atmosphere. The basic method converts wavefront measurement errors to corrective action by using a linear optical model given by

$$\begin{bmatrix} w_1 \\ \vdots \\ w_i \end{bmatrix} = w(0) + S \bullet \begin{bmatrix} g_1 \\ \vdots \\ g_k \end{bmatrix},$$

where the vector w represents the on axis wavefront error in the exit pupil for a given set of perturbations g in the optical system's degrees of freedom. The aberrations inherent in the unperturbed optical design are given by ***w(0)***. The matrix ***S*** is referred to as the sensitivity matrix with elements $S_{ik} = \partial w_i / \partial g_k$.

The solution for a given set of corrections is found by applying the inverse of the sensitivity matrix, $\boldsymbol{S}^{-1}$, to a set of time averaged measurements of the quasi-static wavefront error. Typically w is represented by the Zernike set of polynomials because the first terms of this series closely match natural optical aberrations associated with misalignments and low order bending modes of the telescope primary mirror. In ground based telescope applications, a dedicated wavefront sensor located at the periphery of the science field of view is used to measure w from a reference point source (*e.g.* a star). These measurements must be time averaged in order to ensure that aberrations from atmospheric turbulence do not dominate and that w represents a good estimation of the telescope aberrations.

The LSST presents several challenges for the alignment problem: 1) The LSST is composed of three large mirrors. This increases the number of degrees of freedom in the optical system over the more typical two-mirror telescope. 2) The LSST design is optically very fast with a system f-number of 1.234, leaving little room for error to maintain the required image quality. The sensitivity to alignment errors is further enhanced by the high image quality performance demands for the LSST. 3) The baseline sequencing of LSST exposures is on a faster timescale that would be generally used to time average the effects of atmospheric turbulence on the optical wavefront errors caused by surface deformation and misalignments. 4) The wide field of view in combination with the fast beam and rapid repointing of the telescope restricts the use of probes to place wavefront sensors where suitable reference stars exist.

The remaining sections presented here represent a summary of a proof of principle analysis that considered the degrees of freedom (Section 4.2.5.2), resulting field dependent optical wavefront errors (Section 4.2.5.3), and the wavefront sensing strategy during initial alignment from assembly tolerances and during routine operation (Section 4.2.5.4). This analysis was carried out on an earlier design of the LSST, and with a separate M1 and M3 mirror construction, but is equally valid for the present baseline design and represents our baseline approach to aligning and controlling the LSST optical system.

4.2.5.2 Degrees of Freedom and Wavefront Field Dependencies

For the purpose of analyzing the LSST's alignment and surface control requirements we have considered 125 degrees of freedom (DoF). These include rigid body displacements of the three mirrors and camera as well as 36 Zernike modes to model each of the three mirror surfaces. It should be noted that in the actual telescope system the mirror surfaces will be controlled using

influence functions of the individual actuators, but for the purposes of this feasibility analysis the Zernike representation is deemed adequate. By contrast, a typical two-mirror telescope may have 20 or so controlled degrees of freedom in its AOS.

The vertex of the primary mirror is the reference for aligning the LSST optical system; therefore we do not consider M1 translation in our DoF. Because all the optical elements in the LSST are rotationally symmetric we do not count element rotation in our DoF. The camera refractive elements are relatively insensitive to misalignments and bending; therefore we have assumed the camera optics and focal plane to be a single rigid element. The coordinate system used for the rigid body perturbations is shown in Figure 4.2.5-1. In this system a positive tilt about the x-axis of M3, for example, results in a positive decenter of the optical axis at the focal plane.

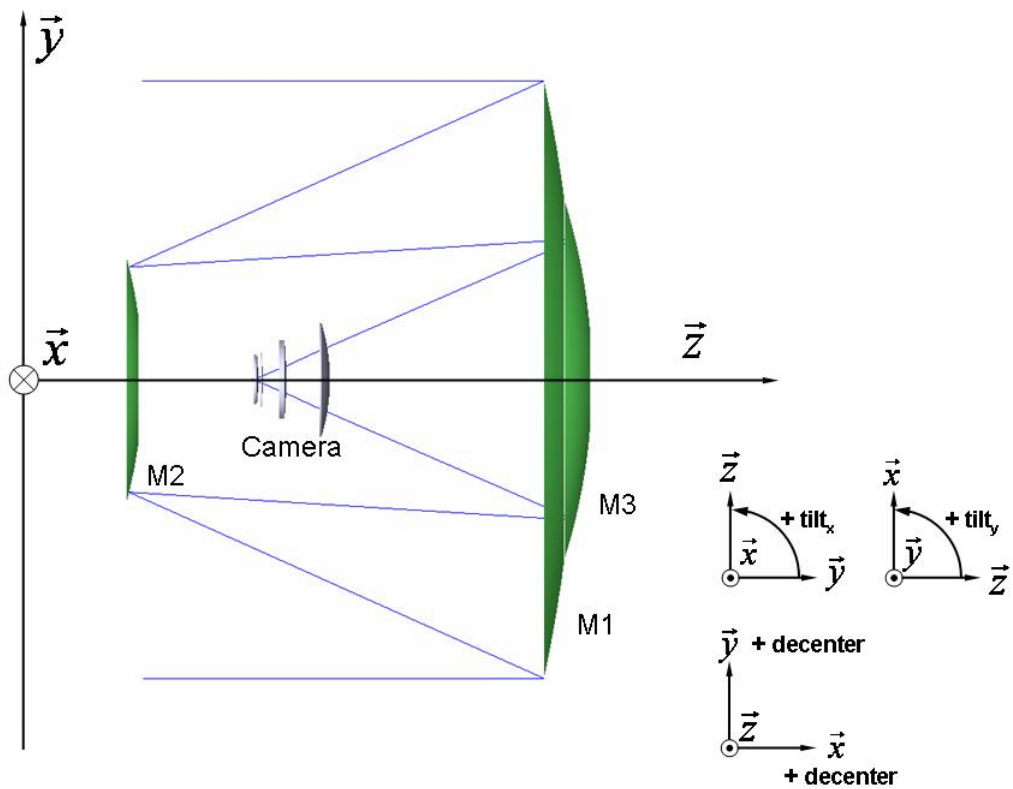


Figure 4.2.5-1 Coordinate system and sign convention used for defining the degrees of freedom for rigid body displacements in the LSST alignment analysis.

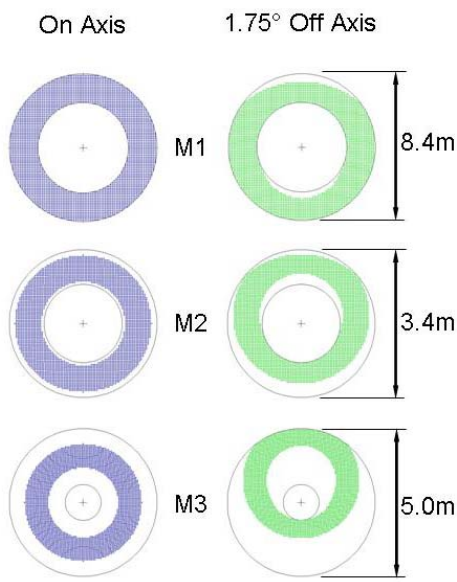


Figure 4.2.5-2 Optical footprint on each of the three mirrors for field points on axis (blue) and at full field of view at 1.75 degrees off axis (green).

The entrance pupil of the optical system is located at the front surface of M1, the primary mirror. As the distance of an optical surface to the entrance pupil increases, the different field positions have an illumination footprint that moves on that surface. The changes in the illumination footprint location for each of the three mirrors at two field positions, on-axis and at the field edge of the full FOV, are shown in Figure 4.2.5-2. These variations cause a given error, whether from misalignment or surface error, to have a unique signature in wavefront error as a function of field position in the focal plane. Two examples of the field dependencies generated are shown in Figure 4.2.5-3, which show field dependent astigmatism resulting from a low order surface mode on M3. In one instance a 3-theta surface mode (trefoil) on the surface of M3 results primarily in astigmatism, a 2-theta aberration, at the focal surface having 3-theta field dependency.

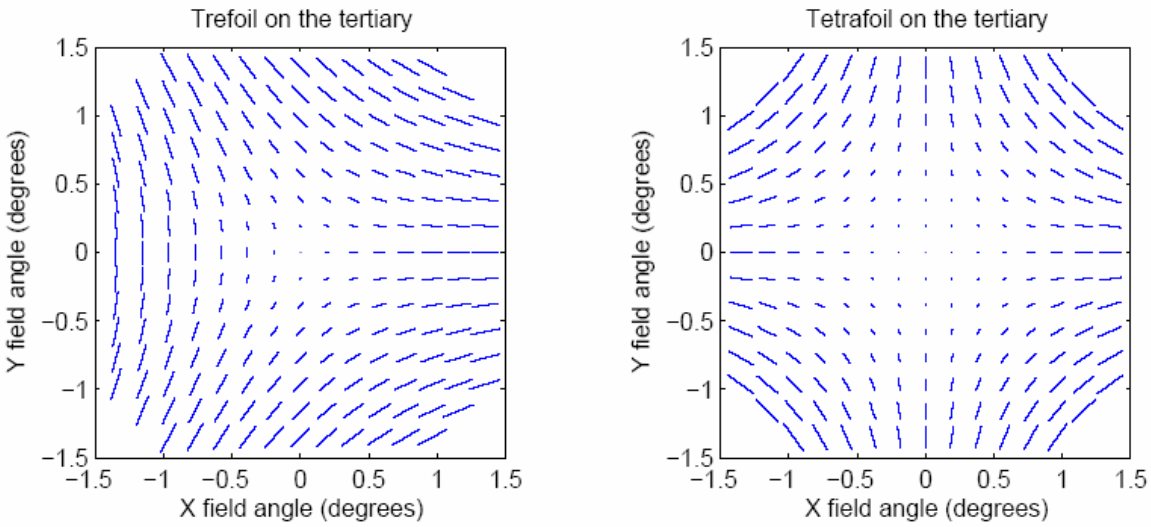


Figure 4.2.5-3 Two examples showing the field dependencies of astigmatism caused by a higher order aberration, trefoil (left) and tetrafoil (right), on the tertiary. The blue sticks represent the orientation and strength of astigmatism in the focal plane as a function of field position.

The key issue for the LSST is to isolate the sources of the field dependencies for the most general case, when many contributions are present and in the presence of noise. In the example in Figure 4.2.5-4 a mixed set of astigmatism field dependencies in the focal plane (left) are decomposed into their component sources (right). In this example, the astigmatism in the focal plane is produced by a mixture of astigmatic bending modes (Z5, Z6) on M1, tilts in M2, and 3-theta bending modes (Z9, Z10) on M3. The next section focuses on one possible method for determining the error sources and amplitudes from an arbitrarily perturbed optical system in the degrees of freedom previously discussed.

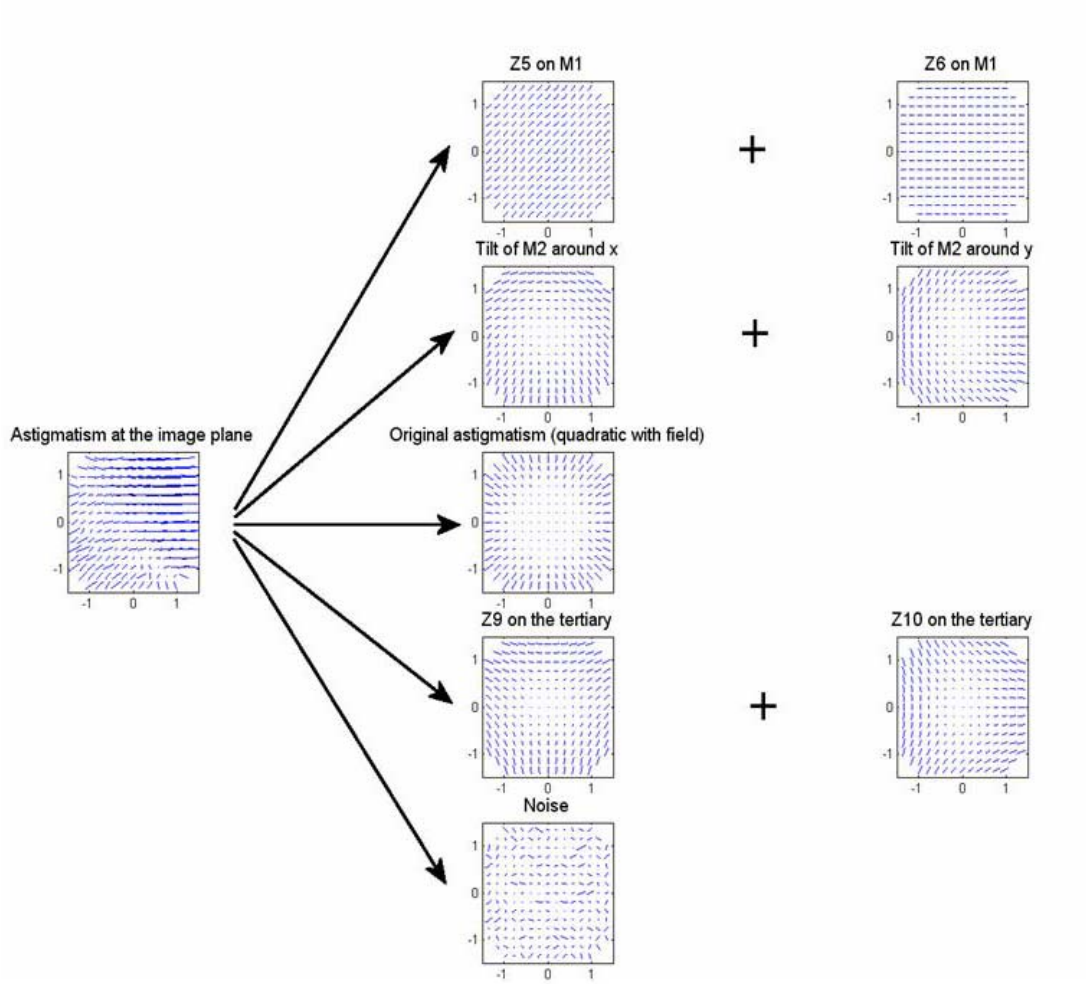


Figure 4.2.5-4 An example decomposition of astigmatism in the focal plane to component sources using the Zernike representation.

4.2.5.3 Error Reconstruction

One of the central elements in the LSST AOS is the "reconstructor". This is the mathematical tool that takes the wavefront error field dependencies and converts these to misalignments and surface errors that can be used to determine the necessary corrective action. The general method of error reconstruction is shown in Figure 4.2.5-5. By characterizing the field dependencies in the wavefront errors for misalignments and bending modes separately we have been able to show that the process of converting field dependent wavefront errors as measured in a limited number of filed positions does converge to acceptable tolerances in image quality.

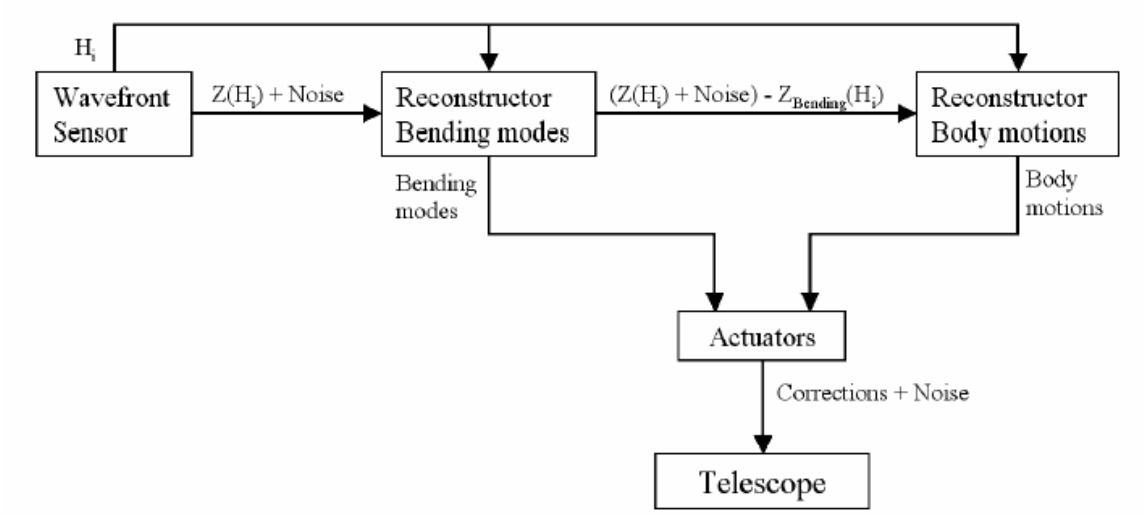


Figure 4.2.5-5 Functional diagram of the LSST Active Optics System

First the wavefront sensors measure the Zernike coefficients Z at different field points (H_i). Then with the Zernike coefficients $Z(H_i)$ a least-squares fit is performed with the reconstructor of the bending modes, which gives the corrections to the bending mirrors and the values of the Zernike coefficients coming from them. The least-squares fit is performed first without considering the body motions because the bending modes produce low order aberrations which are similar to the aberrations created with the body motions. But they create also higher order aberrations, which allow differentiating them from the body motions.

Once the bending modes are estimated, it is possible to subtract these Zernike coefficients from the set originally computed from the wavefront sensors images, to calculate the Zernike coefficients produced by the body motions modulo the noise and the centered aberrations from the perfectly aligned system. After the expected bending modes and body motions have been calculated, the corrections are applied to the telescope optics and camera to align them.

The alignment strategy on the real system can be divided into two main phases: the initial and the operational alignment. The initial alignment is the task performed when all the elements are assembled and aligned together for the first time either during construction or after maintenance. The operational alignment is the task performed during a night while the telescope is in use to correct small misalignments. Both phases are detailed hereafter as they differ in their application.

4.2.5.3.1 Initial Alignment

During the initial alignment, all the elements are assembled within relatively rough tolerance ranges. A typical value for the tolerance range for all the rigid body motions is ±1mm for decenters or spacings. The tolerance ranges for all the degrees of freedom chosen for the initial alignment are summarized in the Table 4.2.5-1.

For that phase, a large number of field points are needed to evaluate the corrections over the wide range of tolerances we anticipate after initial assembly and after routine maintenance (*e.g.* reassembly after mirror recoating or major camera servicing). We have conducted a Monte Carlo simulation using the above process to test the convergence of the optical system. To exercise the method for the initial alignment with Monte Carlo simulations, 48 field points were chosen to be able to sense all the field dependent aberrations. The assumption is that the camera focal plane is

used as the wavefront sensor. After the first iteration, the averaged merit function was found equal to 4.37 μm with a standard deviation of 0.89 μm for 50% of all the models (see Figure 4.2.5-7). Several iterations of this process were repeated in order to get more than 90% of all the Monte Carlo models within the alignment budget. Therefore, if the elements can be manufactured and positioned within the tolerances shown in Table 4.2.5-1, the initial alignment can be realized.

Table 4.2.5-1 Assembly tolerances used in simulation of initial alignment.

Body Motion	**Decenter**	**Tilts**	**Focus**
M1	±1 mm	±0.0136 deg	±1 mm
M2	±1 mm	±0.0337 deg	±1 mm
M3	±1 mm	±0.0229 deg	±0.5 mm
Camera	±1 mm	±0.0654 deg	--
Bending Modes	**Low order (Z5-Z11)**	**High order (Z12-Z37)**	
M1	±2 λ	±0.5 λ	
M2	±2 λ	±0.5 λ	
M3	±2 λ	±0.5 λ	
Spacings	**M1-M2**	**M2-M3**	**M3-Camera**
	±1 mm	±1 mm	±1 mm

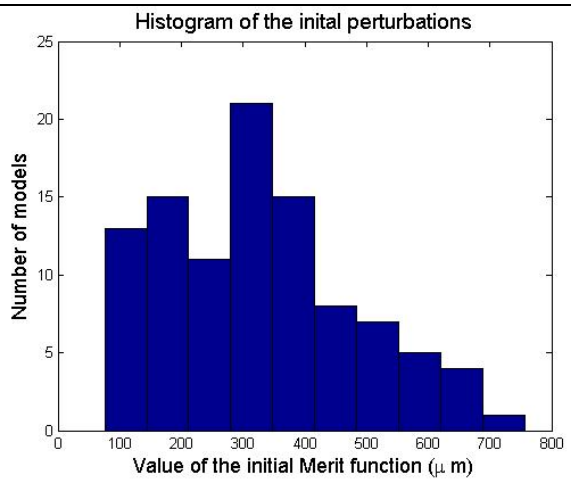


Figure 4.2.5-6: The initial merit function histogram of RMS image size across the field of view from the Monte Carlo simulations of the perturbed optical system. The optical system was perturbed in all degrees of freedom within the tolerances in Table 4.2.5-1.

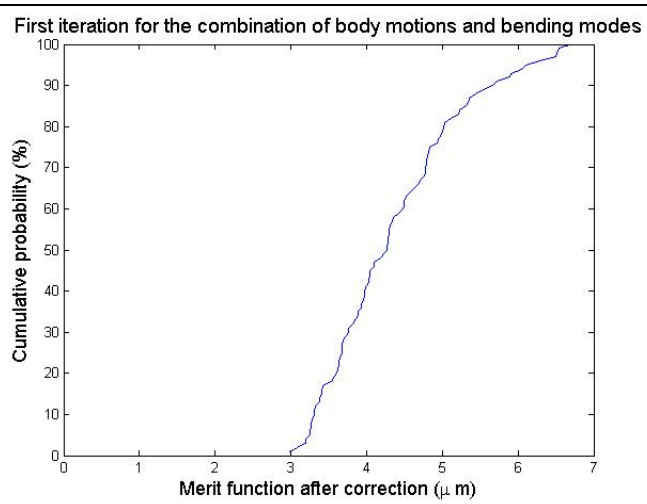


Figure 4.2.5-7: The cumulative distribution (in %) of all the models versus the averaged value of the merit function after a first iteration of correction through the reconstructor. In this case, 50% of all the Monte Carlo models reach an averaged merit function of 4.37 microns.

4.2.5.3.2 Operational Alignment

The operational alignment mode maintains the image quality during the normal science observation time of the telescope. The tolerances are smaller than for the initial alignment because, in this configuration, the alignment has already been done and the purpose is only to maintain it. The tolerance ranges for all the degrees of freedom chosen for the operational alignment are summarized in the Table 4.2.5-2.

Table 4.2.5-2: Estimated capture range requirement for maintaining operational alignment.

Body Motion	**Decenter**	**Tilts**	**Focus**
M1	±100 µm	±5”	±100 µm
M2	±100 µm	±12”	±100 µm
M3	±100 µm	±8”	±50 µm
Camera	±100 µm	±23”	--
Bending Modes	**Low order (Z5-Z11)**	**High order (Z12-Z37)**	
M1	±2 λ	±0.5 λ	
M2	±2 λ	±0.5 λ	
M3	±2 λ	±0.5 λ	
Spacings	**M1-M2**	**M2-M3**	**M3-Camera**
	±100 µm	±100 µm	±100 µm

Since, during operation, the camera is dedicated to the science and cannot be used for wavefront sensing, another WFS system is required for sensing the field dependencies of the wavefront error. To achieve this requirement, a small portion of the available field of view distributed over a relatively limited number of field points is allocated to the WFS system to calculate the corrections. How many field points are necessary to calculate the corrections to maintain a good image quality? With too few field points, it will not be possible to sense all the field dependencies of the aberration and maintain a good correction and with too many field points the loss science area becomes unacceptable.

We have conducted several Monte Carlo simulations studying the number and position of field points needed and considering different levels of noise in wavefront measurement. The results are summarized in Figure 4.2.5-8. The convergence is very dependent on the number of points and on the amount of noise. For that study, a flat random noise was distributed over all the Zernike coefficients with an amplitude variation of ±0.01, ±0.05 and ±0.25 wave at a wavelength of 644 nm. For a noise level below 0.1wave, eight to ten points in the field of view appear to be enough to maintain the alignment. Such a noise distribution was deemed acceptable for this study although in the real world, it may follow a different pattern and may affect preferentially the lower Zernike coefficients. Our current estimation is that our noise level would be between ±0.05 and ±0.25 wave so that 10 to 20 field points may be required (we adopt a quantity of 15 WFS in the rest of this chapter to simplify the text).

The plan is to continue this modeling during the R&D phase. One specific source of noise comes from the LSST operational model, where exposures for certain filters are based on a 15 sec length. This duration is too short to average completely the effects of the Earth’s atmosphere on the wavefront measurements. When the shutter is closed during read-out of the CCDs, the WFS system is also required to stop its measurements. Our plan is to investigate three possible methods to remedy this condition. First, we could average WFS measurements over multiple acquisitions of a same field. In the current operational model, at least a pair of images is taken per field. Second, we could allow secular trending for adjacent fields to average the atmosphere because

the telescope perturbations between two adjacent fields should be minimal. Third, we could exploit for each exposure the correlation of the field dependencies over the 15 WFS because the isoplanatic angle of the atmosphere is smaller than the angular separation between the wavefront sensors.

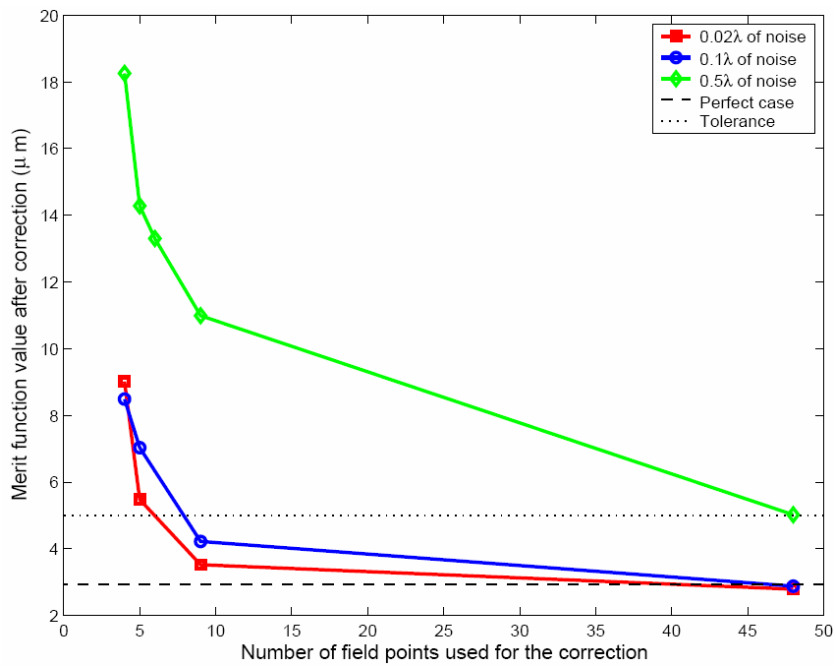


Figure 4.2.5-8: The limiting merit function (mean RMS image size over full FOV) as a function of wavefront sample points. The noise was applied to each Zernike term at each field point using a uniform distribution having the width indicated.

These 15 active WFS field points have to produce wavefront error measurements simultaneously for a given pointing on the sky. Because the star distribution in the field of view will not be identical for each observed field, a number of extra wavefront sensors will be needed to ensure that 15 WFS are always available per exposure. Star availability is discussed briefly in the next paragraph, and we plan to study this aspect in more detail during the R&D phase in correlation with the distribution of these sensors in the field of view.

4.2.5.4 Wavefront Sensing

In this section, we describe what type of wavefront sensing techniques we are investigating.

4.2.5.4.1 Initial Alignment

To map the wavefront field dependencies for the initial alignment reconstruction we intend to use the curvature wavefront sensing method from the science array itself. Using the science array allows a large number of field points to be measured without needing auxiliary equipment or sacrificing a large amount of the field of view to dedicated wave front sensors.

Curvature wavefront sensing relies on solving the intensity transport equation that takes the form:

$$\frac{I_+ - I_-}{I_+ + I_-} = \frac{f(f-L)}{L}\left(\delta_r \frac{\partial \varphi}{\partial r} - \nabla^2 \varphi\right)$$

[where *I*- and *I*+ are intra and extra focal intensity images symmetrically placed a distance *L* about best focus, $\partial\varphi/\partial r$ is the slope normal to the pupil boundary, δr is a delta function centered at the boundary and φ is the wavefront phase in length units. For the initial alignment *I*- and *I*+ would be obtained with the science array by adjusting the camera focus by ±L around best focus. The effects of the atmosphere would be averaged out by taking multiple short exposures at each focus position. A separate image in focus would be used to determine the locations of suitably isolated reference stars for use in the intra and extra focal images. These images would be fed to a separate software pipeline to extract the Zernike coefficients needed by the reconstructor described above.

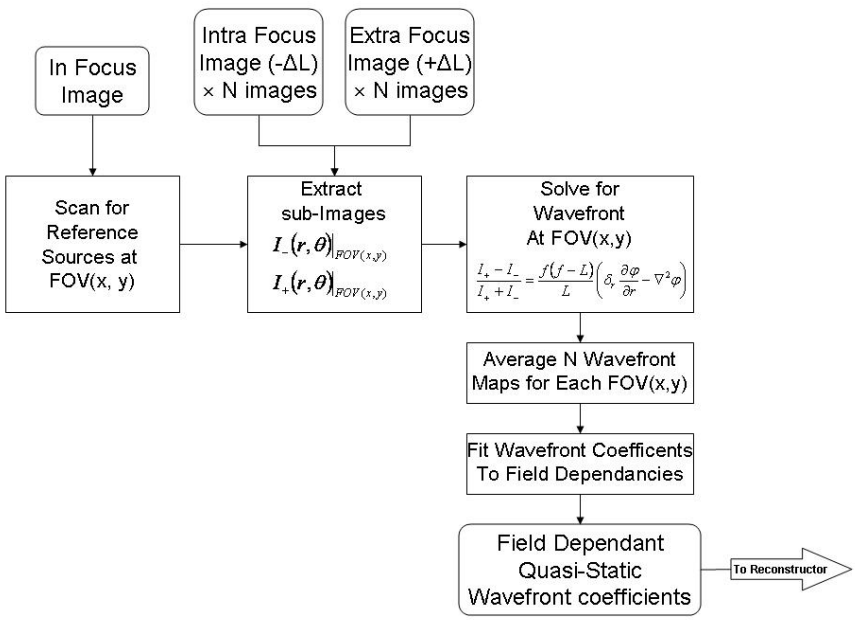


Figure 4.2.5-9 Z (Hi) process flow for initial alignment wavefront curvature data.

The flow of the image analysis pipeline for the initial alignment is shown in Figure 4.2.5-9. First an in focus image of the field is obtained, which is used to identify clean reference stars. Curvature wavefront sensing is particularly sensitive to overlapping objects when the system is defocused by ±L. The reference stars' locations are recorded to be used in extracting sub-images at multiple field locations (x,y) from the intra and extra focal images. A series of intra and extra focal images follow the acquisition of the in-focus image with individual exposure times of sufficient length to average out the effects of atmospheric turbulence on the wavefront estimation. The exposure time used for averaging out the atmosphere is typically 30-60 seconds. For each pair of intra and extra focal sub-images the intensity transport equation is solved providing a wavefront map of the telescope pupil at each field location. The field dependent aberrations of the telescope are represented by an average of the wavefront estimation for each field position. By fitting the field dependencies to fixed coefficients that are passed on to the reconstructor algorithm we prevent the reconstructor from having to be recomputed for each unique sampling of the field by the reference stars.

4.2.5.4.2 Operational Alignment

The wavefront sensing for operational alignment is more of a challenge. This is because the science array is naturally unavailable for providing wavefront information, yet we are still required to sense the wavefront field dependencies to maintain the LSST optical performance.

Phase retrieval was the first technique that was simulated for wave-front sensing on the Large Synoptic Survey Telescope. This technique would not require any additional hardware and could be implemented with just the images on the focal plane array. It simply uses the far-field spot along with an assumption of the intensity in the pupil plane. For the simulations, the intensity in the pupil plane was assumed to be uniform and an estimate of a uniform phase at the pupil plane was the starting point. This approach, as shown in Figure 4.2.5-10, worked well when no atmosphere was present (simulations were well sampled), however, the simulations were not able to handle modest amounts of atmosphere.

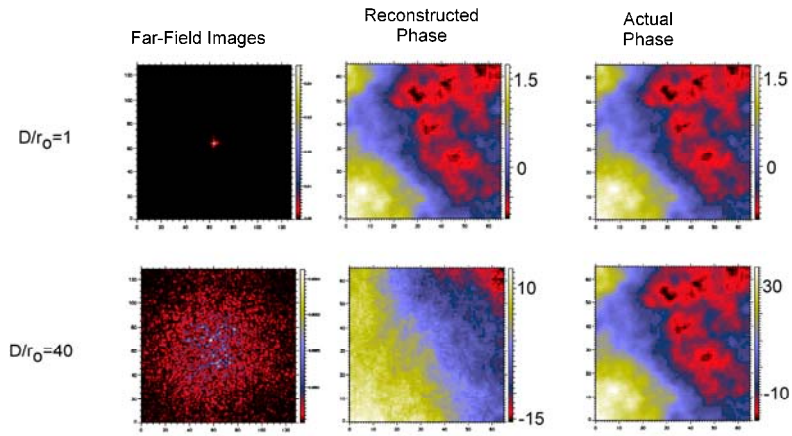


Figure 4.2.5-10 Phase retrieval simulations with weak and strong turbulence.

Then, a phase diversity method was simulated for wave-front sensing on the Large Synoptic Survey Telescope. This technique requires the measurement of the intensity in two planes, usually at focus and a few waves out of focus. The simulations were performed using a least squares fitting of Zernike modes approach. Phase diversity was shown to handle undersampling from Nyquist by a factor of eight at D/ro values in the range of 20-40. This technique was abandoned because the LSST is undersampled from Nyquist by as much as a factor of 80 for the shortest wavelength and the worst turbulence conditions, which precludes the use of this technique without additional optics to magnify the images. Other potential methods are available and we describe below the ones that we are currently considering.

We are looking into a Shack-Hartmann WFS (SHWFS) that could be deployed at the selected field points in the focal plane. This type of WFS is based on the measurement of the first derivative of the wavefront, and is well understood with a robust wavefront estimator. The Shack-Hartmann wave-front sensor is usually composed of a lenslet array placed in the pupil of the optical beam being measured. The lenslet array is composed of an array of identical lenslets, each of which measures a small portion of the optical beam. A CCD camera is placed one focal length behind the lenslet array. A plane wave incident on the lenslet array produces an array of focal spots located on the optical axis of the individual lenslets. An aberrated beam produces a local gradient in the phase, $\nabla\phi$, across each of the lenslets, displacing the focal spots from the optical axis as shown in Figure 4.2.5-11. Each individual spot is displaced from the center of the lenslet by a distance $s = (\nabla\varphi)(f.l.)\lambda/2\pi$, where λ is the wavelength of the incident light, as illustrated

in Figure 4.2.5-12. Therefore, the Shack-Hartmann wave-front sensor measures the wave-front gradient with a spatial resolution equal to the sub-aperture size.

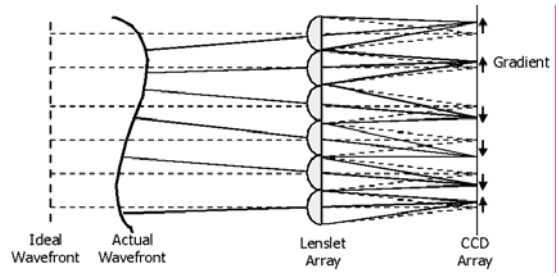


Figure 4.2.5-11 Illustration denoting hypothetical Hartmann spots from a plane wave, dashed lines, and a severely aberrated beam, solid line.

If the source of illumination for the Hartmann sensor is a point source, as in a traditional Hartmann sensor, then the spot displacements are typically found using a “center-of-mass” calculation on the intensity. If there is an extended scene illuminating the pupil, then the scene displacement at each of the sub-apertures is measured by performing cross-correlations between the scenes in the separate sub-apertures. The latter technique was pioneered by the solar adaptive optics community and is the more likely technique to be used for the LSST telescope if the pupil of the telescope is reimaged beyond the focal plane array.

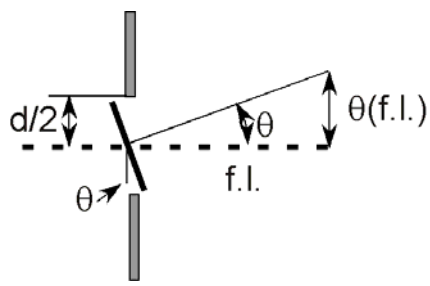


Figure 4.2.5-12 Principle of operation of each lenslet. The angular displacement of the spot for each individual lenslet indicates the average slope vector of the wave-front across the given lenslet.

A SHWFS preliminary design has been established with the condition of minimizing the impact on the LSST science focal plane array by avoiding any moving parts, with near diffraction limited performance over LSST spectral bands, and occupying as small an area as possible. It is composed of a relay system with nine optical elements that forms a 4 mm diameter image of the pupil onto a lenslet array (see Figure 4.2.5-13). This is equivalent to a 5 mm focal length aspheric lens.

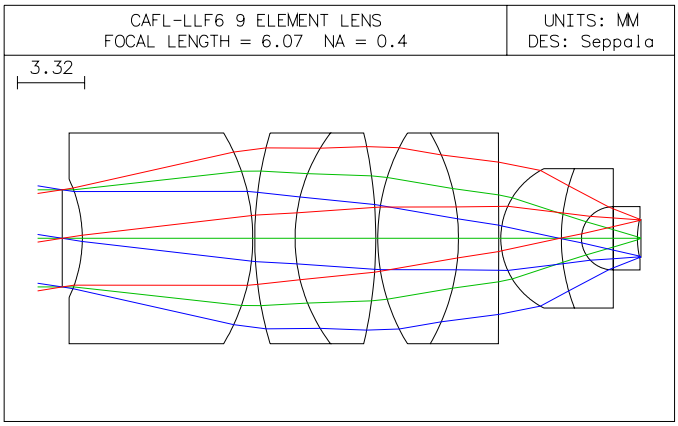


Figure 4.2.5-13 Preliminary design of a Shack-Hartmann wavefront sensor. It is composed of nine optical elements with a numerical aperture (NA) of 0.4. It reimages the pupil of the telescope onto a lenslet array.

An array of 9 identical SHWFS would be mounted next to each other on a 3x3 grid and would occupy an area of 21x21mm in the focal plane. This configuration allows an increase in the acquisition range of sky from 37.5" per SHWFS to a total of 112" on the sky (Figure 4.2.5-14). The drawback of this method is the inefficient use of space in the focal plane. A 21mm^2 area corresponds to a ~45 arcmin^2 on the sky which is ~16 times more than the total ~2.8 arcmin^2 available. However, having this array of Shack-Hartmann sensors presents the advantage that if multiple stars can be found then multiple measurements of the mirror aberrations can be averaged to reduce the noise and help take out residual atmospheric aberrations.

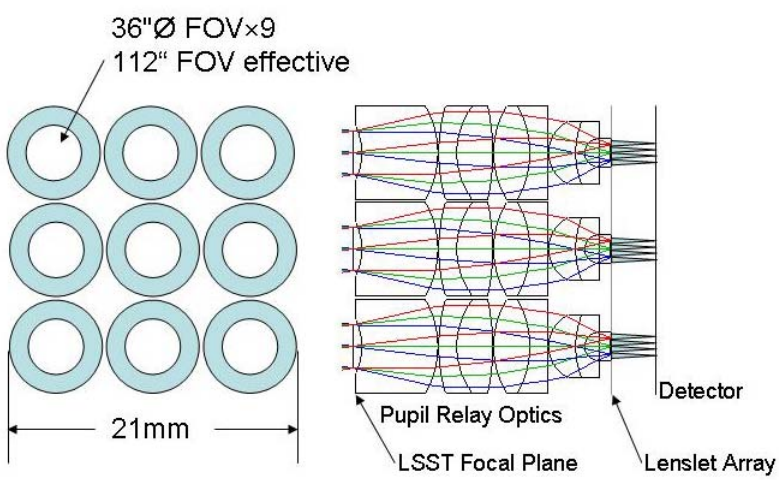


Figure 4.2.5-14 Wavefront sensor configuration with 3x3 Shack-Hartmann sensor. The total field available is 112", having each sensor with a field of view of 37.5".

The second technique we are investigating is the use of a curvature wavefront sensor (CWFS). In its traditional mode of operation, the SHWFS is expected to deliver a grid of point sources, with a spot size essentially distributed over 4 pixels (as a quadcell). The number of photons per pixel on the wavefront sensor versus the magnitude of the star is evaluated in Figure 4.2.5-15 below by taking into account some realistic assumptions on the overall transmission of the system. If we set our minimum number of photons per pixel to 50 to obtain a reasonable signal to noise ratio, it would require at least a star magnitude of ~19.5 to run the wavefront sensor.

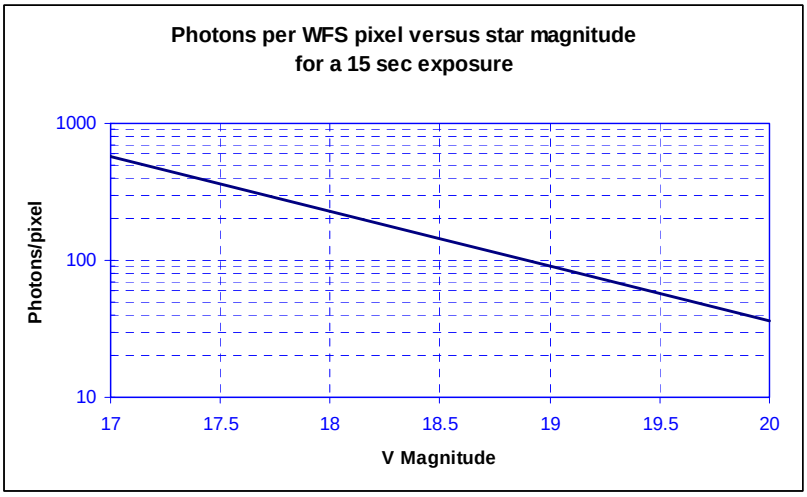


Figure 4.2.5-15 Number of photons per pixel of the Shack-Hartmann wavefront sensor (assumption of a quadcell per spot) versus star magnitude for a 15sec exposure

The surface density S of such magnitude stars near the galactic pole is S ~ 0.3 per sq-arcmin. Assuming the stars are randomly distributed on the sky, then the probability of finding a useable star within the WFS field of view solid angle A is equal to $P = 1 - e^{-n}$ where n is equal to S*A. With the 3x3 grid configuration of WFS, this leads to a probability of ~ 57% of finding a suitable star near the galactic pole. This probability increases as the field positions move toward the galactic equator. For instance, for a galactic latitude of 30 deg, the probability is above 95%.

Simulations of a Shack-Hartmann wave-front sensor have started with a configuration of 23 lenslets across the LSST pupil as shown in Figure 4.2.5-16. In this case the pupil size on the lenslet array is 4 mm such that each lenslet represents 0.37 m on the primary mirror and there are 4 x 4 CCD pixels per lenslet in the simulations. These simulations will be continued during the R&D phase.

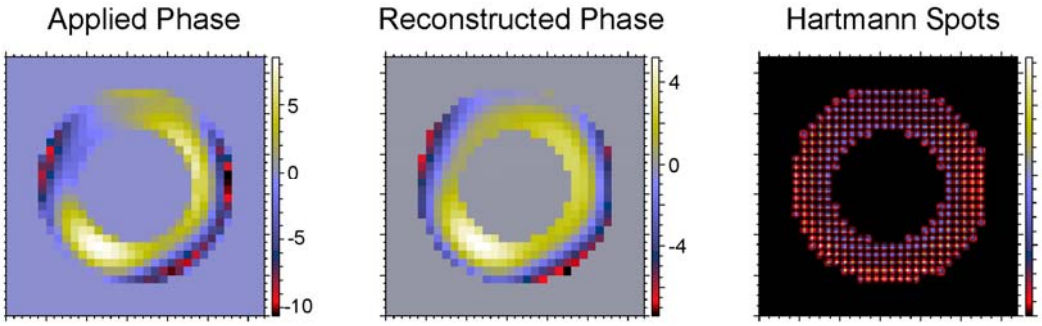


Figure 4.2.5-16 Shack-Hartmann simulation images of the applied and reconstructed phase. The image on the left shows the spot distribution on the Shack-Hartmann sensor.

The second technique we are investigating is the use of a curvature wavefront sensor (CWFS). The principle of the curvature sensing method was explained above for the initial alignment. In a traditional implementation of the CWFS, a pair of detectors is used to record simultaneously intra-focal and extra-focal images of the same star. This requires the use of a beamsplitter or equivalent to feed both detectors.

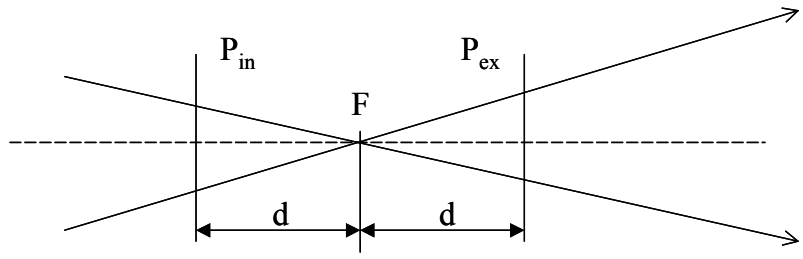


Figure 4.2.5-17 Curvature sensing consist of taking the difference between the illuminations observed in 2 planes Pin (intra-focal) and Pex (extra-focal) separated from the focal plane by the same distance d.

Instead, we are currently proposing to use different stars in order to remove any additional optical element. In that configuration, both detectors are fixed next to each other, both defocused by the same amount, with one in an intra-focal position, and the other in an extra-focal position. Each detector produces images of a different field. The two detectors are close enough so that the variations of the telescope aberrations over that distance are negligible. The process is then to extract from the CWFS images small areas centered around suitable stars and use that information to solve the intensity transport equation to estimate the wavefront error. As it stands, it would be the simplest method to implement opto-mechanically in the focal plane and would provide the most efficient use of the area allocated for wavefront sensing. A concept is presented in Figure 4.2.5-18.

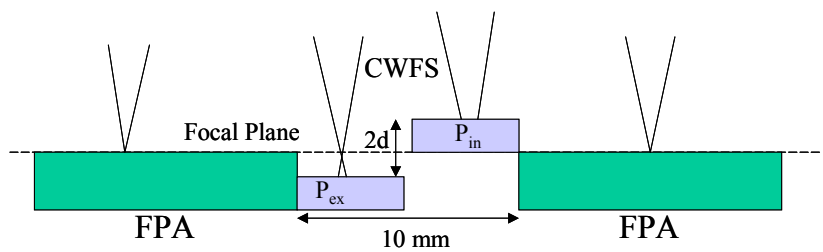


Figure 4.2.5-18 Concept of a CWFS where two detectors, located next to each other and in between the science detectors of the focal plane array (FPA), are placed in an extra-focal position (Pex) and in an intra-focal position (Pin) with a total separation of 2d.

We have developed simulations using curvature sensing algorithms based on Zernike decomposition of the intensity transport equation and an iterative technique based on the simplex method. In the first technique, the intensity transport equation (ITE) can be converted to a system of algebraic equations by assuming an annular Zernike decomposition of the phase. The annular Zernike coefficients of the phase are determined by using singular value decomposition to solve the system of algebraic equations.

In the Simplex iterative method, the wavefront is represented with an annular Zernike basis set. A minimization of the error metric, the mean square difference of intensities between measured and simulated defocused images, is performed. “New” annular Zernike coefficients are found using the simplex algorithm . These coefficients are then propagated to form intensity images after each iteration.

An example of these simulations is presented in Figure 4.2.5-19, showing an intra-focus image and an extra-focus image and the reconstructed phase. These images are 128x128 pixels in size. A defocus of ±500 microns was used for this simulation. The exposure time was set to 30 sec and it includes also atmospheric turbulence.

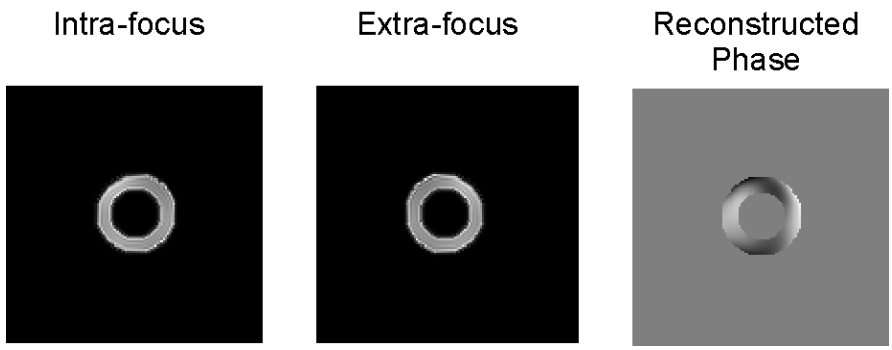


Figure 4.2.5-19 Example of intra and extra-focal images of a single star. The defocus distance is 500 microns. The reconstructed phase is displayed on the right.

The simplex method works well for a phase profile containing a small number of Zernikes. When a power law Zernike set was tried, the iterative technique based on the Simplex method did not perform adequately and so new iterative techniques will be investigated during the R&D phase. The hypothesis is that these techniques are becoming trapped in local minima and unable to find the global minimum. Therefore, we will explore iterative techniques based on simulated annealing and genetic algorithms, which have mechanisms to escape local minima and find the global minimum.

The effect of the choice of a defocus distance on the quality of the reconstructed phase was studied. The iterative technique was shown to give significantly better performance than the approach using Zernike fitting of the intensity transport equation for distances close to the focus, < 2 mm. The performance is measured by computing the variance of the reconstructed phase. The variations of the variance versus the defocus distance are in Figure 4.2.5-20. The results indicate that this technique should be used for distances between 300 to 500 microns from focus. However, they were achieved assuming a fixed number of 64 pixels across the aperture, and no atmosphere. The actual case of a fixed focal plane pixel size with atmosphere would cause there to be 40 pixels across the aperture at 500 microns and would likely make the results at distances close to the focus increase more rapidly than shown in the figure below.

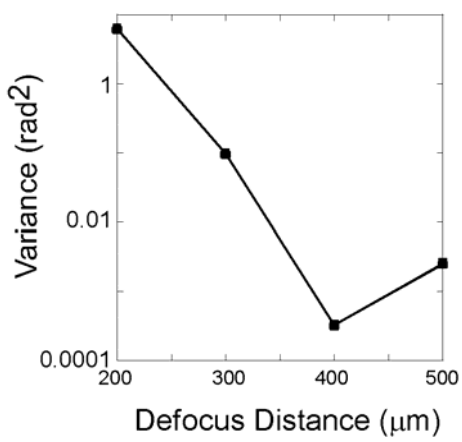


Figure 4.2.5-20 Effect of defocus distance on the variance of the reconstructed phase for the iterative technique

The star magnitude limit is estimated to be approximately the same as presented above for the SHWFS concept. The CWFS has no transmission loss due to additional optics. However it uses defocused images, and is also prone to be more sensitive to noise because of using the second

derivative of the wavefront (the Laplacian) to retrieve the wavefront error. Consequently, our current estimation is that it would require at least a star magnitude of ~19 to run the wavefront sensor.

The star availability for the CWFS is improved because the whole surface allocated can be used for finding a suitable star. If we assume using 1K x 0.5K detectors with 10micron pixels (equivalent to a ~1x1 cm^2 for both detectors), the FOV available per detector is around 3.3 x 1.7 = 5.6 sq-arcmin. This leads to a probability better than 80% of finding a suitable star near the galactic pole. However, a defocused star image uses more area on the detector than a focused star. The issue is then not to have too few stars but too many causing pupil overlap for fields too crowded with stars. This is also one of the reasons why we have tried to restrict the defocus distances to ~ 500 microns to limit the size of the donuts. We have started a study of pupil overlap and its effect on the phase reconstruction. Multiple parameters are involved like the overlap distance or the relative intensity difference between overlapping donuts. An example is provided in Figure 4.2.5-21 where two stars separated by 30 pixels have been simulated. One star has 40% of the intensity of the other. The intra and extra-focus images are shown on the left. The applied and reconstructed phases are shown on the right.

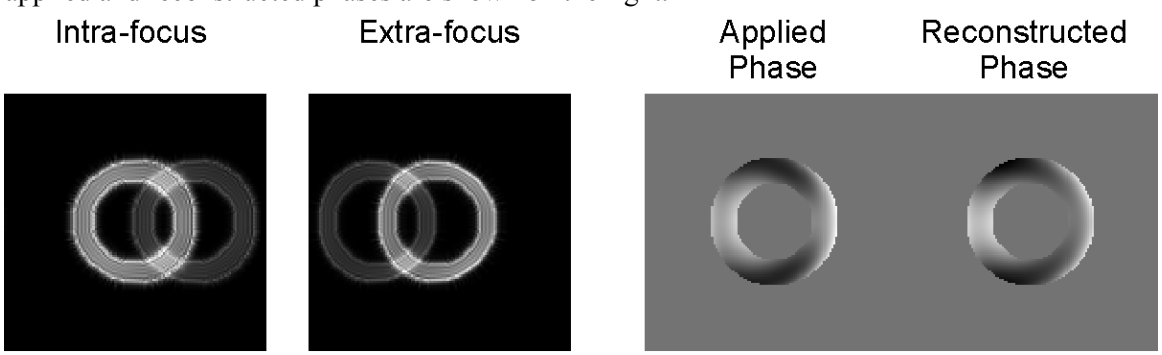


Figure 4.2.5-21 Pupil overlap simulations. The intra and extra-focal images on the left show two stars separated by 30 pixels, with an intensity ratio of 40%, and with a defocus distance of 500 microns. On the right side are displayed the input phase and the reconstructed phase for comparison.

The iterative technique proved to be able to handle pupil overlap, with an increase in residual mean square error from 0.15 rad to 0.8 rad as the overlapped image was changed in intensity from 10% to 40% of the original image. The effect of a reduced number of pixels across the aperture was also studied. As the number was taken to 40 pixels across the aperture, images taken 500 microns from best focus, the reconstruction suffered at smaller values of r_o. More simulations will be conducted during the R&D phase to optimize the parameters of the system.

Finally, we are looking into another way of measuring only low-order aberrations from the defocused images. The set-up would be similar to the CWFS. However, instead of solving the intensity transport equation, the aberrations would be extracted from second moments measured directly on the defocused star images. Andrei Tokovinin from CTIO is currently leading this effort and is developing a simulation package for testing purposes. The size of the donut is set mostly by the amount of defocus, and its elongation is caused mainly by astigmatism. Both parameters can be extracted from the second moments of the intensity distribution. Such a method may allow us to reduce pupil overlap by being able to use a shorter defocus distance. We expect to expand these simulations during the R&D phase.

In summary, although LSST presents several challenges in terms of alignment, our studies have proven that the active optics control of the large three mirrors and camera is in principle feasible to reach and to maintain LSST image quality specifications. Several concepts need to be analyzed

in more detail to identify the most suitable way of implementing the wavefront sensing method and this is part of the work that we are planning to continue during the LSST R&D phase.

4.2.5.5 Active Alignment System

An active alignment system employing laser tracker technology has been defined to address rigid body position control. The system will perform multiple tasks: establish first-order look up tables for basic telescope alignment (mount-model development), start of night operational setup (reduce the wavefront sensing system capture range), and maintain alignment of the M1/M3, M2, and Camera subsystems. The goal of the operation system is to narrow and/or remove rigid body measurements from the wavefront sensor system.

A design study is being used to model the performance of laser tracker technology within the LSST model. The software package Spatial Analyzer is used to measure fiducials located on the major subsystems and create geometries used to determine positions. Figure 4.2.5-22 shows the LSST model within Spatial Analyzer.

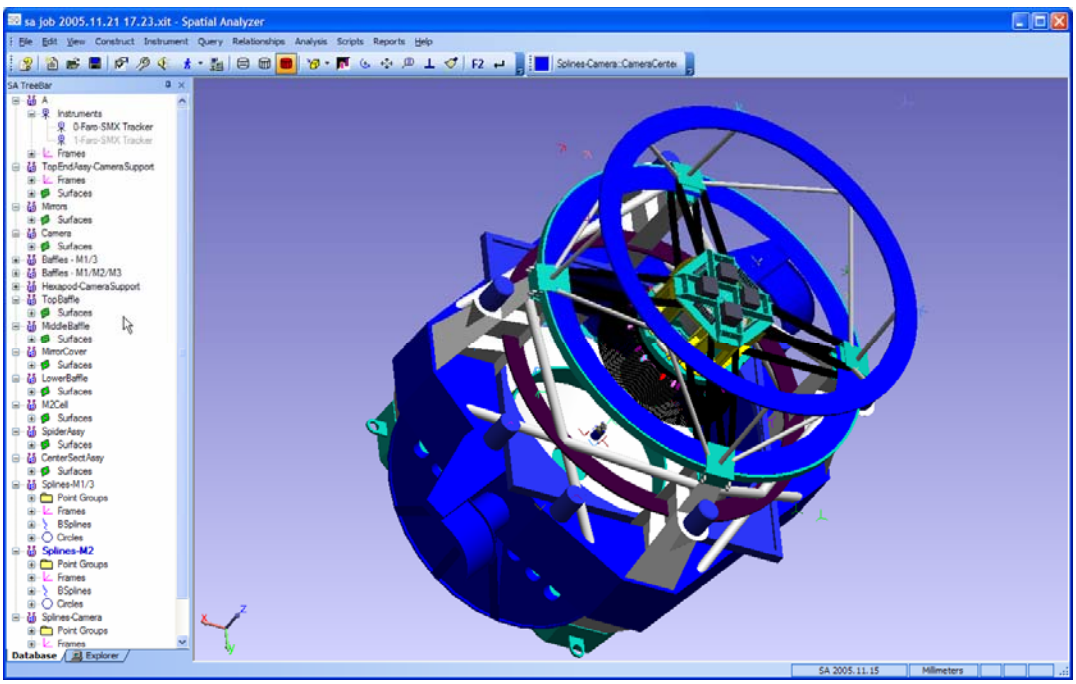


Figure 4.2.5-22 LSST Model in Spatial Analyzer.

Measurement uncertainties associated with the manufacture-stated accuracies, and fiducial measurement conditions (angle, distance, air volume, etc.), are included in the modeling to accurately represent predicted measurement performance. Figure 4.2.5-23 shows a laser tracker suspended within the M1/M3 center hole, with measurement lines-of-sight to the M2 and Camera subsystems. The laser tracker system location is dictated by the available geometry and performance requirements: no allowable obstruction of the telescope beam, single line of sight to all three subsystems, and access for maintenance/calibration. Also depicted in Figure 4.2.5-23 are the measurement uncertainties associated within single fiducial measurements. The measured fiducials are represented by point clouds rather than a specific point.

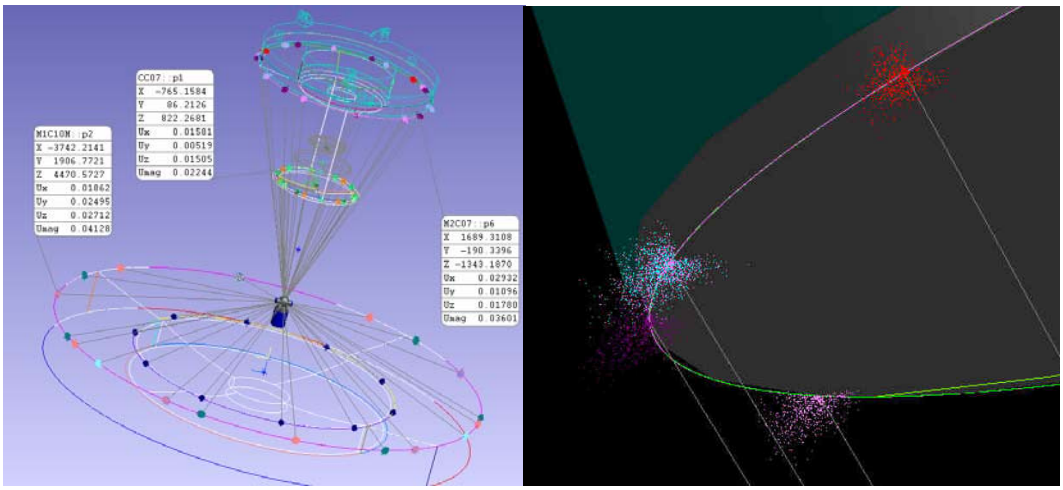


Figure 4.2.5-23 Laser tracker measurement modeling in Spatial Analyzer software.

4.2.6 Telescope Control System

4.2.6.1 Control Architecture

We present the general architecture of the control system leaving aside the detailed design of the individual feedback loops. An active telescope control system should:

(i) acquire a target on the sky,

(ii) track it during the observation,

(iii) correct the wavefront aberrations due to both atmospheric effects and telescope deformations.

There are four major sources of disturbances:

(i) the gravitational deformation of the telescope,

(ii) the thermal expansion of the telescope,

(iii) the wind induced deformation of the telescope, and

(iv) the wavefront deformation due to atmospheric turbulence.

Fully understanding the effects of these sources on the image quality of the telescope needs significant amounts of further investigation and simulations.

The thermal and gravitational effects are potentially large but tend to be rather slow. The temperature changes on a mountaintop during the night are usually very slow. Considering a 1°C/hour maximum slew rate and 5°C swing, the bandwidth of this disturbance is less than $2*10^{-5}$ Hz. The major gravitational effect on the primary mirror is axial deformation. The axial component of the gravitational force f_{ax} is proportional to the sine of the elevation angle h, which in turn is a trigonometric function of the geographic latitude ϕ, star declination δ, and hour angle η:

$$f_{ax} \propto \sin(h) = \sin(\varphi)\sin(\delta) + \cos(\varphi)\cos(\delta)\cos(\eta)$$

While tracking, everything - except the hour angle – is essentially constant, so the bandwidth of gravitational disturbances can be estimated as about 10^{-5} Hz. Assuming smooth enough bearings,

actuators and sensors, the correction of thermal and gravitational deformations should not interact with the structural dynamics of the telescope.

The refraction index fluctuations in the atmosphere don't influence the shape of the telescope, but their effects appear in the optical measurements used to determine that shape. Consequently, although some of the atmospheric effects will be corrected by the telescope control system due to their inseparability from telescope deformations, they have no significant implications on the architecture of the control system.

Figure 4.2.6-1 presents a diagram of the control system architecture. The bandwidth and demands rates are shown, to give an indication of the system's dynamics.

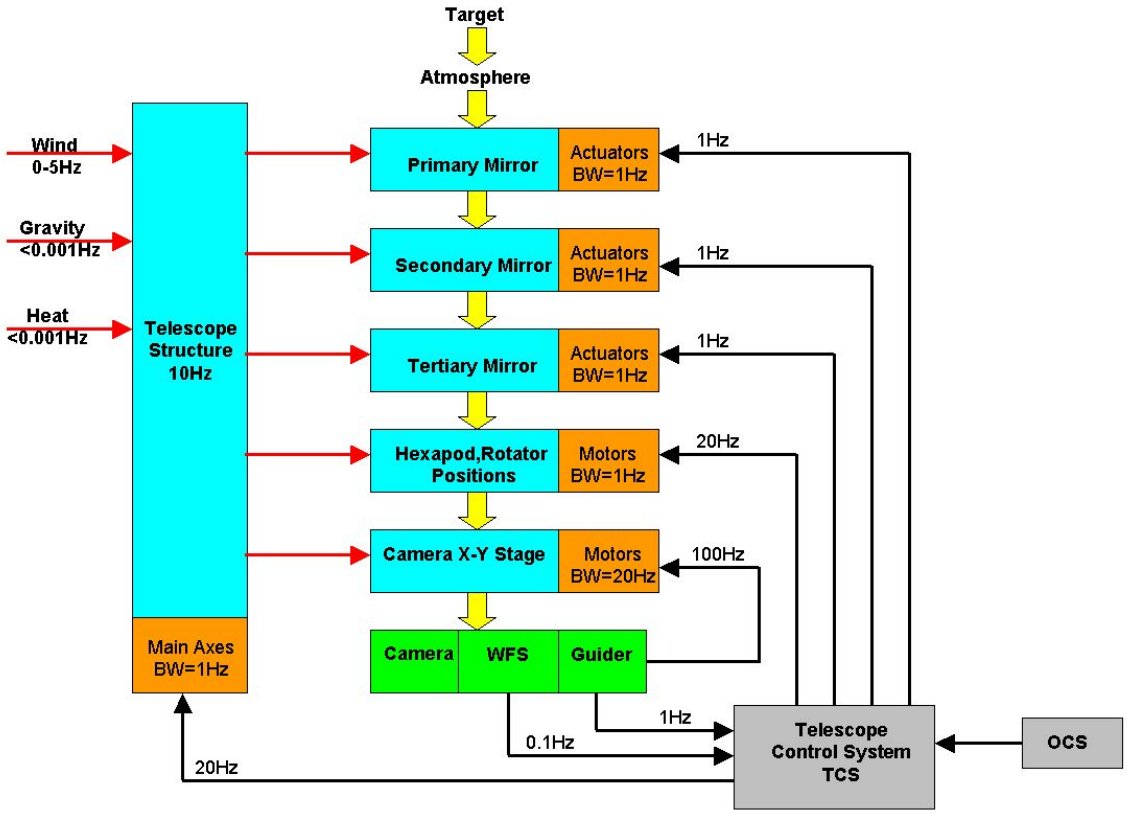


Figure 4.2.6-1 Control Architecture.

4.2.6.2 TCS Design

The main purpose of the Telescope Control System (TCS) software is to accept the target position of a celestial object, which can be given in a variety of coordinate systems, and calculate mount, rotator and optical surface positions, so that the target is imaged perfectly at a given point in the focal plane. Furthermore, the TCS is characterized by the need to integrate a number of heterogeneous subsystems, which exhibit complex interactions. These interactions, although not hard real-time bounded, need a high level of synchronization.

The Telescope Control System (TCS) is the central coordination facility for the delivery of high quality field images to the camera. It is responsible for the precise pointing and tracking calculations necessary to observe a certain field. The TCS does not itself operate any mechanical component; rather it delegates this responsibility to the various telescope subsystems and manages them according to the observation requests.

The TCS design is based on a distributed system model. Under this model, the components interact through well defined interfaces, to accomplish the desired system behavior. The main components in the proposed implementation, are tied together by the use of an Ethernet Bus, thus permitting the efficient exchange of commands and status among them.

The distributed nature of the TCS is complemented by the control model based on a supervisory control strategy. Under this model, a supervisor agent computes the "setpoint" to be applied to a controllable device. The time critical loops are closed locally at the device level, and the device makes status information available for monitoring purposes.

The TCS itself will be controlled either directly by a telescope operator, or by commands initiated by the Observatory Control System (OCS). Its role therefore, is to act as intermediary between the observer(s) and the telescope hardware, translating high level user commands into low level subsystem commands. Consistent with our control model, the TCS will return status information to be distributed system wide.

The following diagram gives an overview of the TCS dataflow:

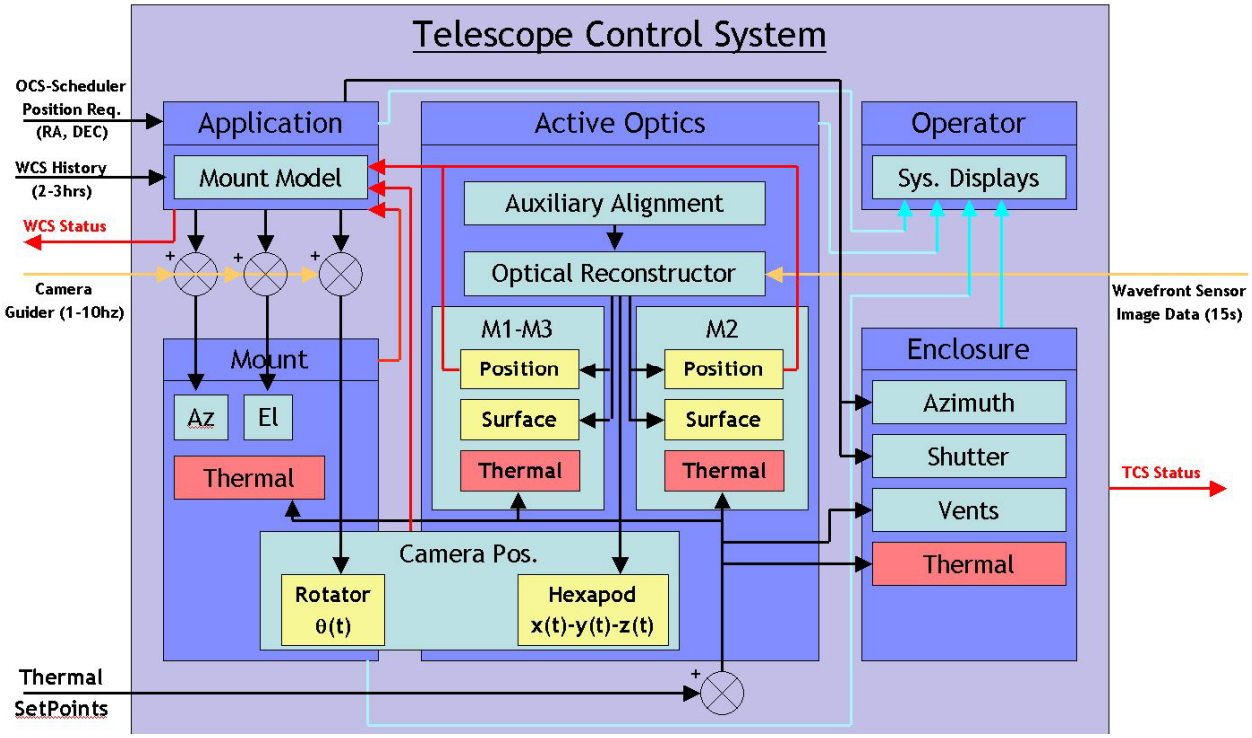


Figure 4.2.6-2 Overview of the TCS dataflow

Based on the diagram, a brief description of the main components follows:

Mount. The LSST telescope mount supports the optics and the camera. The design is based on an alt-azimuth setup with a single focal plane for the camera. The mount component exerts control over the following elements:

- Alignment of the altitude and azimuth axis into the target position.
- Control of the instrument rotator.
- Control of the azimuth and rotator cable wraps.
- Control of mirror covers.

The basic interaction with the TCS consists of accepting azimuth, elevation and rotator position parameters for the target position to guide the two mount axis and rotator, together with the time

these targets should be attained, and the time these demands were sent from the TCS. These parameters may be sent once or they may be continuously delivered at up to 20 Hz. The mount responds with the encoder readings, drive demands, servo errors, alarms and status. The mount reacts also to events generated by a handpaddle that triggers axis motions.

Enclosure. The main function of the enclosure (Dome) is to shelter the telescope from dust and inclement weather, but it also serves to reduce the effect of wind gusts during observations. The telescope enclosure will normally co-track the telescope in azimuth during observations, but in fact can move independently. The enclosure component exerts control over the following elements:

- Enclosure opening by means of a shutter.
- Enclosure rotation in azimuth.
- Vents opening

The basic interaction with the TCS consists of accepting position requests at up to 1 Hz. The enclosure responds with encoder values and status. The enclosure reacts also to events generated by a handpaddle that triggers shutter and rotation motions. Finally, the enclosure supports other equipment like cranes, not related to the actual observations.

Active Optics. The active optics component will control the optical elements in the target optical path, to perform optical alignment, maximize energy concentration, minimize residual aberrations and contribute to image stabilization. The LSST optics control is more complex than similar telescopes, in that the focal plane is optimized not only by the M1, M2 and M3 mirrors, but by optics associated with, or under control of, the camera.

The basic interaction with the TCS consists of accepting azimuth and elevation coordinates, environmental and aberrations parameters. The optics component exerts direct control on M1, M2 and M3, and computes setpoints to be applied to the optics under control of the camera. The component responds with status information.

Application. The application component is where the business logic resides and gets executed. A central part is the “Kernel”, responsible for the computation of the demands to the mount and rotator. The component coordinates the different subsystems by making available setpoints and parameters at the right rate in the right sequence. This component is also responsible for the interactions with other parts of the LSST, like the OCS and CCS. It accepts observing commands and returns status information to be available system wide.

Operator. The operator component is the primary means of interaction for a user with the TCS. It presents information on the state and status of the system in textual and graphics form, and accepts user inputs by means of specialized GUIs. Given that normal operation will be done by means of the OCS, the operator component will be tailored to maintenance and diagnostics operations, instead.

Optical Reconstructor. The optical reconstructor component computes optics aberrations, normally in the form of zernike coefficients, from the images, or image segments, that will be provided by the CCS at a rate to be determined. The reconstructor will generate surface and position corrections, to be applied to the active optics components. Even though the final details of the interactions between optics, WFS and CCS are as yet to be determined, the present control model should apply as well, in that setpoints will be generated for the optics, CCS and mount components.

4.2.6.3 TCS Implementation Plan

The TCS will be developed following a carefully formulated plan which identifies traceable deliverables that provide accurate assessment of project progress. The goal of this program plan is

to deliver a set of design documents leading to the final implementation of the TCS software. The plan consists of the following steps.

Standards. For the development of the LSST software, a series of standards will be adopted. At the present time the Unified Modeling Language (UML) has been adopted for the early phases of analysis and design. In what follows, mention will be made to the tools utilized in each case.

Requirements. The first key step is to analyze all documented requirements and elicit, where possible, all hidden or missed requirements. A good technique is to view the proposed system as a black box. At the highest level the system boundary must be clearly defined. This is to determine what is outside and what is within the system's responsibility. Items outside the TCS include other systems, the roles people take on when they use the system or some more abstract concept such as time. We refer to these items as actors. From a structural perspective, we show these actors in the context of the system, and we show the interface devices or components that are used by the various actors interacting with the system. The delivery then is a System Scope diagram. We have already developed a basic diagram like this, and this step will serve to validate and/or modify our original design.

Functional Requirements. Then the functional requirements are analyzed. This will ensure that the right solution is implemented that fully meets the project requirements. In UML, Use Cases are used for this purpose. The purpose of a Use Case is to define testable system functionality from an outside-in perspective. The properties of a Use Case are that it captures some user visible function, and that achieves a discrete goal for the user. The scope of a typical Use Case Diagram includes all major functionality (Use Cases) associated with an actor. The deliverables here are the Use Case Diagrams.

Areas of Responsibility. In parallel with Use Case definition, the major areas of responsibility need to be defined. These can be used to assign tasks to project teams, identify existing components that meet system requirements, and/or to break the system up into smaller pieces. UML uses packages to show the major areas of responsibility and their interdependencies. Packages group functionalities that provide consistent and coherent services and provide a higher level of granularity than objects. They provide services to support system functionality, which are provided through interfaces. They can contain other packages and their interdependencies can be modeled. The deliveries at this step are Package Dependency diagrams, and interface documents that are utilized when assigning implementation contracts.

4.2.7 Observatory Telemetry System

4.2.7.1 Overview

The LSST Observatory Telemetry System (OTS) is a service provided to all systems, that serves as a central point that will monitor, analyze, and archive conditions and state of the telescope, optics, enclosure, observatory environment, camera, and data management system. It will supply operational conditions (science metadata) for calibration and analysis of scientific data, it will be the database of conditions for real time and periodic analysis, and it will record system actions to support troubleshooting and maintenance.

The OTS will be an important resource during integration, commissioning, operations, maintenance and data processing, and an important tool for meeting LSST's goals of operational efficiency and data quality. As such this system will be fully operational during the early integration phase, and in the long run it will operate in a manner consistent with a complex industrial machinery environment, where minimizing downtime is critical.

All Systems will provide real time status and diagnostics data to the Observatory Telemetry System Database server, at the highest sampling rate possible, without any restrictions on data bandwidth and storage.

The LSST Observatory Telemetry System will be capable of providing:

- Data Storage
- Data Analysis
    - Real Time Tactical Displays
- Event Detection
    - Real Time Fault, Error Detection and Notification.
- Publishing
    - Historical Trend Statistical Analysis
    - Science Metadata
    - Early failure detection
    - Automatic Report Generation
    - Internet based data mining tools

4.2.7.2 Architecture

The LSST Observatory Telemetry System will gather real time information from all systems and store it in the OTS centralized Facility Database Server as shown in Figure 4.2.7-1.

Once the data is stored the OTS Data Analysis Module will process it. This Module will simplify, organize and share, the large amount of information generated by the operation of the LSST system.

The OTS Event Manager Module will monitor the system for out of limit conditions and alarms, and notify the appropriate personnel for immediate attention and action if necessary.

The OTS Publisher Module will be the interface to the outside world, it will be responsible for web publishing and report generation, and it will provide an interface to data mining.

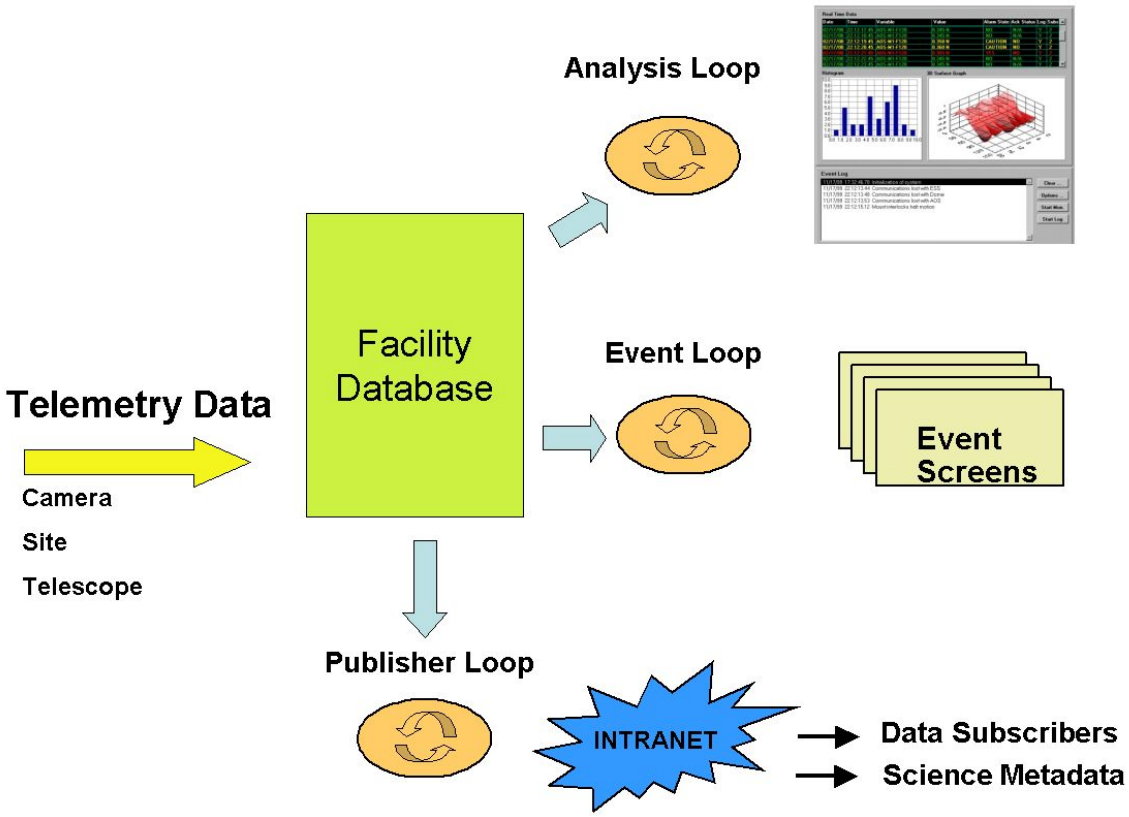


Figure 4.2.7-1 Observatory Telemetry System Architecture

4.2.7.3 Design

The LSST Observatory Telemetry System will be designed and implemented into all components of the LSST complex. The data producers will specify the data types and bandwidth required, with the system being flexible enough to allow for changes in the data definitions and data rates. This specification will be synchronized by maintaining a configuration database for each data producer, that is available to the OTS Database Server.

The data from each subsystem will fall into three categories:

- Continuous Data: This encompasses the data that needs to be produced regularly at a fixed sampling rate; this data will drive the real time monitors and alarm system.
- Asynchronous Data: This encompasses data produced by the normal communications and fault/error messages between computers
- On-Demand Data: This encompasses the data that is produced during a special test or calibration sequence.

4.2.7.4 Implementation

The LSST Observatory Telemetry System shall be implemented as a group of independent data producers, each sending telemetry data to a central data logger. The implementation will utilize the same data distribution services middleware, that is responsible for the exchange of messages and status between the distributed elements of the LSST system.

As an example the Telescope System will provide the following types of data:

- Telescope Mount AZ, EL, Rotator Dome, and Shutter:

    - Position, Rates, Error, Torque, Currents
    - Raw Encoder data
    - Limit Switch/Interlock States
- Mirror Axial Support Actuators: Forces, Positions, and Error
- Lateral Support Actuators: Forces, Positions, and Error
- Environmental Control System:
    - Temperature Sensors: Outside and Inside Air, Telescope, and Optics
    - Weather Station data
    - Wind Sensors at telescope
    - Chiller Parameters: Glycol Temperature, Pressure, and Flow Rate
- Other Engineering Sensors:
    - Strain Gauges and Load Cells on Load Bearing Areas
    - Absolute Linear Gauges on Optics and Cell Interfaces
    - Accelerometers for Special Vibration Analysis

4.2.8 Telescope Integration

Telescope integration and testing is a process that begins with the initial design and specification of subsystems and continues through to the final summit assembly. The telescope will be designed with the final assembly process and the performance requirements as guiding tenets. In addition, the telescope will be procured in subsystems that, to the extent possible, can be pre-assembled and fully tested at vendor factories. This approach will allow significant debugging and subsystem commissioning to occur long before the parts arrive at the remote summit location. At the summit, integration will be staged with clearly defined tests and commissioning tools that will verify each step meets its performance goals, thereby limiting the complexity of the final on sky testing and commissioning phase.

4.2.8.1 Design and Development

The critical subsystems within the telescope as well as the camera will be designed to include measurement fiducials. These will be surfaces or targets that reference the critical dimensions of the part or sub system at a location accessible during both the initial fabrication and later in the assembled system or subsystem. The fiducial will be a permanent feature or in some cases a custom dedicated metering fixture that can be temporarily installed when it is necessary to extend the measuring point. For example, the reflective mirrors will have a vertex tool, a temporary lightweight fixture that allows a target to be placed at the vertex of the optical surface where there is no substrate. Permanent targets are also to be attached to the substrate. All of these must be designed into the system so in process metrology tools can account for and register the fiducials.

For the three reflective mirrors, the telescope structure, and the camera, the permanent fiducials will also provide operational alignment reference. The telescope system will have an auxiliary system that provides rigid body alignment feedback during operation to limit the misalignments and reduce the load on the higher resolution WFS. To achieve both the integration and operational alignment tasks, the elements and the alignment systems will be engineered together as a single system. The critical tight tolerance systems and elements will require a series of fiducials that may include retro reflectors as well as simple alignment marks.

The system being investigated to support integration and operational alignment needs is a laser tracker. Several tools are commercially available and optical manufacturing has included custom versions for profiling large surfaces. Figure 4.2.8-1 shows one example of such a laser tracker and retro reflectors used in some systems.

Figure 4.2.8-1 Example of Laser tracker head and retro reflectors under consideration for integration and operational alignment

Published performance for one such system is found in Table 4.2.8-1

Table 4.2.8-1 Commercial Laser Tracker Measurement Performance

Absolute	+/- 20 microns + 1.1 micron / meter Accuracy
Differential	+/- 7 microns + 1 micron / meter Repeatability 0.06 arc seconds resolution +/- 9 microns + 2 microns / meter accuracy +/- 3 microns + 1 micron / meter repeatability
Range	35 Meters

4.2.8.2 Factory Testing

The LSST telescope will be designed and developed as a single integrated system and procured as large sub-systems. These subsystems will be aligned with general industrial capability such that contracts remain, to the extent possible, within the scope of existing organization and single industry expertise. This insures that procured subsystems incur little additional risk and unnecessary project management. Each sub system will have normal acceptance criteria and for systems like the telescope mount and the large mirrors the contractors will pre-assemble the system in their factories for additional testing and performance verification.

Factory assembly and testing have been successful in many previous astronomical telescope development efforts. For the SOAR telescope the factory commissioning work allowed for rapid assembly and very short summit integration efforts. Figure 4.2.8-2 shows the 4.2 meter SOAR

telescope mount during its factory pre-assembly testing. Complete assembly at the fabricator and testing with a complete complement of surrogate masses allowed extensive testing and efficiently resolved numerous integration and commissioning issues. A similar approach will be taken with LSST.

Figure 4.2.8-2 Factory assembly and testing of SOAR telescope mount (structure and control system). A similar approach will be used for LSST.

4.2.8.3 Summit Integration

Integration of the LSST will be controlled as a series of well planned efforts performed in parallel and series as dependencies allow. The basic integration plan is shown in Figure 4.2.8-3. This Gantt chart shows the general flow and the potential critical paths for integrating the system and reaching first light in December of 2012. To support this plan the facility construction must commence early and support staged completion so the coating support facility is available for occupancy earlier then the remaining facilities. Also the dome must be started during the facility

construction so that mount telescope mount integration can begin just after facility and dome completion.

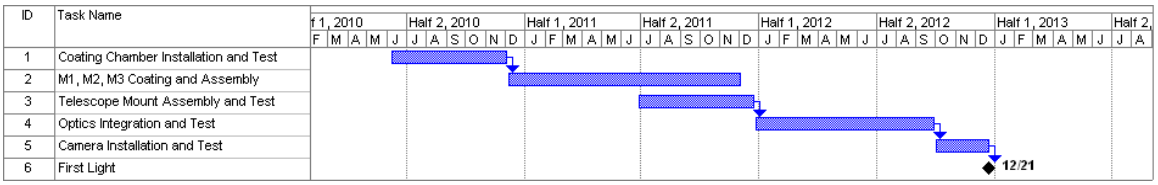


Figure 4.2.8-3 LSST summary integration Gantt chart for critical tasks.

4.2.8.3.1 Optics Integration and Test

The three reflective optics will be integrated into the telescope sequentially. With the deployment of each mirror the telescope configuration will be adjusted to support a camera for individual mirror testing. Figure 4.2.8-4 shows the optical configuration to support testing of each new mirror surface.

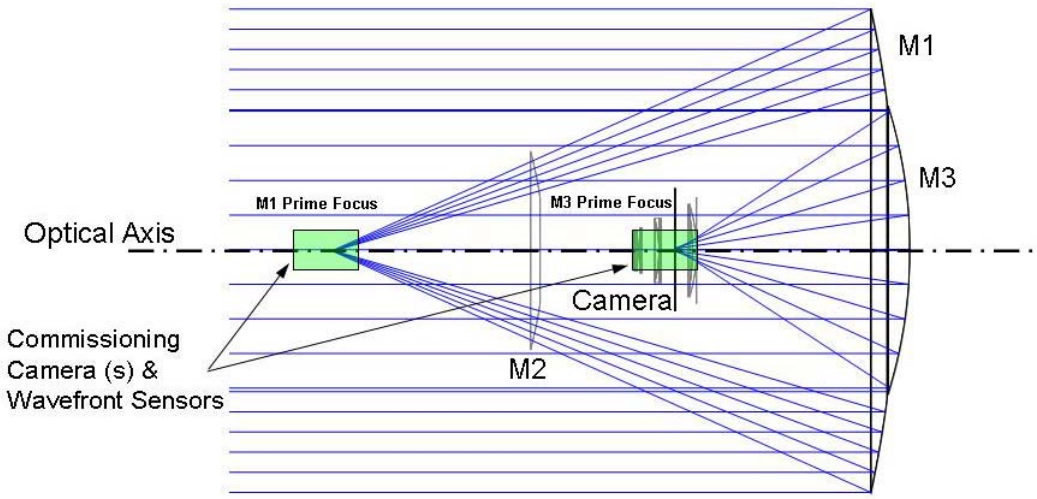


Figure 4.2.8-4 Optical configuration showing two positions for prime focus cameras to test M1 and M3 and the commissioning camera in place of the LSST Camera to test M2 and the three optic system

The primary and tertiary monolithic mirror will be installed and a custom telescope top end installed to deploy a primary mirror prime focus camera and corrector. This preliminary camera and set-up will allow the mirror control algorithms and the elevation dependency models to be verified and tuned for on telescope performance. The initial actuator influence functions will have been established during factory testing under an interferometer but the elevation dependency model will have been preset to analytic values. This first integration step allows the effort to be focused on the primary mirror alone without the added complexity of additional active mirror surfaces. The top end can be replaced and the tertiary surface can be tested independently like the primary mirror.

Following the commissioning of the primary and tertiary mirrors the secondary mirror can be tested. Secondary mirror testing will rely on the presence of M1 and M3 but with these two mirrors commissioned the degrees of freedom in the test are reduced significantly. Additionally, the optical system has a benefit that the three reflective surfaces provide a spherical focus on axis. This layout, shown in Figure 4.2.8-5, significantly eases corrector needs for this layout.

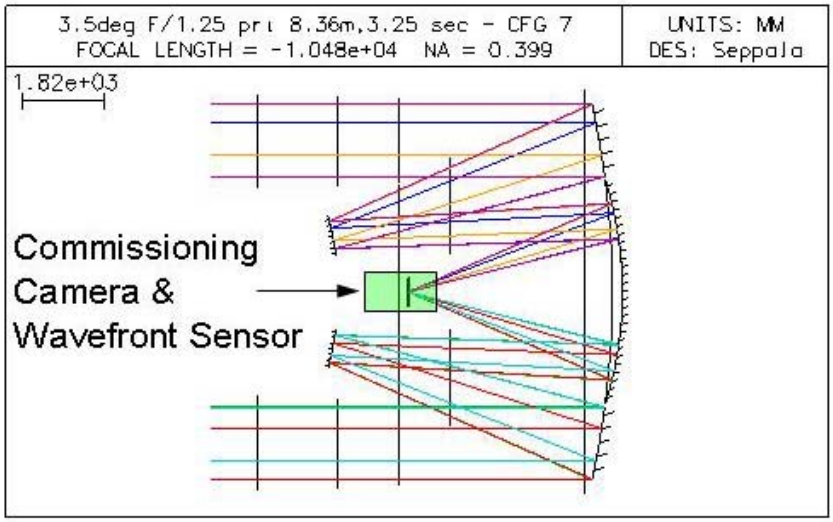


Figure 4.2.8-5 LSST optical design delivers spherical wavefront at camera entrance on axis.

4.2.8.3.2 Integration Personnel

A core integration team will be on the summit for the duration of assembly, testing, and commissioning efforts. Extra teams will augment this group to focus on specific subsystems on schedule with specific hardware deployment. The specialized teams will have been part of the subsystem assembly and testing at the factory so they will arrive at the summit with direct experience prepared for reassembly a second time.

4.3 Camera

4.3.1 Overall Description

The LSST camera is a wide-field optical (0.4–1 μm) imager designed to provide a 3.5° FOV with better than 0.2 arcsecond sampling. The image surface is flat with a diameter of approximately 64 cm. The detector format will comprise a mosaic of 16 Mpixel silicon detectors providing a total of approximately 3.2 Gpixels. The camera includes a filter changing mechanism and shutter. It is positioned in the middle of the telescope where cross-section area is constrained by optical vignetting and heat dissipation must be controlled to limit thermal gradients in the optical beam. The camera must produce data of extremely high quality with minimal downtime and maintenance.

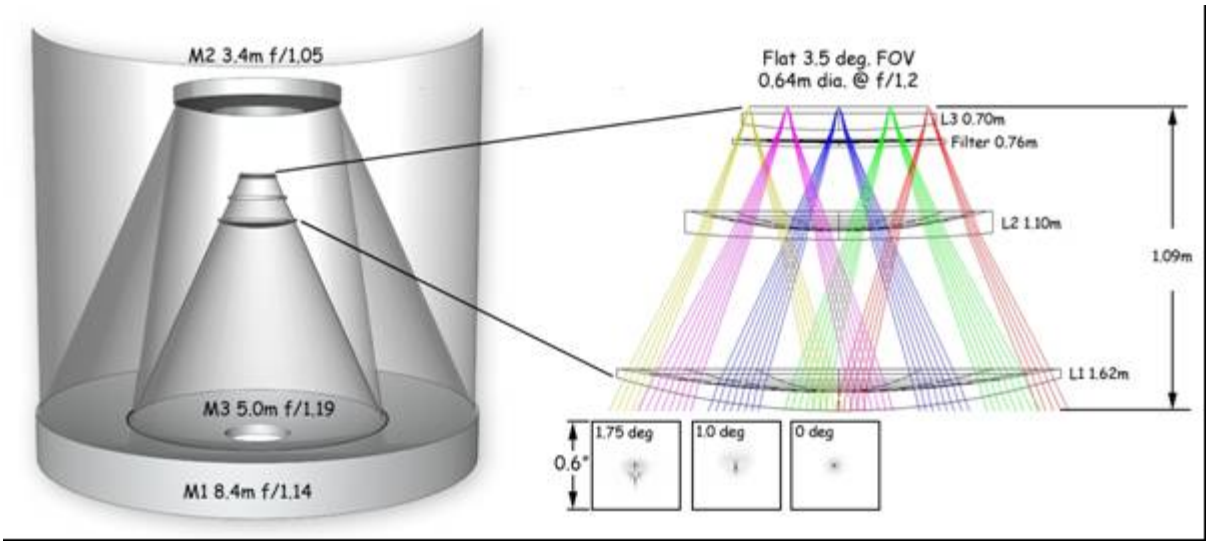


Figure 4.3.1-1 Optical layout of the LSST telescope (left) and camera elements (right).

The camera concept currently under development is shown in the accompanying figures. Figure 4.3.1-1 shows the optical layout of the camera corrector optics, filter and focal plane. Figure 4.3.1-2 shows a cross-section illustration of the major components of the camera, while Figure 4.3.1-3 shows an cutaway view of the camera and its components. The focal plane array operates at a temperature of approximately −100°C to achieve desired detector performance. The focal plane array is contained within an evacuated cryostat which incorporates detector front-end electronics and thermal control. The lens L3 serves as an entrance window and vacuum seal for the cryostat. Similarly, the lens L1 serves as an entrance window and gas seal for the camera housing, which is filled with a suitable gas to provide the operating environment for the shutter and filter change mechanisms. The filter carousel can accommodate 5 filters for rapid exchange without external intervention.

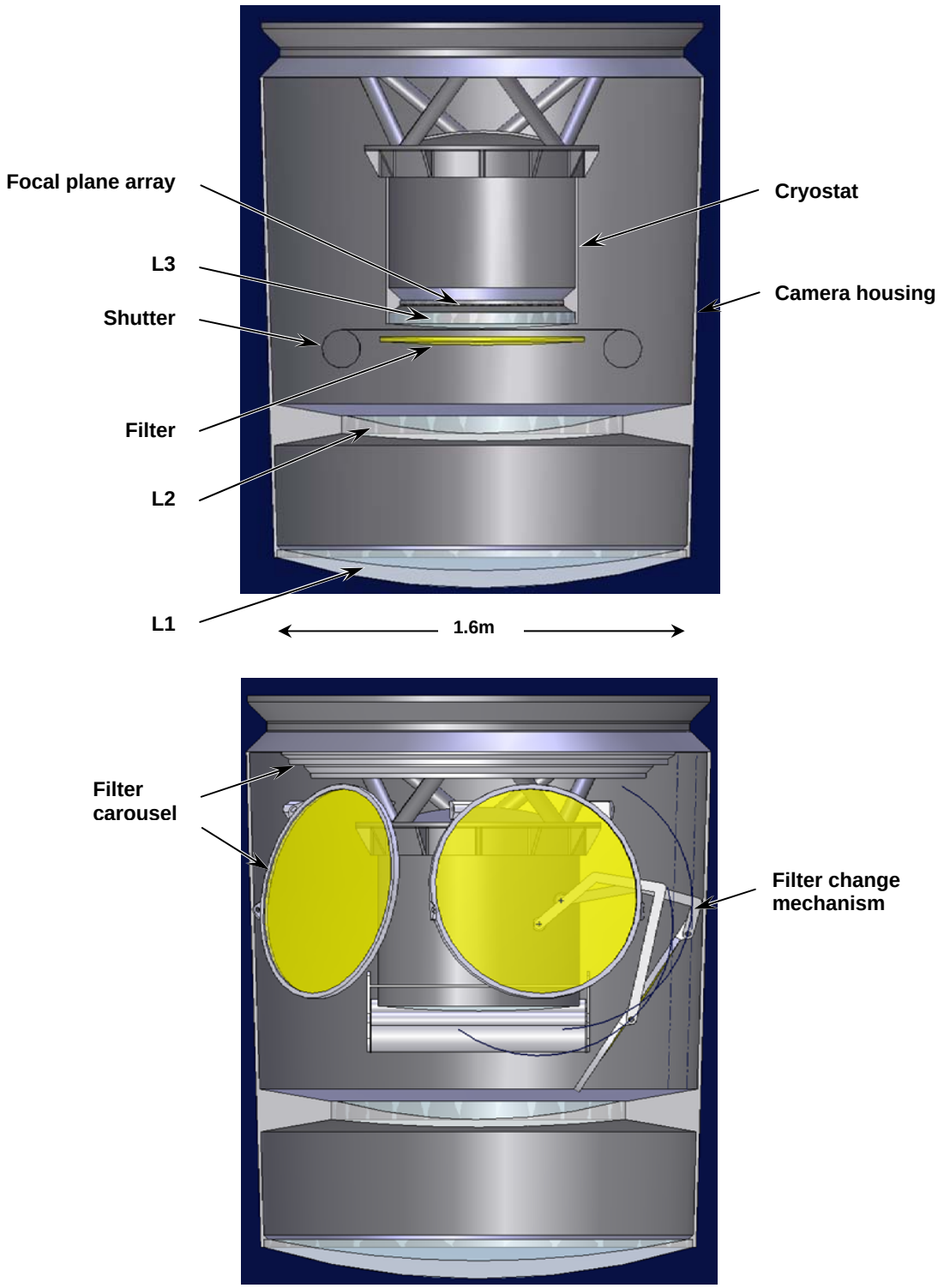


Figure 4.3.1-2 Cross-sectional views of camera showing major camera components. In the upper illustration, the filter carousel and changing mechanism have been omitted for clarity; the lower illustration has been rotated 90° to show two views of the shutter mechanism and its relationship with the filter change mechanism.

The camera mechanical mount provides proper support and registration to the telescope and incorporates provisions for adjusting camera position and orientation to compensate for alignment variations with telescope elevation. In addition, the camera axial position must be adjusted to optimize focus at different filter wavelengths (the axial position of L2 is similarly adjusted). Additional camera interfaces include electrical power, thermal cooling, and fiber optic connections for control and data interfaces. The following sections describe these features in detail and discuss the considerations leading to the current design concept.

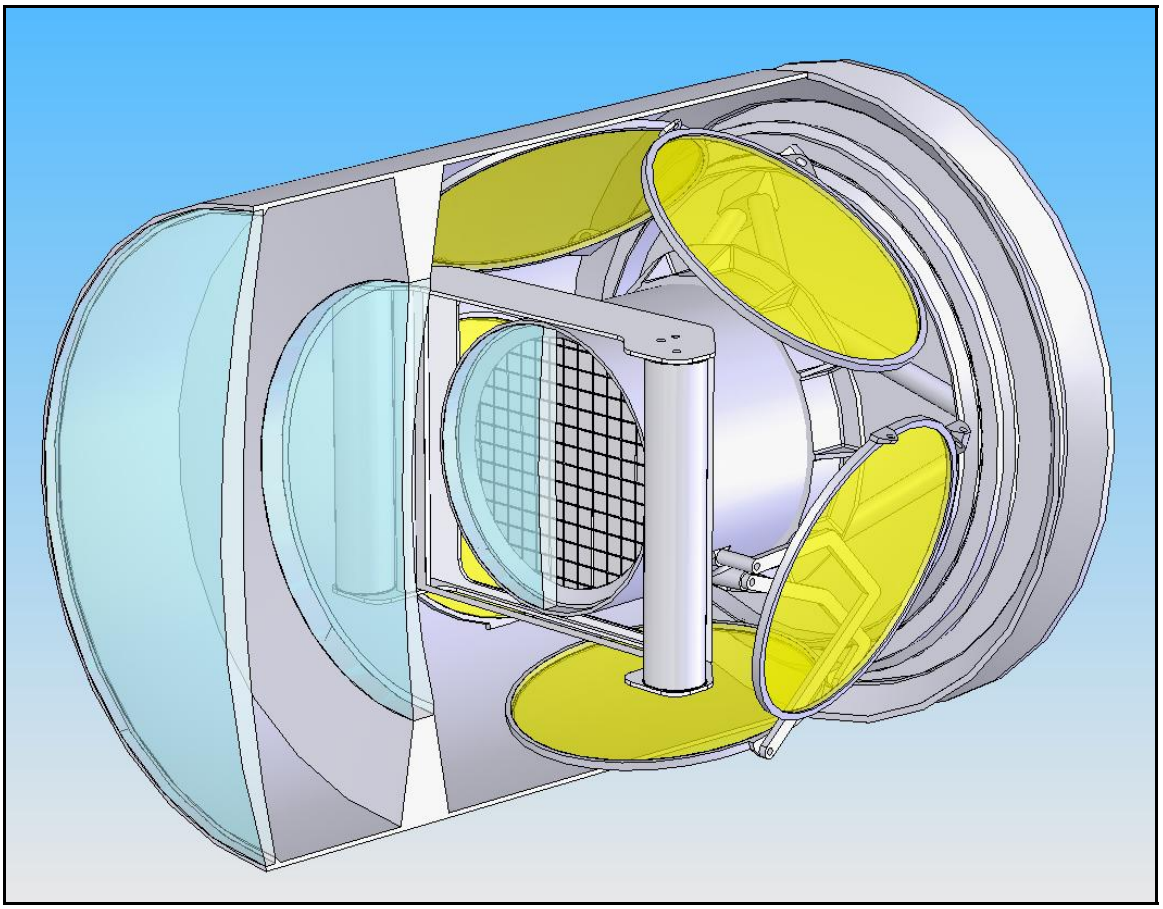


Figure 4.3.1-3 Cutaway view of the camera housing and components

Section 4.3.2 presents the requirements and considerations which drive the sensor design and selection, and a complete strawman sensor design concept is presented and discussed. Section 4.3.3 discusses the requirements and considerations leading to the conceptual design for camera electronics. Section 4.3.4 provides the requirements, considerations and conceptual design solutions for packaging the sensors and focal plane assembly. Section 4.3.5 discusses the design considerations and requirements for the mechanical structures and mechanisms. Section 4.3.6 presents the thermal requirements and design considerations, and describes the conceptual design details for implementing the thermal requirements. Section 4.3.7 discusses design considerations and requirements for the corrector optics and color filters, while the data acquisition and control concepts are described in Section 4.3.8.

4.3.2 Sensor Design

4.3.2.1 Requirements and Design Approach

The LSST science goals lead to a set of challenging performance requirements for the focal plane sensors, listed in Table 4.3.2-1:

Table 4.3.2-1 Science requirements driving sensor design

Science Requirement	Design Implications
High QE out to 1000nm	thick silicon (> 75 μm)
PSF << 0.7”	high internal field in the sensor high resistivity substrate (> 5 kΩ·cm) high applied voltages (> 50 V) small pixel size (0.2″ = 10 μm)
Fast f/1.2 focal ratio	sensor flatness < 5μm peak-to-valley focal plane package flatness < 10 μm p-to-v package with piston, tip, tilt adjustable to ~1 μm
Wide FOV	3200 cm^2 focal plane > 200-CCD mosaic (~16 cm^2 each) industrialized production process required
High throughput	> 90% fill factor 4-side buttable package, sub-mm gaps
Fast readout	highly-segmented sensors (~6400 output ports) > 150 I/O connections per package

A list of requirements is given in Table 4.3.2-2.

Of central importance are high quantum efficiency (QE) extending into the near infrared and small contribution to the instrument point spread function (PSF) budget. The requirement for a large $A\Omega$ leads to a large focal plane area, which can be economically realized only by using silicon as the sensor material. To achieve high QE in the near-IR, the sensor must be thick because the absorption depth of silicon increases rapidly in this wavelength range. However, increasing detector thickness degrades the spatial resolution of the sensor due to two effects. Diffusion of the photogenerated charge increases because of the longer transit time to the collecting electrode. Thick sensors also require sufficient substrate bias to fully deplete the device; otherwise, lateral diffusion in the undepleted field-free region severely degrades the PSF. A second cause of PSF broadening results from the fast focal ratio of the LSST optics (f/1.2). For red wavelengths where the absorption length is long, the light becomes defocused before it is fully absorbed, further broadening the PSF. Other less pronounced drawbacks of thick sensors include higher dark current and increased contamination from cosmic rays.

Table 4.3.2-2 Requirements table (from "THE LARGE SYNOPTIC SURVEY TELESCOPE DESIGN AND DEVELOPMENT PROPOSAL" submitted to the National Science Foundation by the LSST Corporation, December, 2003)

	Allowable range	**Target**	**Units**
Pixel size	8 – 12	10	μm
Flatness deviation	10	5	μm
Aggregate fill factor (entire array)	90	95	%
Frame read time	3	2	s
Read noise	10	6	e^-
Full well	70000	90000	e^-
Output-output crosstalk	.05	.01	%
Nonlinearity	7	5	%
Dark signal (95th percentile)	4	2	$e^- s^{-1}$
Charge memory (residual image after one readout)	.05	.02	%
QE at 400nm	55	60	%
QE at 600 nm	80	85	%
QE at 800 nm	80	85	%
QE at 900 nm	60	85	%
QE at 1000 nm	25	45	%

We have performed a detailed study of sensor thickness, and the guidelines are summarized in Sections 4.3.2.1 and 4.3.2.2.

Sensor format, readout speed, readout segmentation and manufacturing considerations are discussed in Section 4.3.2.3. Based on this, a "strawman" design, illustrated in Section 4.3.2.4, was developed. Signal processing and readout of the strawman CCD are briefly described in Section 4.3.2.5, and in more detail in Section 4.3.3. The QE and PSF analysis and sensor thickness considerations are equally applicable to either CCD or hybrid PIN-CMOS sensors. Also, the sensor size, the pixel size, the packaging and assembly into the focal plane are largely independent of the sensor technology. We have elected to pursue CCD technology as the baseline, and PIN-CMOS as the backup. Considerations leading to this approach are discussed in Section 4.3.2.7. The necessary sensor testing program is outlined in Section 4.3.2.8. From the LSST science requirements and the state of CCD technology it is clear that some developments will be needed. These are emphasized in Section 4.3.2.9, while the development plan with a request for R&D funding is described in Section 5.3.1.

Quantum Efficiency (QE) and Sensor Thickness

Sensor quantum inefficiencies arise from: *a*) reflection loss, *b*) incomplete charge collection, and *c*) incomplete light absorption. The reflectivity properties of the final LSST sensors will depend on the antireflection coating used on the illuminated surface and on the interface properties of the opposite (charge-collecting) side. Charge collection is expected to be near unity over most of the wavelength range, but surface defects will cause a falloff toward the blue end where absorption takes place very near the surface. For the red and near-IR region, reflection losses will reduce the QE and multiple internal reflections will cause interference fringes in the optical response, Groom *et al*. [2] (although for thick sensors and low f-number the fringing is reduced). The first two causes of QE inefficiency (*a* and *b*) can be reduced by sensor design and

processing. The upper limit of QE is determined by *c*) incomplete light absorption, which determines *internal quantum efficiency.*

The absorption length is defined as the depth at which the light intensity falls to 1/*e* of the original incident intensity. Note that absorption length increases rapidly for photon energies near the band gap of silicon (about 1100 nm). There is also a decrease of absorption length with temperature, especially in the red, since transitions near the indirect band gap must be phonon-assisted to conserve momentum. Internal QE as a function of wavelength, sensor thickness and temperature is plotted in Figure 4.3.2-1 and Figure 4.3.2-2. The sensor temperature range to control thermally generated dark current is indicated in Figure 4.3.2-2. To satisfy the LSST QE requirement of 25-45% at 1000nm, a sensor thickness larger than ~ 75 μm is required. Figure 4.3.2-3 shows temperature dependence of QE as a function of wavelength. For photometry measurements at long wavelengths with accuracy better than 1%, the sensor temperature will have to be controlled to within about 0.25 °C. However, temperature variations across the focal plane of a few degrees can be tolerated provided they remain stable between calibrations.

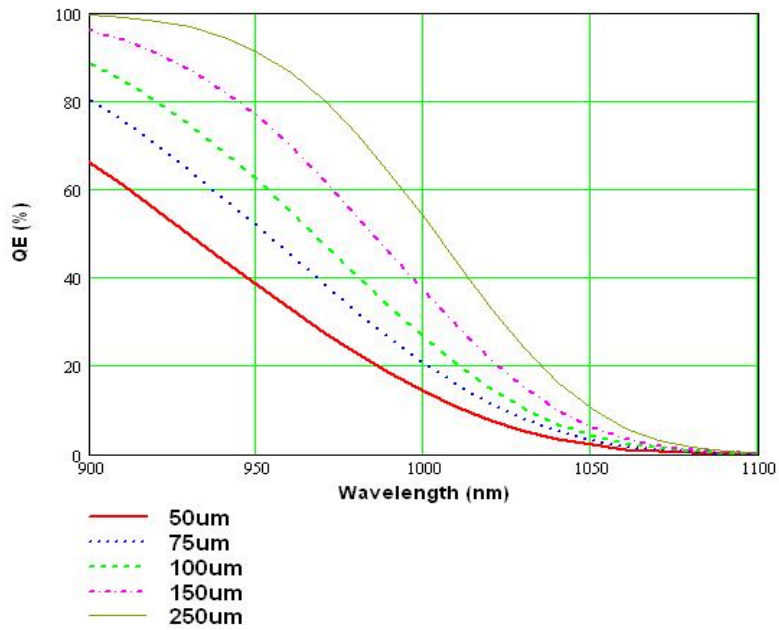


Figure 4.3.2-1 Internal quantum efficiency of silicon as a function of wavelength, for thicknesses of 50, 75, 100, 150, and 250 μm.

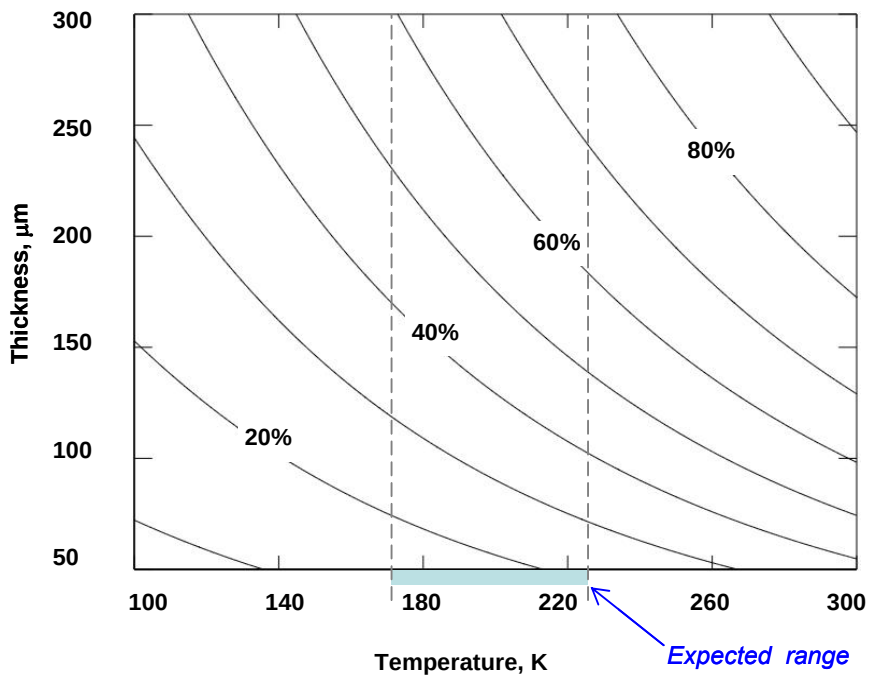


Figure 4.3.2-2 Contour plot showing the dependence of QE on temperature and silicon thickness, for 1000 nm wavelength. The blue bar shows the expected operating temperature range for the LSST camera.

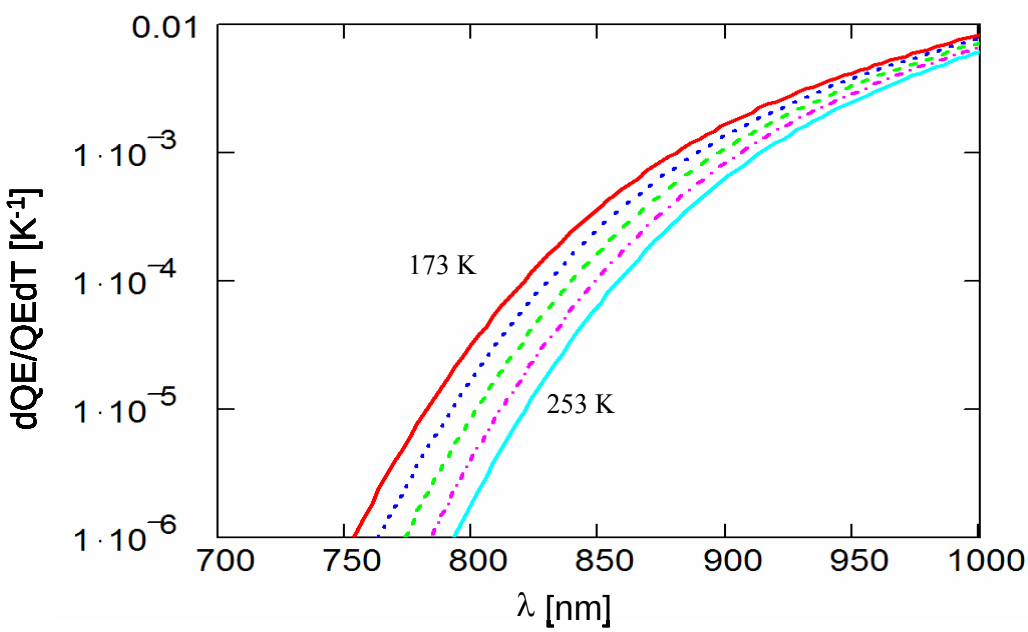


Figure 4.3.2-3 Temperature coefficient of quantum efficiency as a function of wavelength. Thickness = 100 μm. At long wavelengths the temperature coefficient approaches 1% per degree C.

4.3.2.2 Point Spread Function (PSF) and Sensor Thickness

Several of the science missions of the LSST, including weak gravitational lensing, depend on its ability to resolve faint point sources. The signal-to-noise ratio for a point source depends inversely on the PSF, since the smaller the area over which the image is spread, the less sky noise is acquired with the signal. Contributions to the PSF come from aberrations of the telescope optics (including fabrication and alignment tolerances), atmospheric seeing at the site, and sensor contributions. The LSST specification for the detector contribution to the point spread function is shown in Table 4.3.2-3 [3]:

Table 4.3.2-3 Allowable detector contribution to point spread function

	Allowable	**Target**
Pixel FWHM (charge spreading)	< 10 μm	< 7.5 μm

At short wavelengths, for which charge generation is near the illuminated ("back") surface, the main contribution to charge spreading comes from diffusion. At near-IR wavelengths, for which the charge is generated all along the path of light in the sensor, there is additional broadening due to the divergent "cone" of light that enters the sensor. The light becomes progressively defocused and the deeper the point of absorption, the wider the effective spot size.

Charge spreading mechanisms in the sensor are illustrated in Figure 4.3.2-4.

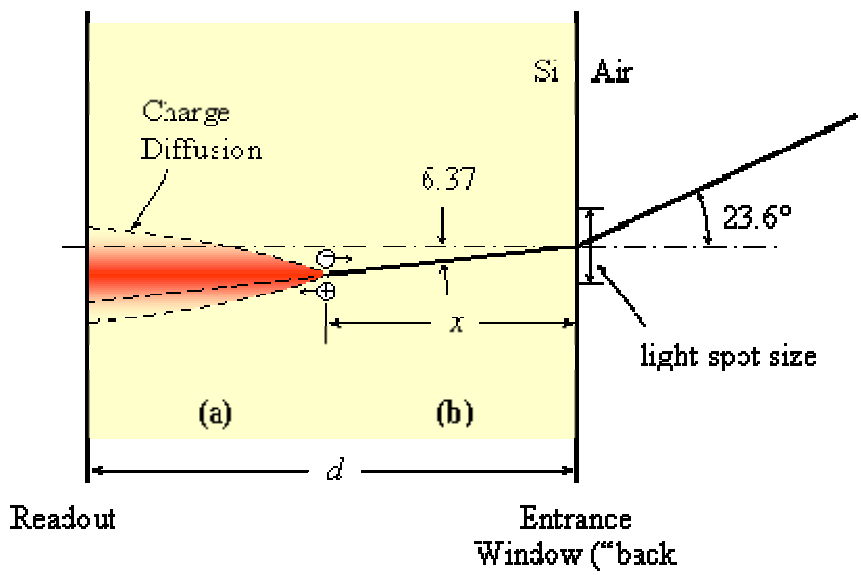


Figure 4.3.2-4 Diagram illustrating two detector-related contributions to the point spread function. (a) Photogenerated charge experiences diffusion as it drifts to the collecting electrodes, broadening the PSF approximately as the square root of the drift distance. (b) At longer wavelengths the optical beam penetrates deep into the silicon and forms a conical volume over which charge is generated. The conical half-angle is arctan(1/2nf), where n is the index of refraction of silicon and f is the focal ratio of incident light.

4.3.2.2.1 Charge Diffusion

Visible photons are absorbed in a thin layer near the illuminated (back) side of the sensor and the resulting charge moves to the front under the influence of the applied electric field. Charge

drifting through a distance *d (= sensor thickness)* in a constant electric field *E* will experience lateral diffusion resulting in a Gaussian charge distribution of width

$$\sigma_D \approx d\left(2V_{th}/V_{op}\right)^{1/2} \qquad (4.3.2\text{-}1)$$

where $V_{th}= kT/q$ is the thermal voltage, and $V_{op}=Ed$ is the sensor operating voltage. It is important to apply a bias (operating voltage) that will deplete the entire sensor thickness. If not, there will be a field-free, neutral layer of thickness d_{FF} near the illuminated ("back") side of the sensor. Charges generated in the field-free region diffuse isotropically and will have a profile of rms width $\approx d_{FF}$ when they encounter the electric field that sweeps them to the collecting side of the sensor. The bias needed to fully deplete a device of resistivity ρ_{sub} and thickness *d* is

$$V_{depl} = d^2/\left(2\varepsilon_{Si}\mu\rho_{sub}\right) \quad (4.3.2\text{-}2)$$

where μ is the mobility of free majority carriers in the substrate and ε_{Si} is the permittivity of silicon. The constant field expression (Eq. 3.3.2-1) for diffusion becomes accurate enough, once the operating voltage exceeds the depletion voltage ($V_{op}/V_{depl} \geq 1.1$).

The PSF FWHM due to the diffusion alone as a function of sensor thickness, operating voltage and substrate resistivity is plotted in Figure 4.3.2-5. Depletion voltage is indicated. It should be noted that for a 100 µm thick sensor, the operating voltage will have to be in the range of 30-40 volts to satisfy the LSST PSF requirements. It is also apparent that the substrate resistivity will have to be > ~ 5 kohm cm, so that the sensor can be operated overdepleted (which improves the PSF).

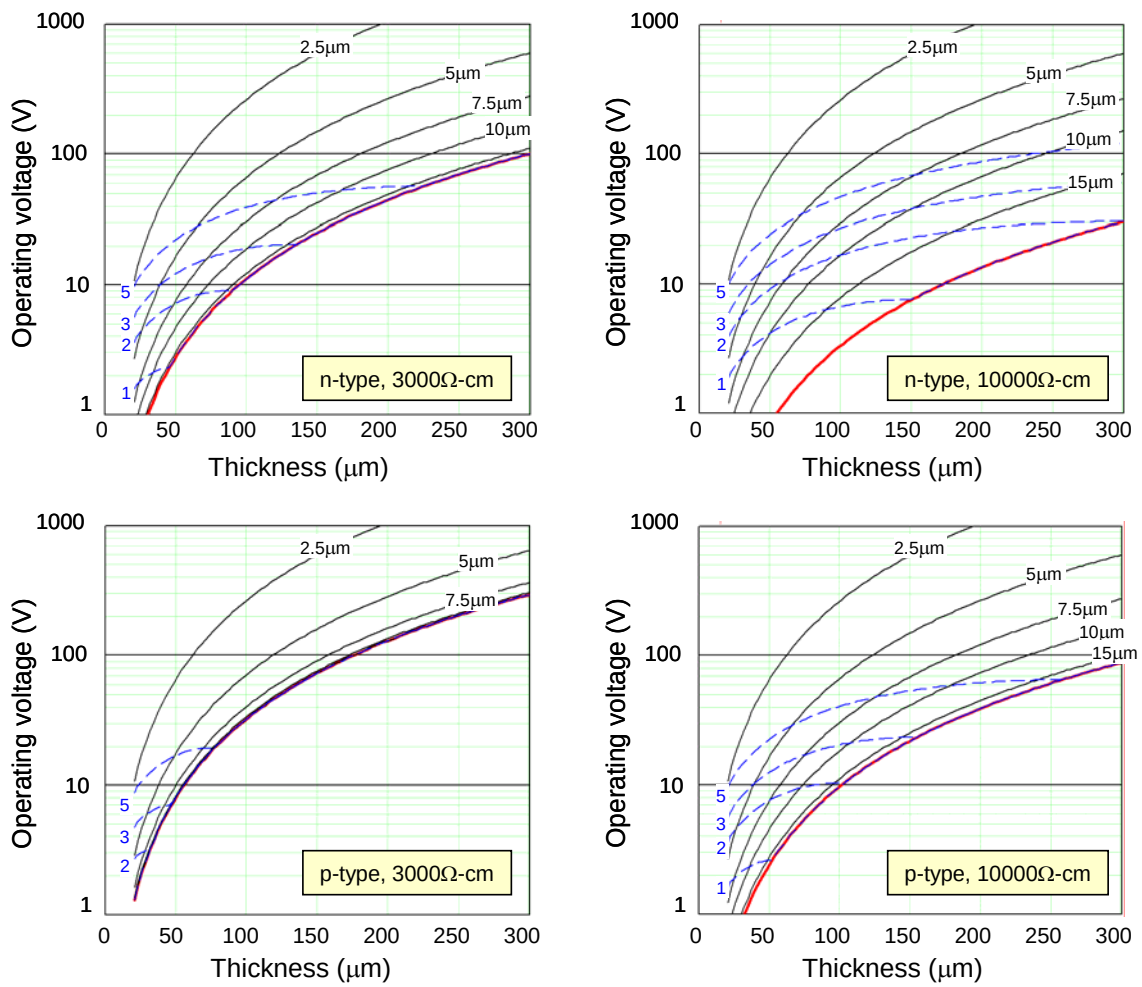


Figure 4.3.2-5 Contour plots showing the dependence of PSF on sensor thickness, operating voltage, resistivity, and conductivity type. Solid black lines are contours of constant g-band diffusion (FWHM shown). Temperature 173 K. Red lines mark the boundary of full depletion. The dashed blue lines show the contours of maximum electric field in kV/cm. Note that the high resistivity n-type material allows small PSF to be obtained at the lowest E-field.

4.3.2.2.2 PSF broadening due to divergent optical beam in silicon

In a low-*f* number optical system like LSST, light is incident on the sensor at large angles from the normal. Figure 4.3.2-6 shows an optical raytrace result from the LSST f/1.2 configuration , for rays incident on the sensor from the right. Because of the high index of refraction of silicon, the light "cone" inside the silicon has a half-angle of only 6.4° compared to the 23.6° maximum angle in air, as indicated in Figure 4.3.2-4. For long wavelengths where *absorption length* is much greater than the sensor thickness, the light is absorbed almost uniformly, while shorter wavelength light is strongly absorbed at the surface. A set of Monte-Carlo simulations combining the effects of beam divergence and diffusion has been performed. A study of the charge spot size as a function of the displacement of the focal plane was also made using the raytrace simulations. When the sensor is displaced in the direction towards the incoming rays, a situation shown in Figure 4.3.2-6 arises where long-wavelength rays come to best focus *within* the sensor.

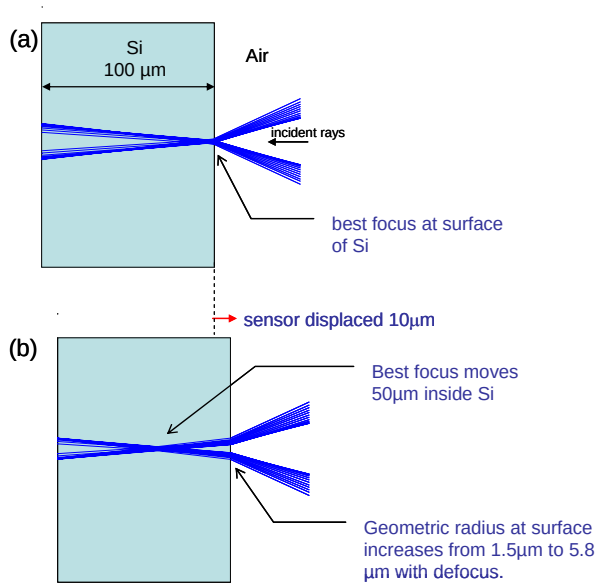


Figure 4.3.2-6 Effect of displacement of the focal plane: (a) position of best focus for short-wavelength light; (b) focal plane displaced 10 μm in direction of incoming rays. Refraction causes position of focal point to move about 5 times farther than sensor displacement.

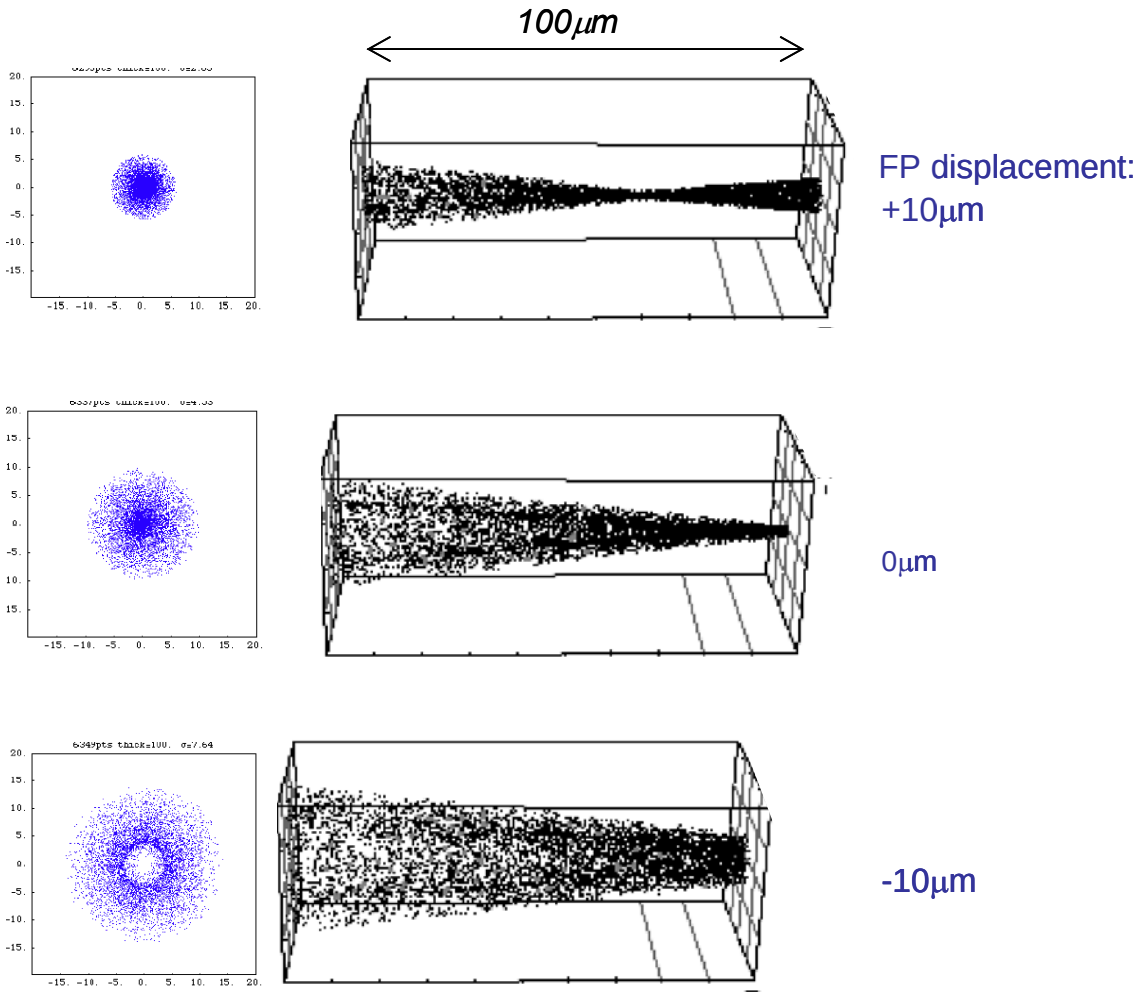
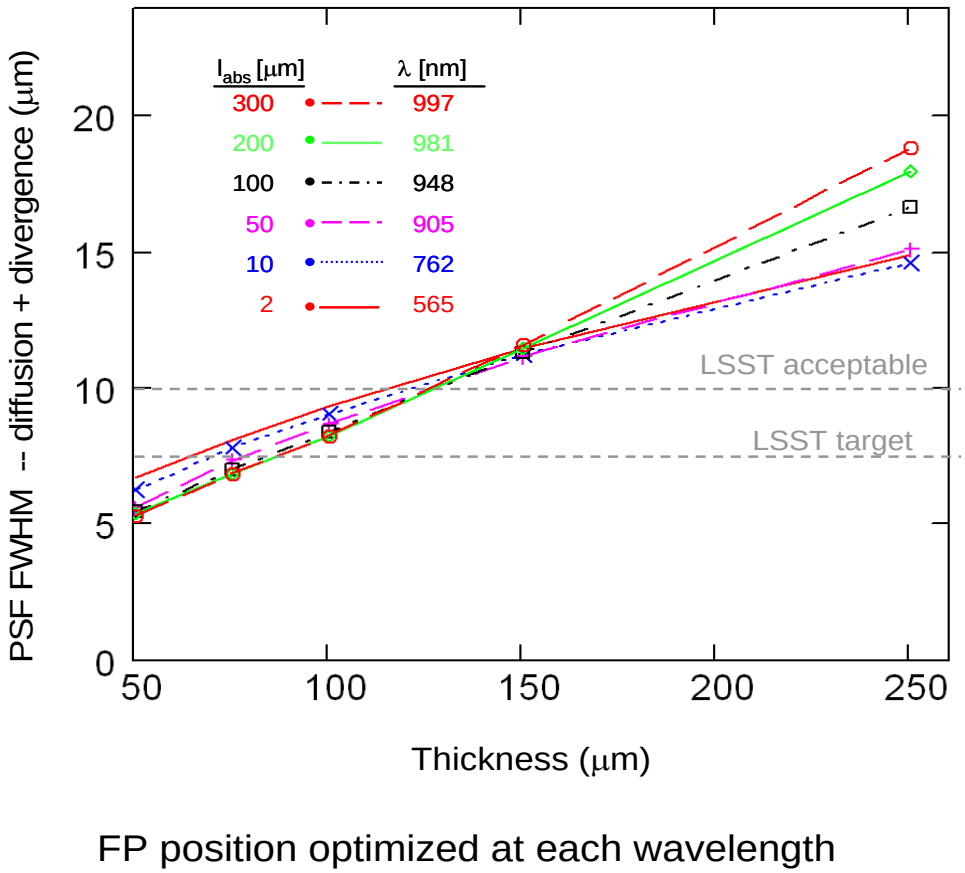


Figure 4.3.2-7 Absorption of long-wavelength light in a 100 μm-thick sensor. Focal plane position shown displaced by 0, and ±10 μm from best short-wavelength focus. Spot diagrams (with no diffusion) shown on the left.

As illustrated in Figure 4.3.2-7, for negative focal plane displacements the charge distribution is strongly non-Gaussian (the central obscuration in the telescope creates a PSF with a flat top); the FWHM values were calculated from the second moment of the charge distributions assuming Gaussian fitting. The combined effects of diffusion and divergent optical beam on PSF, and its dependence on the focal plane displacement are summarized in Figure 4.3.2-8 and Figure 4.3.2-9.



FP position optimized at each wavelength

Figure 4.3.2-8 PSF as a function of thickness. Effects of diffusion and beam divergence have been included. The focal plane position is varied at each wavelength until the best PSF is achieved. Sensor temperature 173 K; average electric field 2 kV/cm.

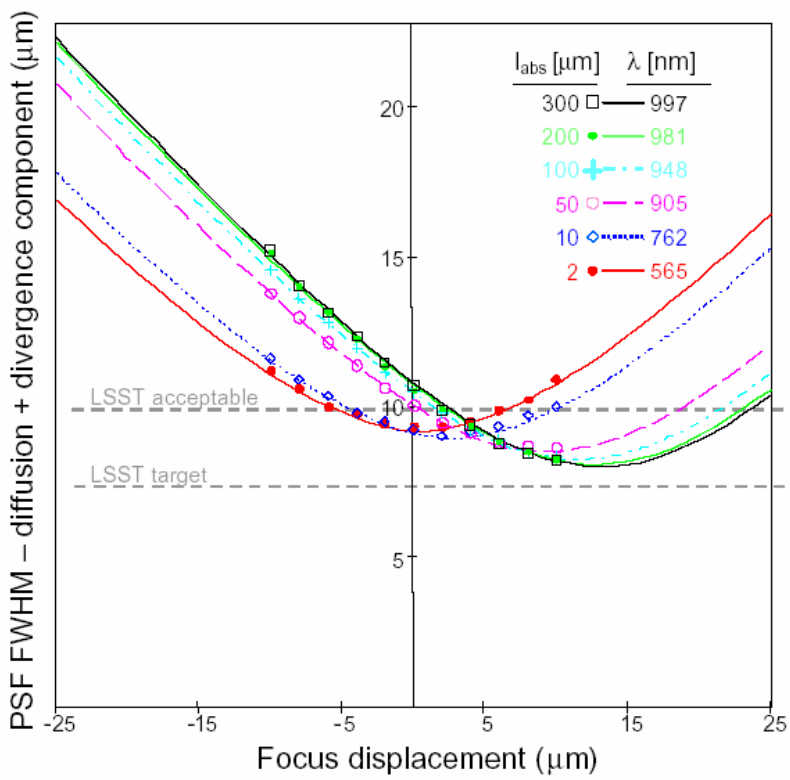


Figure 4.3.2-9 Point spread function dependence on focal plane position. Displacement of 0 μm corresponds to the point where light rays come to focus at the silicon surface. Positive displacements move the focal point into silicon volume. Sensor thickness = 100 μm; average electric field 2 kV/cm; operating temperature 173 K.

4.3.2.2.3 Detector Figure of Merit

The LSST has several science missions including detection of faint optical transients, galaxy cluster counting, and weak lensing by measurement of galaxy ellipticities. These goals require:

- detection of faint pointlike objects in multiple co-added exposures;
- accurate photometric redshift determination;
- stable and well-characterized point spread function;
- astrometric accuracy at the 0.01 pixel level;
- accumulation of large statistics data sets to minimize systematic error.

There is no single figure of merit that characterizes the detector's performance against all these science goals. The telescope's large $A\Omega$ product together with the fast detector readout speed give LSST wide sky coverage with multiple revisits to each field. Individual image quality can be characterized by the signal-to-noise ratio for point sources, or equivalently the limiting magnitude giving useful S/N. In 10-second exposures the background sky brightness will limit the S/N in LSST; under these conditions the limiting magnitude will be proportional to $\sqrt{QE / R_{sky}} / PSF$, where R_{sky} is the background sky flux. It is clear that an optimum thickness must exist, since both the QE and PSF increase for increasing thickness. The dependence of QE and sky brightness on wavelength will result in different optimum thickness for the different filter bands.

If we consider only the detector contribution we find [1] that for the g, r, i, and z filters (filters are described in Section 4.3.7) the optimum thickness is below 50 μm (assuming theoretically minimum diffusion, *i.e.*, a fully depleted detector) while Y band shows a broad optimum between

50 and 100 μm, depending on operating temperature, electric field, and the filter passband. In all cases the limiting magnitude depends only weakly on sensor thickness, varying by no more than 1/2 magnitude over the entire practical thickness range (~ 50 to 250 μm). When the contribution of atmospheric seeing to the PSF is included, the limiting magnitude becomes even less dependent on thickness and the optima shift to slightly higher thickness values (see Figure 4.3.2-10).

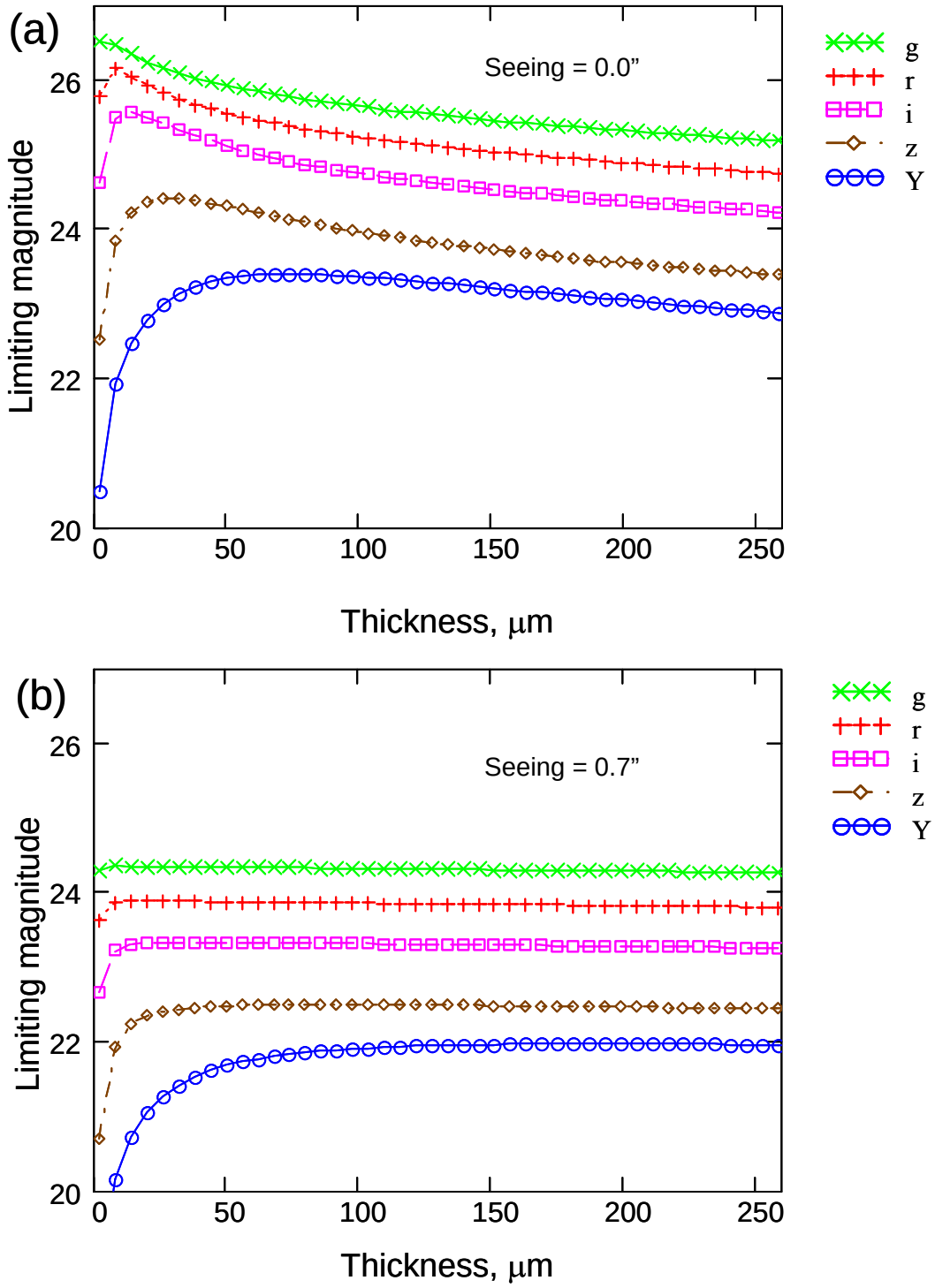


Figure 4.3.2-10 Limiting magnitude (S/N=10) for 10-second LSST exposures, for the filter set shown in Table 2. Sensor operating temperature 173K; substrate resistivity 10kΩ-cm, n-type; average electric field 2000 V/cm; focal plane position adjusted for best focus in each band. Reflection losses at sensor and atmospheric extinction not included. (a) Atmospheric seeing not included; (b) including seeing of 0.7" FWHM.

Somewhat different figures of merit are relevant for detecting extended objects and measuring their shape, and for determining redshift photometrically. The models we have developed for QE and PSF will be important in tradeoff studies during the R&D phase where end-to-end simulations of the telescope, atmosphere, and detector will be done.

Bearing in mind the weak dependence of S/N on thickness, engineering considerations such as packaging, cooling, and manufacturing yield may play a determining role in the ultimate thickness choice.

4.3.2.3 Sensor Format, Readout Speed, Readout Segmentation and Fabrication Considerations

The LSST focal plane has a number of requirements and desired characteristics that preclude the use of any existing sensor design. In addition to the need for extended red response and small point spread function (discussed in the previous sections), the fill factor, full well capacity, and readout speed and noise of the sensor have the most impact on LSST science performance. Due to the large (~3200 cm^2) imaging area of the focal plane, a large number of individual sensor units will need to be produced, so reliability, reproducibility, and compatibility with industrial fabrication methods are of paramount importance as well.

A large format CCD with highly segmented readout is the proposed approach to address these issues. Large format will minimize the number of gaps between sensors in the assembled mosaic. To provide for a 2 second readout of the entire focal plane, it will be necessary to have multiple outputs per CCD operating in parallel. As the number of outputs per CCD is increased, the readout speed per output can be reduced, thus minimizing the noise bandwidth. However, a correspondingly higher number of electronic signal processing channels will be required. Output-to-output crosstalk must be minimized, and reaching the required level is expected to be a challenge for the electronics development, discussed in Section 4.3.3.

A further advantage of segmentation is that it can be used to reduce the impact of bloomed charge from bright stars. In a 10-s LSST exposure the charge from 16th magnitude and brighter stars will exceed the pixel full well capacity, resulting in blooming up and down the column. By choosing an appropriate segmentation the length of the affected columns can be kept small, so that blooming from a saturated star is contained to within no more than .005% of the imaging area. Segmentation also substantially reduces the power dissipation of the clock drivers.

It is assumed that the LSST sensors will be developed and produced by a commercial vendor or vendors. However, using expertise that exists within the collaboration a “strawman” design of a large format, highly segmented CCD design has been produced and will be discussed in detail in the next section. The geometries of the frontside electrodes have been chosen to be consistent with today’s production CCD technology. It is anticipated that this strawman design can be used as a starting point for negotiations with potential CCD vendors leading to a contract for the production of the first prototypes (see Section 5.3.1).

4.3.2.4 CCD Strawman Design Outline

4.3.2.4.1 Overview

The overall features of the strawman sensor design are as listed in Table 4.3.2-4:

Table 4.3.2-4 Sensor features

Parameter	***Value***
Pixel size	*10 μm*
Format	*4000 × 4000 pixels*
Segmentation	*Eight 4000 × 500 pixel sub-arrays, 4 outputs each*
Total no. of output amplifiers	*32*
Anticipated gain	*3 – 5 μV/e⁻*
Parallel clocking	*4-phase (4 poly layers)*
Serial clocking	*3-phase*
Contiguous column length:	*500 pixels (100 arcsec)*
Guard ring	*100 μm*
Pin count	*208*
Fill factor	*96.5%*

We chose the 4k x 4k format to be the largest footprint consistent with good yield. Each amplifier will read out 500,000 pixels (one-quarter of a 4000 × 500 sub-array), allowing a pixel readout rate of 250 kHz per amplifier. The gaps between sub-arrays are approximately 100 μm. In image space, therefore, the sub-arrays are 13.3 arcmin long by 1.67 arcmin wide, with gaps of 2 arcsec separating them in the long dimension.

The strawman CCD architecture is illustrated in Figure 4.3.2-11. All pins are located along the two side edges. Shared serial and parallel metal busses connect the respective clock phase lines together.

The design was accomplished using a single metal layer, in order to demonstrate compatibility with as broad a range of fabrication facilities as possible. There are a few benefits to be had in reduced pinout and reduced drive impedance by going to a 2-metal process. Also, there are minor improvements in pinout and fill factor by changing to a 2-phase serial structure if this can be accomplished without greatly increased risk of incomplete charge transfer.

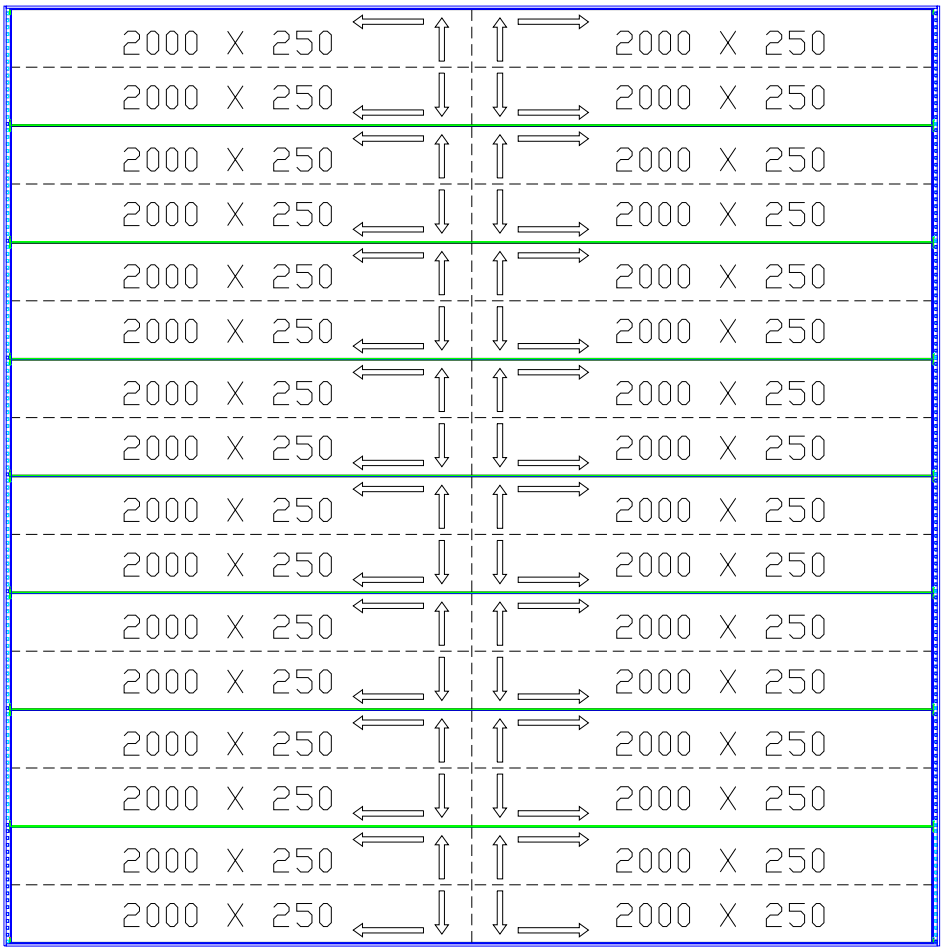


Figure 4.3.2-11 Outline of the 16 Mpixel strawman CCD, showing the partitioning and charge movement for the hardwired split parallel and serial registers. All pinout for the device is along the left and right edges. The fill factor achieved in this design iteration is 96.5%.

4.3.2.4.2 Output Structures

A single-stage FET output stage of 30 × 5 μm was chosen because of its compactness and minimal pinout. It will be the case for LSST that the external preamplifier will be located physically close to the back of the packaged device, with minimal video line impedance, so it is anticipated that a dual-stage output structure will not be necessary. The desired first stage gain is in the range of 3–5 μV/e$^-$.

The serial registers are bent 90 degrees at the end of each imaging area (see Figure 4.3.2-12), in order to move the adjacent outputs further apart (in this case, about 320 μm between sense nodes). This is to minimize crosstalk between outputs, as we anticipate this will be an issue for fully depleted high-resistivity substrates. The actual spacing for acceptable crosstalk (90 db or better) should be considered to be an open issue, subject to modeling and experimental verification.

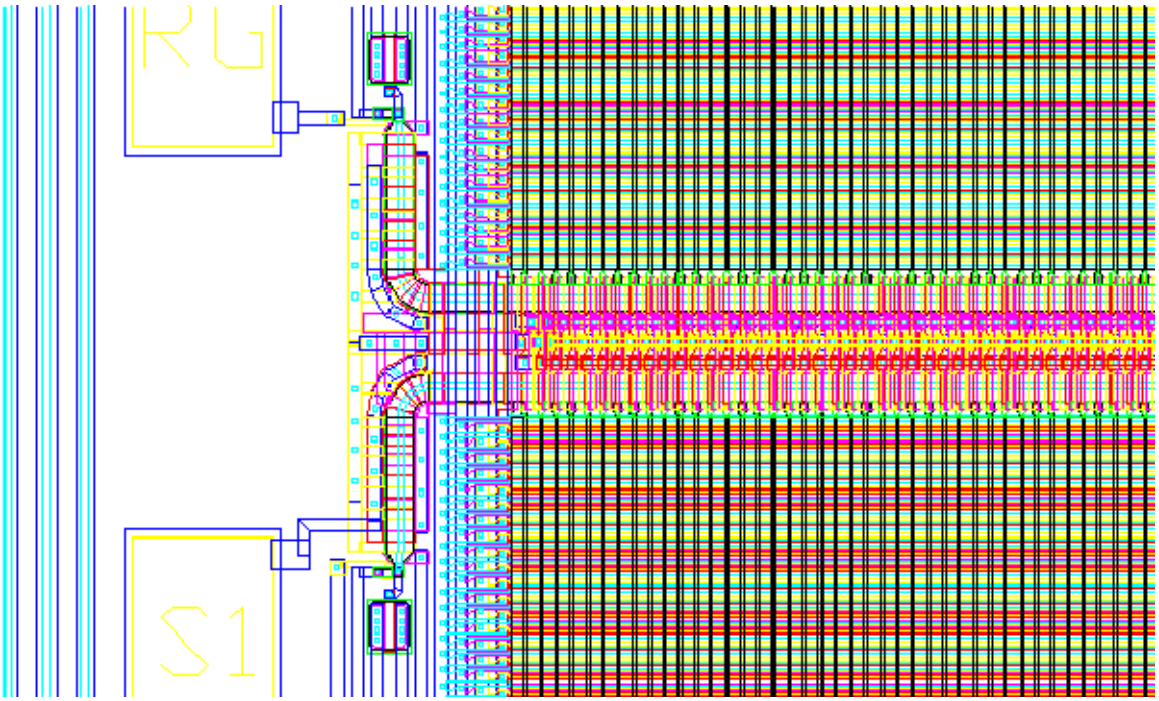


Figure 4.3.2-12 Serial registers for two of the interior imaging areas, with a 90 degree bend introduced in order to provide more separation between output FETs.

4.3.2.4.3 Guard ring and backside bias contact

The backside-illuminated high-resistivity substrate will require a backside voltage in the range of 30–60 volts be applied (see discussion of thickness, QE, and PSF issues above). We have provided an outer contact ring and implant to accommodate this bias voltage on the wiring side of the imager, with contact to the backside via edge conduction through the undepleted bulk silicon around the periphery. The details of workable structures to achieve this bias contact should be considered another open design issue, subject to modeling and experimentation, as are the details of the guard ring structure necessary to drop this voltage safely to the inner CCD substrate contact. A compact solution, such as the one drawn in the strawman design, is desirable, as growth in the size of these structures cuts into fill factor. However, a significant increase in the area for guard rings will have a relatively small effect on the fill factor (e.g., a 100 μm increase in the guard ring width represents only 1% in the fill factor).

4.3.2.5 Signal Processing and Guidelines for CCD Readout Design

Some essential features of the readout are highlighted here, and a more complete description is given in Section 4.3.3. Charge packets in the CCD are brought to the readout node by the serial shift register, and the charge-to-voltage conversion gain is determined by the node capacitance and the FET source follower gain. A higher conversion gain presents a larger voltage signal on the interconnections and in comparison to the noise in subsequent stages of amplification. However, the conversion gain must be limited to less than 5 μV per electron in order to handle the large full well charge (10^5 e^-) with acceptable linearity. The maximum readout noise requirement of 5 e^- rms (required to ensure sky-noise-limited performance) then corresponds to 25 μV rms.

In order to make the noise of the subsequent amplifier stage (external to the CCD) negligible, its noise will have to be less than ~ 1 e^- rms, or < 5 μV, over the bandwidth of interest. With a

bandwidth of ~2 MHz, corresponding to the highest anticipated readout speed, this translates into a spectral density of ~3.5 nV/√Hz . While there is some margin in this estimate, this requires a careful design of the signal processing chain. The number of CCD readout ports required to achieve the readout speed is large enough (32 for each of the ~200 CCDs in the focal plane) that monolithic circuits (application-specific integrated circuits, or ASICs) are essential for signal processing in this project. Their small size allows them to be placed close to CCDs (within a few centimeters). This makes a low interconnection capacitance possible, which results in reduced power dissipation in the CCD source followers and the crosstalk among signal channels. Analog noise filtering functions, correlated double sampling and dual-slope integration are best included in the readout ASIC.

4.3.2.6 Hybrid Silicon PIN-CMOS

The other candidate sensor type is a hybrid device consisting of a thick, high resistivity silicon photodiode array bump-bonded to a CMOS readout ASIC, referred to as a "Si PIN-CMOS sensor". A hybrid PIN-CMOS sensor has a sandwich construction consisting of a pixilated photodetector layer bump-bonded to a CMOS multiplexer integrated circuit, as illustrated in Figure 4.3.2-13. In this sense it is similar to hybrid near-IR array sensors that have gained wide acceptance in the IR astronomical community. Separation of photon detection from readout facilitates independent optimization of the Si PIN detector array and of the CMOS readout electronics. Bump bonding between the two planes of contacts with a pixel size as required for LSST (~10 μm) presents a technological challenge (18 μm pitch has been achieved). The bias on the PIN array is in the 30–50 V range depending on the thickness (the considerations on the thickness, depletion and diffusion (see Sections 4.3.2.1, 4.3.2.2) apply equally as for CCDs). Virtually all the charge resulting from photon conversion is collected on pixel electrodes (implants), and thus this sensor approaches 100% fill factor over the sensitive area. An area has to be provided at the edges of the sensor for the guard rings required to sustain the bias (as discussed for the CCD strawman design, Section 4.3.2.4).

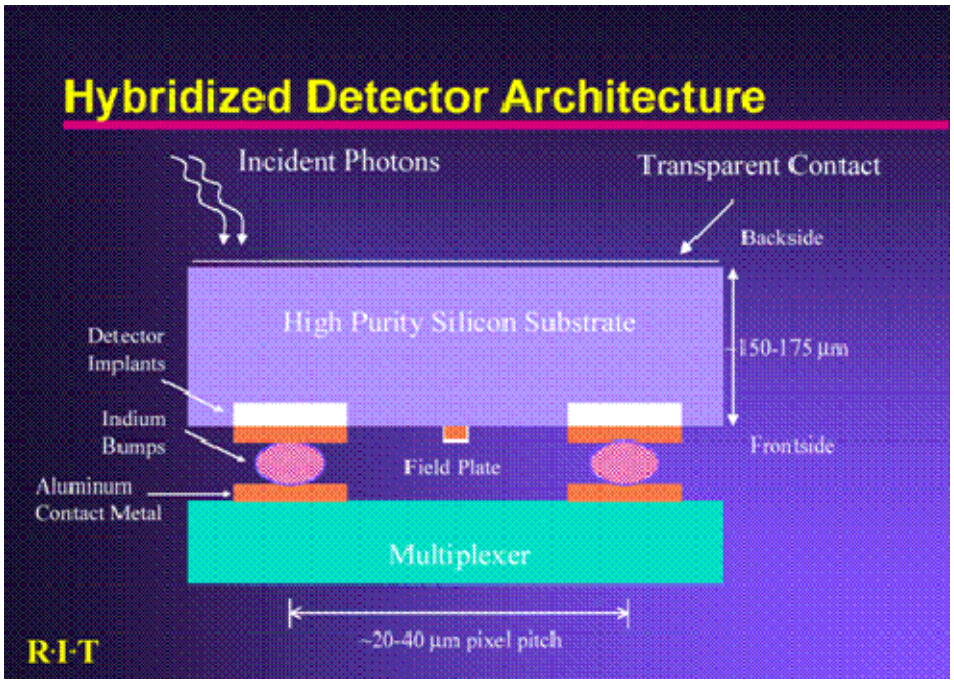


Figure 4.3.2-13 A simplified sketch of a hybrid Si PIN-CMOS senor. (Large pixel pitch, as used in IR sensors are indicated).

The CMOS readout contains a sense amplifier and a reset transistor in each pixel. The readout ("multiplexer") is arranged in a matrix of (row) select lines and (column) sense lines. In a normal readout mode pixels are addressed row by row. The columns provide information in parallel for a row of pixels, and this information is multiplexed to a number of outputs determined from the readout speed requirements. A variety of windowed readout modes may be obtained by utilizing flexible addressing controls in the multiplexer.

The reset transistors at each pixel (normally driven row by row), can also perform the function of an electronic shutter, eliminating the mechanical shutter needed for a CCD array. The resulting improved system reliability is a significant advantage of this type of sensor.

Evaluation of Si PIN-CMOS devices for scientific applications such as LSST is still at an early stage. Devices with 18 μm pixels, and with a smaller overall format, have been undergoing tests. Quantum efficiency was found to be as expected for the PIN layer thickness used, and the dark current was lower than required. The read noise for single CDS samples exceeds LSST requirements at present, and will be further studied. Devices with large formats and small (10 μm) pixel size are in active development by at least one manufacturer.

Note: Monolithic CMOS imagers, another CMOS-based technology, have not as yet achieved low-light performance suitable for astronomy. They have a very thin active region (a few μm), and a fill factor «100%.

4.3.2.7 CCD vs PIN-CMOS technology selection

Both the CCD and hybrid-CMOS technologies offer advantages and disadvantages. The most significant potential advantage of PIN-CMOS over CCDs would be operation of the camera without a mechanical shutter. However, silicon PIN-CMOS devices with the required area and the small pixel size have not been proven in astronomy applications. In addition, technical issues regarding the achievement of a reliable production process for large format and fine-pitched hybrid arrays, and the cost and limited number of potential suppliers also present special

challenges for LSST. At this stage, we are pursuing the development of CCDs as the baseline, with CMOS as the backup. A Vendor Information Package for the sensors has been prepared and was distributed to both CCD and CMOS vendors in January 2005. A request for proposals (RFP) to develop LSST sensors will be issued when the availability of R&D funds to support such development efforts is made clear. The technology downselect will be made when we have made an evaluation of responses to the RFP.

4.3.2.8 Testing Program

Testing will be concentrated on the most challenging parameters, which are the key objectives of the R&D program:

- direct measurement of PSF due to charge diffusion (e.g., ≤ 3 μm rms);
- properties of multiport readout, such as readout noise (≤ 5 rms e–) and crosstalk (< 10–3) at the required readout speed (2 s);
- QE vs wavelength (e.g., ~25–45% at 1000 nm);
- optical quality measurements of sensor flatness (< 5 μm peak-to-valley).

Optical metrology measurements of sensor flatness: Conventional phase measuring interferometry will be used to measure sensor flatness.

It is assumed that the sensors will be fabricated in an industrial process due to the large number required, e.g., ~250–300 including spares. As a part of the manufacturing process most of the basic functional testing is expected to be done by the manufacturer(s).

Testing in the R&D phase will be done at BNL, Harvard, CfA and STScI. Existing facilities at these institutions will require some additional equipment. The methods for critical parameter evaluation are independent of sensor technology. After the sensor technology is selected, some of the equipment will have to be tailored to the sensor type (CCD or PIN–CMOS).

In the R&D phase we will determine from the tests on successive prototypes how much additional testing and on how many units will have to be performed to prepare adequate testing stations for the production phase.

4.3.2.9 R&D Leading to Procurement

In summary, the necessary R&D to satisfy LSST requirements, starting from the well established CCD technology for astronomy-grade sensors, must address the following key ingredients:

- The effective pixel readout speed will have to be about two orders of magnitude higher than in previous telescopes in order to achieve a readout time for the telescope of ~1 – 2 seconds. This leads to ***a segmented CCD readout with multiple output ports operating at moderate clock frequencies***, as described in the "Strawman CCD Design" (Sections 4.3.2.3 and 4.3.2.4 above).
- The CCDs will have to have an ***active region ~100 μm thick*** to provide sufficiently high quantum efficiency at ~1000 nm, and they will have to be ***fully depleted (with no field free region)*** so that the signal charge is collected with minimum diffusion as needed to achieve a narrow point spread function.
- Packaging ensuring ***sensor flatness <~5μm*** peak-to-valley (not routinely achieved with presently delivered devices by industry).
- Extensive ***use of ASICs*** to make the readout of a large number of output ports practical, and to reduce the number of output links and penetrations of the cryostat.

This development program will lead to a fully qualified prototype, which will allow the procurement process to proceed. The development plan and its required resources is described in Section 5.3.1.

4.3.3 Camera Electronics

4.3.3.1 Introduction

For the purposes of this document, "Camera Electronics" is taken to refer to all electronic systems, interfaces, and functions which reside within, or are attached to, the outer camera body. Some may be commercial-off-the-shelf systems (COTS) while some will certainly be full custom and specific to LSST. In the case of COTS, electronic engineering resources may only consist of maintaining proper grounding, shielding, and connection systems, while the full custom systems, such as focal plane readout, will require the bulk of the LSST electronics effort. A general layout of the Camera Electronics systems architecture is shown in Figure 4.3.3-1.

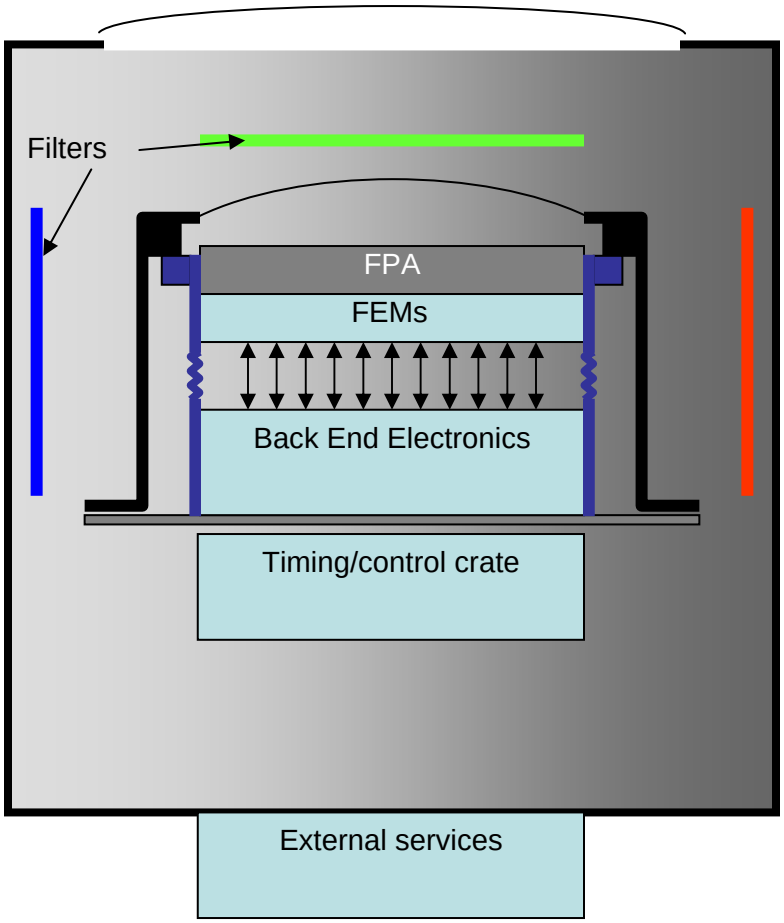


Figure 4.3.3-1 – Layout of camera electronics

As seen in Figure 4.3.3-1, the Camera Electronics divides itself logically into the following general areas.

- Systems residing inside the inner cryostat
- Systems residing mainly outside the cryostat, but within the outer camera body

- Systems directly outside of the outer camera.

A list of these systems and their likely location is shown in Table 4.3.3-1 below.

Table 4.3.3-1 Distribution of camera electronics functionality

System	Description	Location
Front End Modules (FEMs) Back End Electronics	Sensor signal processing Pixel digitization, data collection, optical fiber drivers	Inner cryostat
Timing/control crate	Image sensor timing/control Guide sensor readout electronics Wavefront sensor readout electronics Thermal monitoring & control Raft actuator controls Shutter monitoring & control	Outside the cryostat, in the Camera enclosure
External services	Power conversion / conditioning Vacuum system control & monitoring Camera mechanical actuator controls	Attached to the Camera enclosure

Note that Table 4.3.3-1 indicates a preliminary placement only. System placement may change as more information becomes available. The focal plane sensor readout electronics, residing within the inner cryostat, includes 1) the front end signal processing electronics and 2) the back end digitization and data collection electronics and optical fiber drivers. These are collectively referred to as Focal Plane Electronics which is described in the following section.

4.3.3.2 Focal Plane Electronics

4.3.3.2.1 General Discussion

The LSST Focal Plane represents a quantum leap in size and scope over those in use in telescopes today. While both CMOS/PIN diode image sensors and CCDs remain under consideration, we restrict our attention here to the CCD option described in Sections 4.3.2.3 and 4.3.2.4; for this design option, the readout electronics support 4K × 4K CCD sensors with 32 output ports per device. With the exception of the Front End Module, comprising the front end signal processing electronics, the remaining downstream electronics could accommodate a CMOS image sensor array as well.

The combined requirements of large size, high dynamic range, low noise, and rapid readout time, dictates a highly segmented focal plane with about 6,400 readout ports. This, in turn, dictates a high degree of integration for both the “front end” electronics, those which process the CCD output signals and provide clocks, as well as the “back end” which digitizes the data, buffers it, and sends it off-camera via optical fiber.

A primary choice must be made as to distribution of functionality either within the camera inner cryostat or outside it. A design with most of the electronics outside cryostat would afford greater accessibility at the expense of a much higher number of cryostat electrical penetrations. An analysis of connector requirements indicates that for a system with two hundred 16 Mpixel CCDs, about 20,000 cryostat penetrations would be required. To avoid this excessive number of penetrations, a much more highly integrated strategy with considerable electronics within the cryostat is preferred.

A second design driver is the requirement of low noise, and in particular, low feedthrough of digital activity back to the sensitive analog signals from the CCDs. This dictates front end electronics located very closely to the CCD ports with digital activity somewhat removed. It also dictates use of low-level differential signaling both for analog and digital data transmission.

Finally, the entire LSST focal plane will be synchronous in operation, which means that clocking for all sensors in the array will be synchronous to the level of some tens of nanoseconds. This assures repeatability and robustness against feedthrough and pickup. Timing generation will take place within the Timing/Control crate under command from the off-camera control system (see Section 4.3.8).

4.3.3.2.2 Specifications

Specifications for the LSST Focal Plane and Readout are shown below in Table 4.3.3-2.

Table 4.3.3-2 Focal plane array and readout specifications

Parameter	Value
FPA size	~3.5 gigapixels
Number of CCD sensors	~200
Pixel size	10 μm × 10 μm
CCD size	4k × 4k
CCD output ports	32 ports, 500,000 pixels each
Total no. output ports	~6,400
Full well capacity	100,000 e^-
Read noise	< 5 e^- rms
Dynamic range	16 bits
Readout time	2 s
Nominal exposure time	15 s

4.3.3.2.3 Readout Architecture

A block diagram showing the readout structure for each of the CCD sensors is shown in Figure 4.3.3-2.

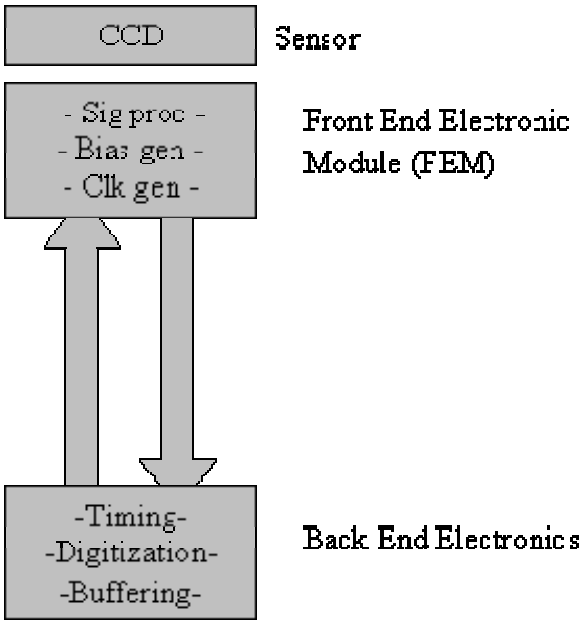


Figure 4.3.3-2 – Readout architecture for single sensor

Directly beneath each packaged CCD module is a Front End Electronics Module (FEM) containing the following functionality.

- Analog signal processing
- CCD clock drivers
- Bias generation

This architecture minimizes the physical distance from CCD output amplifiers to the analog signal processing circuits, thus minimizing power dissipation and risk of noise pickup. The CCD clock drivers are assumed to be simple level translators, with no pattern generation at this point. Beyond generating clock signals, there is no digital activity at this level. All analog signals (down arrows) are buffered and fully differential for purposes of noise immunity. Similarly, all timing signals (up arrows) comply with the Low Voltage Differential Signaling (LVDS) standard. Both the clock drivers and the bias generators will be programmable by means of a slow serial link. The "back end" electronics also resides in the cryostat and provides the functionality of digitization and frame buffering, as well as sending clock signals up to the FEM. Clock signals will not be generated at this level, but will be global signals distributed throughout the focal plane. This is done for reasons of simplicity and ease of synchronization.

Finally, the back end functions are collected onto "Data Cards" as shown in Figure 4.3.3-3.

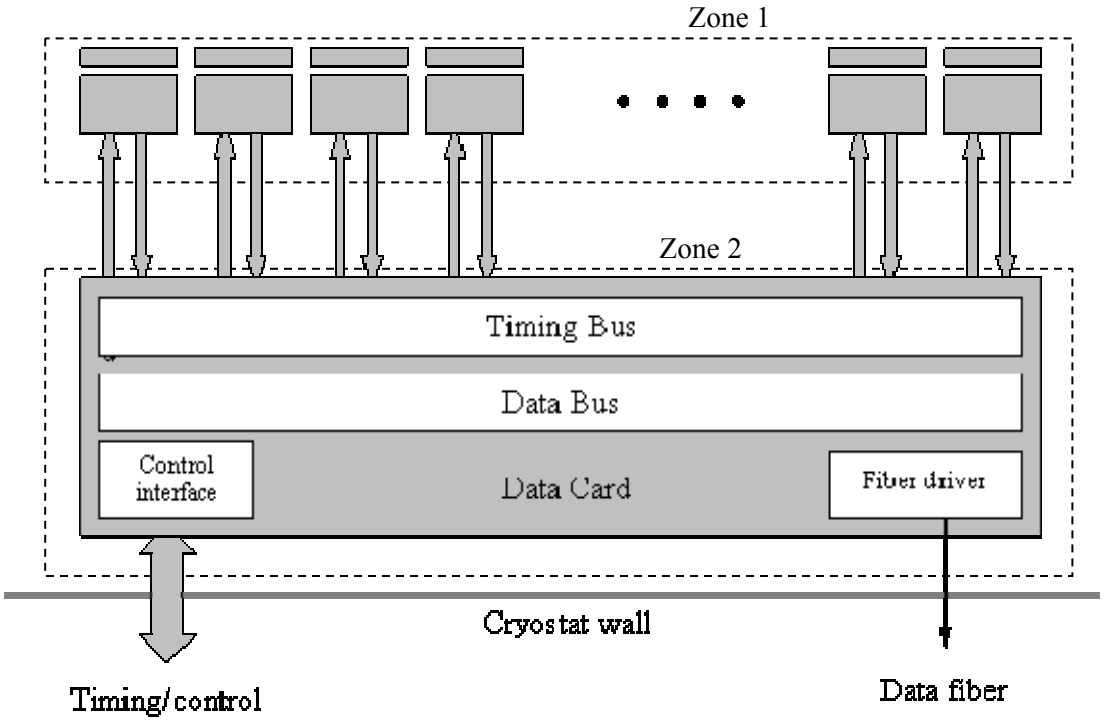


Figure 4.3.3-3 – Readout Architecture for groups of sensor modules. The designations Zone 1 and Zone 2 refer to thermal zones discussed in Section 4.3.3.2.7.

Each Data Card would service approximately 14 Front End Modules, so that 15 cards are needed to service the entire focal plane. Since each CCD delivers 16 Mpixel in 2 seconds, the peak data rate for each Data Card is 14×16/(2 seconds) = 112 Mpixel/s. Assuming each fiber frame is 20 bits, this translates to a modest fiber requirement of ~2.3 Gb/s. Two options exist for moving the serial pixel data through the cryostat wall: traditional electrical (copper) feedthroughs or fiber optic feedthroughs. The latter approach is preferred because no additional electronics would be required to convert to fiber outside of the cryostat.

4.3.3.2.4 ASIC development

In traditional approaches, a large portion of the focal plane readout real estate goes to analog signal processing of CCD output signals. The high channel count of LSST dictates that ASIC front end electronics replace the traditional discrete approach. The advent of low cost CMOS ASIC technology has revolutionized the field of experimental high energy physics over the past several decades. The current and next generations of high energy physics experiments would not be possible without extensive use of custom CMOS ASICs and that community has invested heavily in acquiring ASIC design expertise. Space-born and ground-based astronomy and astrophysics facilities are also benefiting from these developments and we intend to borrow heavily from these two communities for our ASIC development plans. For LSST, the major electronics development will be a multi-channel analog signal processing ASIC.

4.3.3.2.4.1 Analog Signal Processor ASIC

Standard techniques for analog signal processing of CCDs include “Clamp and Sample” and “Dual Slope Integration”. Both are forms of Dual Correlated Sampling and we have chosen the latter for its superior noise performance as well as its flexibility. We plan to develop a multi-

channel dual slope integrator of 32 (or 16) channels with noise performance of ~5 electrons rms at a read speed of 250 kHz. The baseline design is a single range dual slope integrator with external commercial (COTs) ADCs.

The ASIC development will likely be done in a 0.25 µm CMOS process with prototypes fabricated through MOSIS. Based on extensive experience, we anticipate the need for three prototype design/fab/test cycles prior to production. Each cycle will take approximately nine months, for a total development time of about two and a half years, after which, integration tests with CCD sensors will commence.

Preliminary specifications for the ASIC are shown in Table 4.3.3-3.

Table 4.3.3-3 Preliminary ASIC specifications.

Parameter	***Value***
Read noise	*~ 5 e^- rms*
Full well capacity	*~ 100,000 e^-*
Power dissipation	*~ 25 mW or less per channel*
Linearity	*~ 1% or better*
ASIC contribution to crosstalk	*~ 0.1%*
Operation at focal plane temperature	

4.3.3.2.4.2 CCD Clock Translator

In order to insure absolute synchrony over the entire focal plane, CCD clock timing signals will be generated from a single timing state machine located externally to the inner cryostat and distributed to the FEMs as LVDS logic level pulses. These will then be translated to the levels required to operated the CCDs. The required voltage levels will be programmable and be in the range of 10V or more, depending on the final choice of sensor. The translation of the logic levels to CCD clock levels will be done by the CCD Clock Translator.

We will investigate both discrete and ASIC implementations for the CCD Clock Translator. Required signal swing of 10V or more preclude the use of standard CMOS ASIC processes. Specialized “high voltage” processes, tolerant to 20V or higher swings, do exist however, and are commercially used in automotive and other applications. We will investigate use of these technologies for LSST. The baseline plan is an ASIC implementation as that will be most attractive from the point of view of required real estate and power. The alternate plan is to develop a high density hybrid Clock Translator. While both power and real estate will be larger for the hybrid, we still consider this an entirely viable approach. Since we anticipate the need for sensor testing well in advance of completion of a CMOS ASIC clock translator, we will develop the hybrid early in our program.

4.3.3.2.5 Back End Electronics: Data Cards

The Data Cards receive differential analog signals from the Analog Signal Processing ASICs and digitize them using commercial off-the-shelf (COT) 16-bit ADC chips. The latter exist from several vendors and have been in use by the astronomy community for many years. A challenge will be packing density, real estate, and cooling. A Data Card which services 14 CCD modules would, for example, require 112 dual channel ADCs. Use of direct wire bond-on-board and other high density packaging techniques will be investigated to accommodate these needs.

Required technologies for the Data Cards exist commercially and we foresee no need for ASIC developments. Single chip 256 Mb memories are available and will be used as frame buffers.

Commercial FPGAs will handle control and data management functions. And finally, commercial fiber optic components exist for our anticipated data rates.

4.3.3.2.5.1 Analog to Digital Conversion

The required dynamic range for the LSST Camera focal plane is 16 bits and we are currently considering several approaches. The first, and preferred, approach is the use of commercial off-the-shelf ADCs. A number of these are commercially available and appear to match our requirements. To achieve true 16-bit performance, the analog signals must be transmitted from the Front End Boards to the Data cards while maintaining 16-bit level of signal integrity, which is to say, linearity, noise, and crosstalk immunity. We intend to test the viability of this option in a systems test well in advance of any ASIC developments. Use of a single range 16 bit system is the preferred approach.

Alternatively, 16 bit dynamic range can be achieved with multiple range gain stages coupled with a lower resolution ADC. This is the approach used by SNAP [need ref] in the development of their CRIC chip with integral analog to digital converter. This approach will be investigated at some level if true single range 16-bit operation fails to deliver the required performance.

4.3.3.2.6 Outgassing

Both Front End and Back End electronics must address the issue of outgassing, as outgas products would condense on the focal plane surface causing severe optical degradation. Low outgas substrates and material coating techniques do exist and will be investigated for the LSST Camera environment. The outgas management strategy may also involve baffling, shielding, and thermally controlled elements.

4.3.3.2.7 Thermal management

The in-cryostat electronics divides itself conveniently into two thermal zones. Zone 1 comprises the front end modules which reside just below the sensors and therefore will be only some tens of degrees warmer. If the focal plane is at −100 °C, for example, the Front End Modules might be at −80 °C. This zone will likely be closely coupled to the Focal Plane thermal circuit to maintain a fixed temperature differential. The Back End Electronics (Data Boards) comprise the second thermal zone. As these boards will be mounted in close proximity to the back flange of the cryostat, they will be considerably warmer, perhaps in the neighborhood of 0 °C to −20 °C. A separate cooling system will be designed for this zone, realizing that the temperature stability requirements will not be nearly as strict as those of zone 1 or of the Focal Plane itself. The collection of flex-cables between the Front and Back End boards will thus drop the bulk of the temperature difference between the two zones and will be designed and specified accordingly.

4.3.3.2.8 Test Stand

A multi-purpose Test Stand will be required early on in the Camera development program. The main functions of this test stand are test and evaluation of:

- Signal Processor ASIC prototypes
- Clock Translator circuits (ASIC and Hybrid)
- Sensor prototypes
- Commercial off-the-shelf (COTs) components
- All performance parameters vs temperature
- Outgassing tests and development of remediation
- Readout strategies and critical performance parameters.

- Prototype CCDs

We have developed a Test Stand architecture which closely mimics the architecture of the final readout. Thus, design and development of the Test Stand serves the function of prototyping all aspects of the final Camera readout. A block diagram of the Test Stand is shown below.

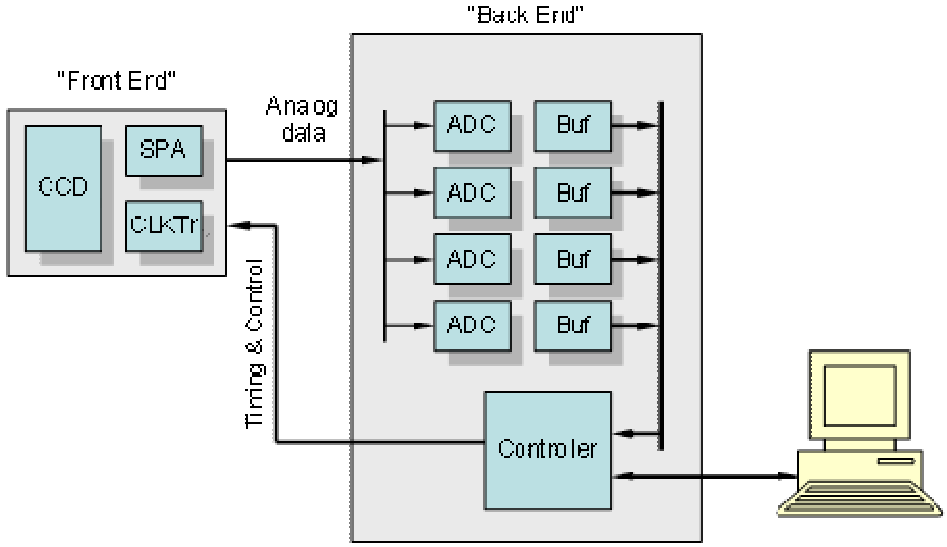


Figure 4.3.3-4 – Test stand architecture

The "Front End" mimics the functionality of the Front End Modules and can be used to test and evaluate any of the front end devices either singly or in combination. Likewise, the "Back End" is used to mimic the Data Cards' functionality. We expect both of these major blocks to evolve as the designs mature.

In operation, analog data from the front end test card is digitized by ADCs on the back end and the corresponding digital data are stored in frame buffers to be read out after digitization into the host PC. The particular communication link between PC and Back End Test Board is not critical, but is likely to be high speed USB.

A fully instrumented cryostat will be set up for each copy of the Test Stand. Once room temperature operation has been established for each component, all critical operating parameters will be characterized at operating temperatures of approximately −100 °C.

A major goal of the Test Stand activity will be to implement a "vertical slice" systems integration test of CCDs with all prototyped downstream electronics. This full readout chain could then be evaluated in the focal plane of an existing telescope.

4.3.4 Packaging and Focal Plane Assembly

4.3.4.1 Introduction

Comprising the focal plane subassembly are: 1) sensors precision mounted to a supporting plate to create a testable 3×3 "raft;" 2) tested 3×3 rafts of sensors precision mounted in a 5×5 array to 3) an integrating structure (together forming the focal plane array or FPA); 4) an image-

stabilization mechanism; and 5) elements of the thermal management system and electronics that install simultaneously with the focal plane array into the camera's inner cryostat.

4.3.4.2 Requirements and constraints

1. The active area of the focal plane array must be sufficient in size to cover the 3.5˚ field of view and 10.5 m focal length of the telescope. These specifications lead to an area requirement of 0.64 m diameter.
2. The sensing surfaces of the assembled focal plane array must be flat to 10 µm peak-to-valley. This flatness specification applies for any orientation of the telescope during observations.
3. The focal plane and the camera optics (L1, L2 and L3) must be precisely located with respect to one another to tolerances determined using an optical design code. The focal plane subassembly must have mechanical or optical datums so that these tolerances may be satisfied at assembly.
4. The sensors must operate at an absolute temperature of −100 °C (see Section 4.3.6). The allowable spatial variation in temperature is 0.1 °C within one sensor and 1 °C across the full array. The allowable temporal variation in temperature subsequent to calibration is 1 °C.
5. The sensors must be protected from mechanical and electrical damage, and from contamination from sources such as airborne dust or volatile condensable chemicals, at all times during assembly, operation and service.
6. Controlled planar motion of the FPA for image stabilization must have sufficient range and bandwidth to track image motion. The actual requirements may not be well known until detailed dynamic simulations can be done or perhaps even until operation, but closed-loop bandwidth on the order of 100 Hz should be achievable (based on previous experience) and sufficient to track image motion estimated to be of order 2 arcsec on the sky or 100 µm at the focal plane.

4.3.4.3 Focal Plane Components and Layout

4.3.4.3.1 CCD Packages

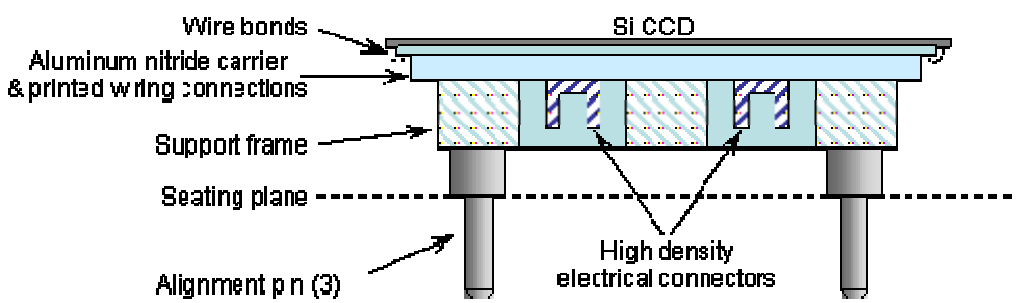


Figure 4.3.4-1 Proposed sensor package showing a CCD on aluminum nitride carrier with electrical and mechanical interfaces.

Each sensor chip will be attached to a support structure consisting of materials with very low and matched temperature expansion coefficients, such as aluminum nitride and invar. Such a package is shown schematically in Figure 4.3.4-1 and Figure 4.3.4-2. It consists of a 4-side buttable silicon detector glued to an aluminum nitride substrate. Bond wires then make the electrical connection

between the photon-sensitive silicon and metal pads on the aluminum nitride. The metal lines route electrical connections to two connectors.

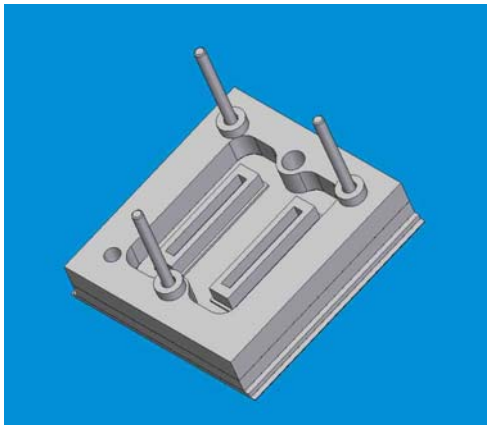


Figure 4.3.4-2 Backside of CCD package

Raft Structure

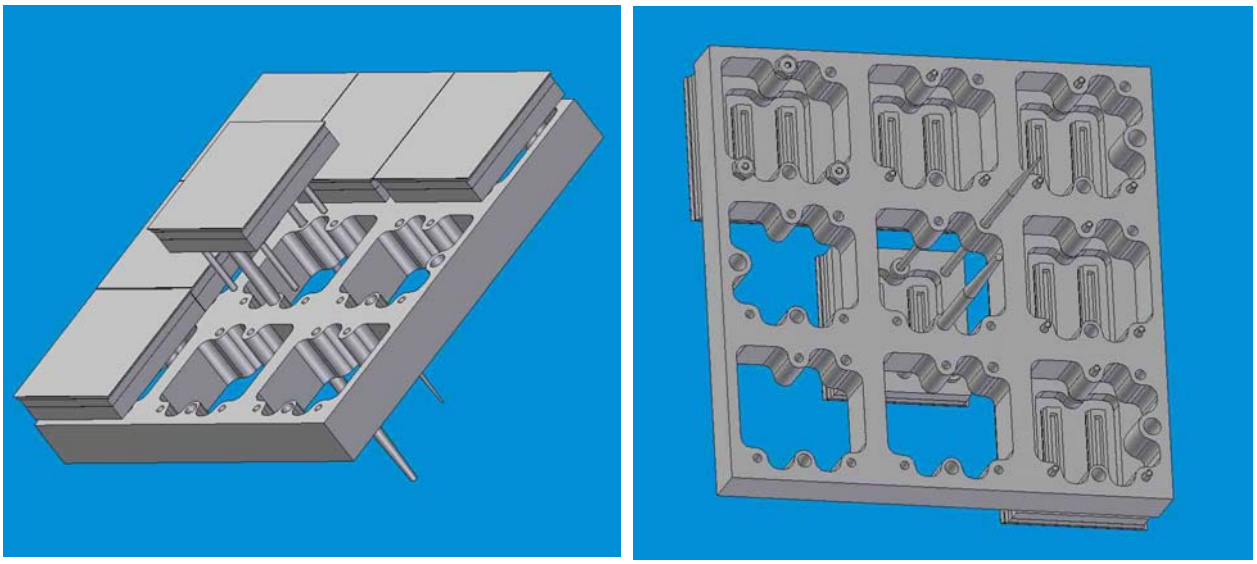


Figure 4.3.4-3 Sensor packages assembled into “raft” structure.

The camera focal plane will be assembled by mounting sensors onto “rafts,” Figure 4.3.4-3, and then mounting the rafts onto an integrating structure as illustrated in Figure 4.3.4-4. Individual sensor packages will be enhanced with temporary guide pins to aid assembly into rafts.

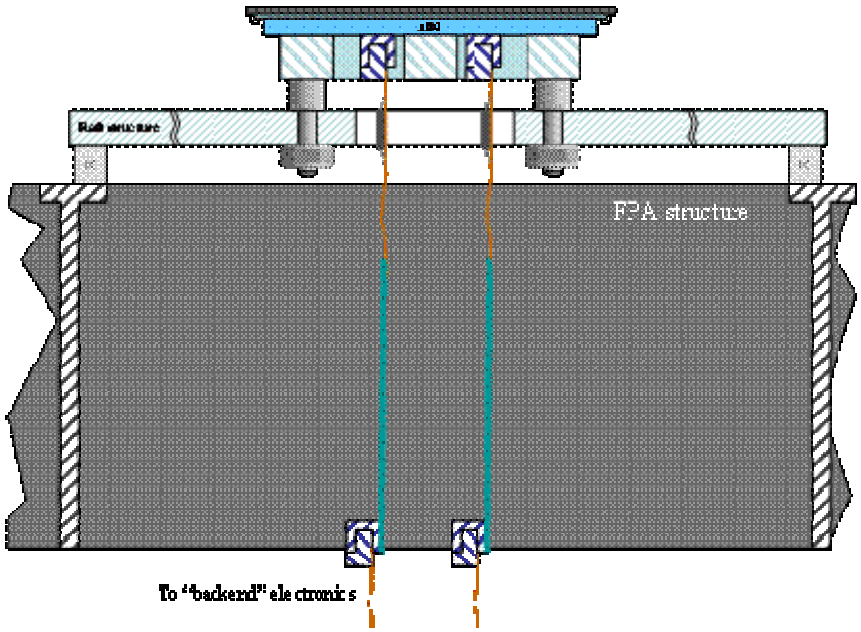


Figure 4.3.4-4 Cross-section through focal plane showing one packaged CCD on raft assembled to FPA integrating structure. Kinematic mounts at the raft-FPA integrating structure interface may include adjustments for raft alignment.

4.3.4.3.3 Focal Plane Layout

A layout of the 3.5° focal plane, using the raft concept described above, is illustrated in Figure 4.3.4-5. The rafts are offset by 1 cm to provide locations for wavefront sensors (see Section XXXinsert Xref to 3.1.6).

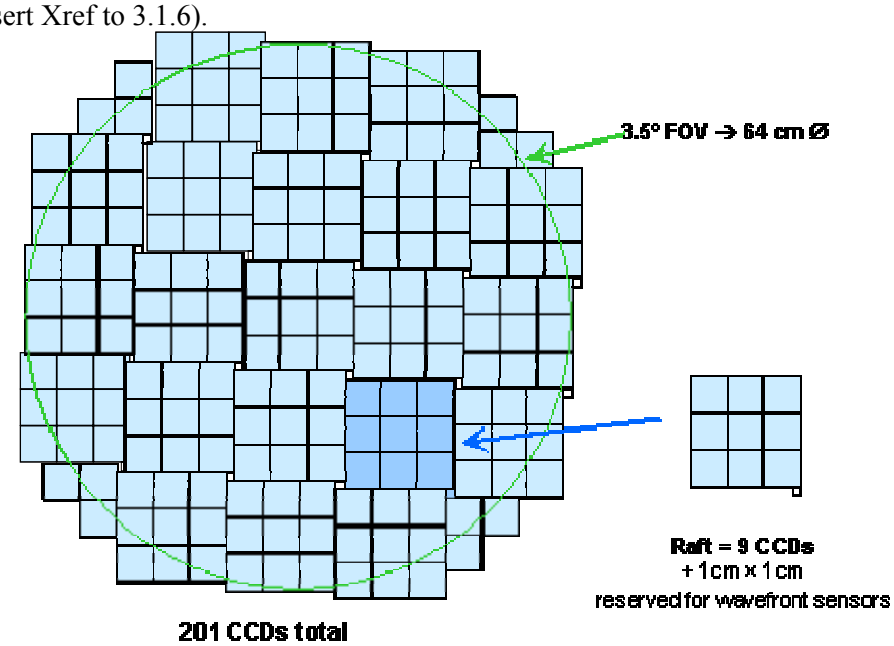


Figure 4.3.4-5 An example of a possible arrangement of 3×3 sensor rafts in the focal plane with areas provided for wavefront sensors by raft offset.

4.3.4.3.4 Integrating structure

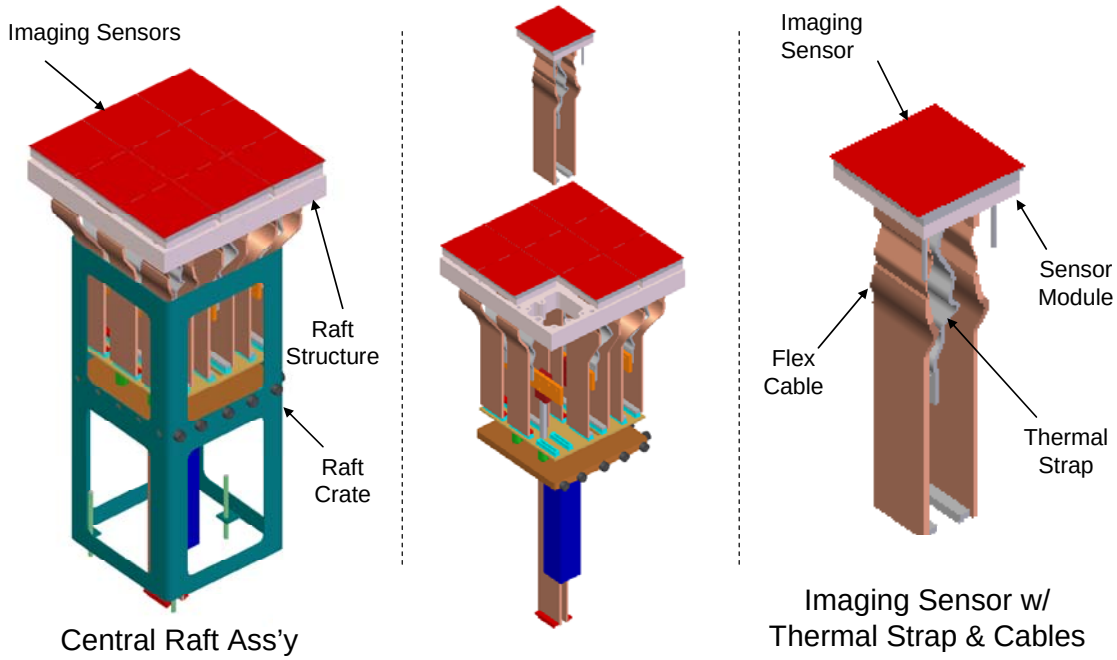


Figure 4.3.4-6 Raft module with integrated front-end electronics and thermal connections.[need to add label for front end elevtronics]

The rafts are assembled and tested with front-end electronics, as illustrated in Figure 4.3.4-6. These raft assemblies are then installed in an integrating structure as shown in Figure 4.3.4-7. The physical space between sensors is minimized to maximize sensitive area ("fill factor"), yet contact must be avoided to avoid permanent damage. If the guide pin technique used to assemble individual sensor packages into rafts is inadequate for installing rafts into the integrating structure, an assembly fixture will be designed and built and a procedure developed to facilitate routine insertion. This fixture would be moved to the telescope site so any faulty sensors could be replaced with spares.

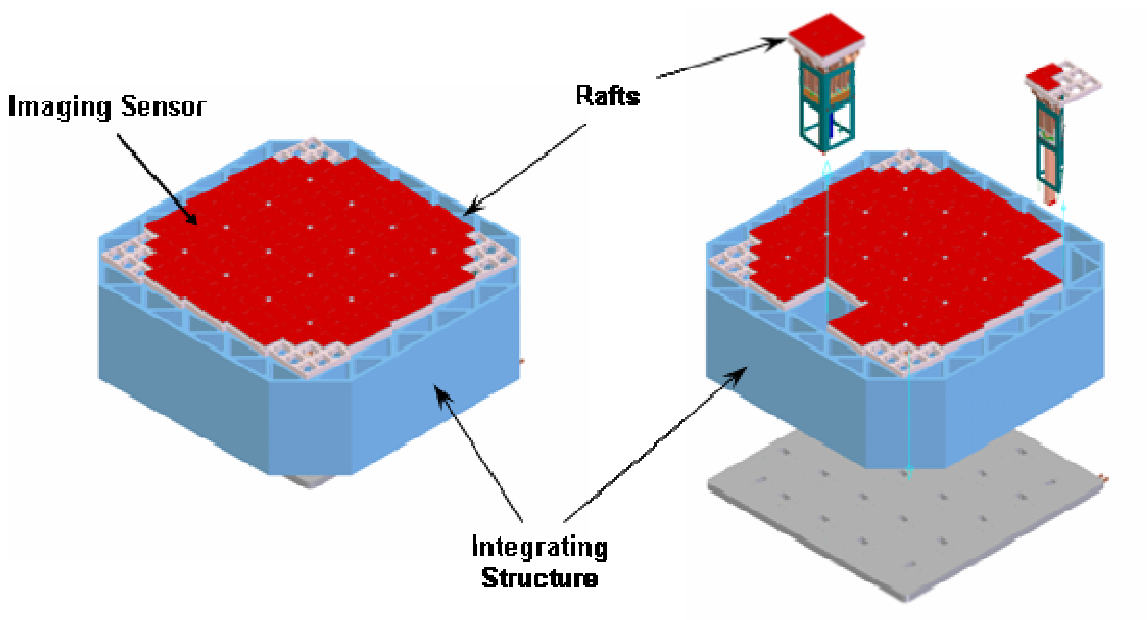


Figure 4.3.4-7 Focal plane assembly concept

4.3.4.4 Focal Plane Flatness

Achieving 10 μm peak-to-valley flatness over the whole focal plane array is challenging because there are several error contributors needing to be of order 1 μm or less. The error budget, which apportions flatness and size tolerances among components (e.g., individual sensors, rafts and integrating structure) and changes that occur in interfaces and components with time, temperature and orientation, is coupled with the approach for building the array. For example, all components can be manufactured with sufficient precision to assemble in snap-together fashion without needing adjustments, but this may be more expensive than providing relatively few simple adjustments to relax difficult tolerances. The final choice will depend on analysis and testing results and on interactions with sensor vendors, for example. A likely approach is presented next along with merits and difficulties. The error budget associated with this approach is illustrated in Figure 4.3.4-8

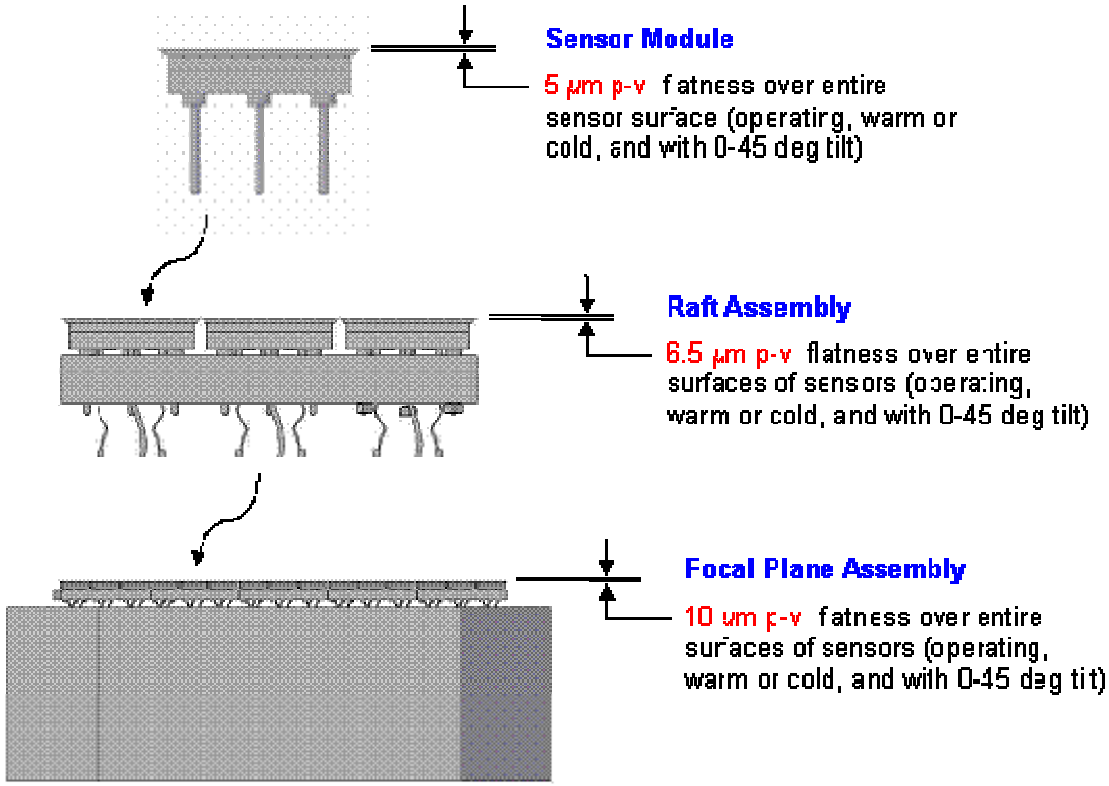


Figure 4.3.4-8 Allocation of focal plane flatness specification. Of the total camera allocation of 10 μm peak-to-valley (p-v),.half is reserved for the sensor module fabrication, and the remainder allocated to the combination of fabrication, thermal and mechanical load errors in the raft substrates and focal plane integrating structure, and assembly tolerances.

There are 201 sensors in the focal plane array, each with three mounting pads. This is a sufficient number to favor developing an industrial process with sufficient precision to avoid custom manufacture or adjustments. This preference is reflected in the error budget, which allocates 5 μm peak-to-valley height tolerance to the sensing surface with the sensor mounted to a perfect plane. Therefore, individual sensors must be sufficiently flat, parallel and at the same height after mounting to achieve 5 μm peak-to-valley. The process of replication, already used in current devices to join the silicon sensor to the aluminum nitride substrate, can achieve suitable precision at low per-unit cost. The sensor vendor will be responsible for controlling the flatness and height of the sensitive surface with respect to the three mounting pads.

Arrays of 3×3 sensors are mounted to raft structures that in turn mount to the integrating structure. Raft structure plates can be double-side lapped all at once to be flat, parallel and the same size to less than 1 μm. Thus, rafts assembled with sensors should be flat, parallel and the same height to 6.5 μm peak-to-valley. In the present scheme, it will be important for the raft structure and the sensor substrate to have a well-matched coefficient of thermal expansion and good thermal conduction to maintain this flatness performance during operation, since the assembly is subject to large temperature changes (ambient to operating temperature).

A 5×5 array of rafts mount to the integrating structure in a similar manner as the sensors with three coplanar pads, giving 75 pads total. These mounting pads incorporate flexural freedom to avoid over-constraint of thermal expansion, which could alter the flatness of rafts and the integrating structure and be potentially damaging. The integrating structure can be manufactured flat to 1 μm at room temperature and it can be made stiff enough to deflect less than 1 μm due to

the changing orientation of the telescope. Non uniform thermal expansion and change in heat load between room temperature and operating temperature could cause distortions in the structure. If the sum of all error terms (including the raft allowance) exceeds the 10 μm peak-to-valley budget, a means of adjustment could be incorporated into the design. The adjustment mechanisms could be realized with remotely driven actuators, manually driven screws, fitting spacers, or component modification (such as scraping or lapping). These adjustments may be a permanent part of the assembly or they may be part of tooling used temporarily.

A way to inspect flatness of the focal plane array will be necessary whether adjustments are made or not. Non contact optical techniques are preferred to allow inspection through a window (with proper calibration) at operating temperature and also at room temperature.

The stability of the integrating structure is of critical importance to achieving and maintaining the flatness of the focal plane array through all operating conditions. There are three main considerations affecting stability: material choice, the structure's shape and size (i.e., geometry), and connections to other entities, namely the rafts and the cryostat. Looking first at material properties, the main drivers are high elastic modulus E, high thermal conductivity k, low density ρ and low coefficient of thermal expansion α. Collectively these give minimum deformation under inertial and heat loads. Assuming they all have the same importance, the material property group $E/\rho * k/\alpha$ may be used to choose the best candidate materials. Among the best are silicon carbide and beryllium (I-70A). The properties of silicon carbide vary among types and manufacturers. In particular CoorsTek SC-30 direct sintered silicon carbide rates substantially (3.77x) higher than beryllium. Working down the list to more conventional materials, SC-30 is 48x higher than aluminum, 57x higher than Invar-36 and 79x higher than steel, using the material property group $E/\rho * k/\alpha$ as a figure of merit. If analysis shows that E and α should be given more weight in this figure of merit, SC-30 has an even greater advantage and Invar-36 becomes more desirable than aluminum and steel. The uniformity of E, α and temporal instability will also be very important. That is to say, the material could grow uniformly with pressure, temperature and time and not cause degradation of flatness. One would expect a ceramic to be more stable and more uniform than a metal but we do not yet know whether there is a significant difference. SC-30 is used increasingly in preference to Beryllium for mirror substrates where instability would surely be noticed. The main reason to carry Invar-36 as an option at this point is primarily economic; it can be conventionally machined by any number of vendors, although with difficulty. Silicon carbide will require close interaction with the probable vendor during the design process. Both materials are good CTE matches to silicon and aluminum nitride.

Figure 4.3.4-9 shows a first-cut finite element analysis of a possible structure design. It is similar in design to a structure for a previous project that was machined from a solid billet of beryllium. The outer dimensions are 771.4 mm square x 200 mm deep and there is space allocation for at least 300 mm depth. Although the results shown are for steel, Table 4.3.4-1 also gives results for SC-30 and Invar-36 scaled from material properties. This analysis did not include the mass of rafts of sensors but the last row of the table gives a mass-scaled estimate for the Z gravity sag across the Ø 640 mm image field assuming a payload of 40 kg. Invar-36 at 0.70 μm is challenging to the error budget but is still workable, whereas SC-30 at 0.12 μm is practically negligible. There are several ways to improve the structure's geometry to reduce the number further, perhaps by a factor of two.

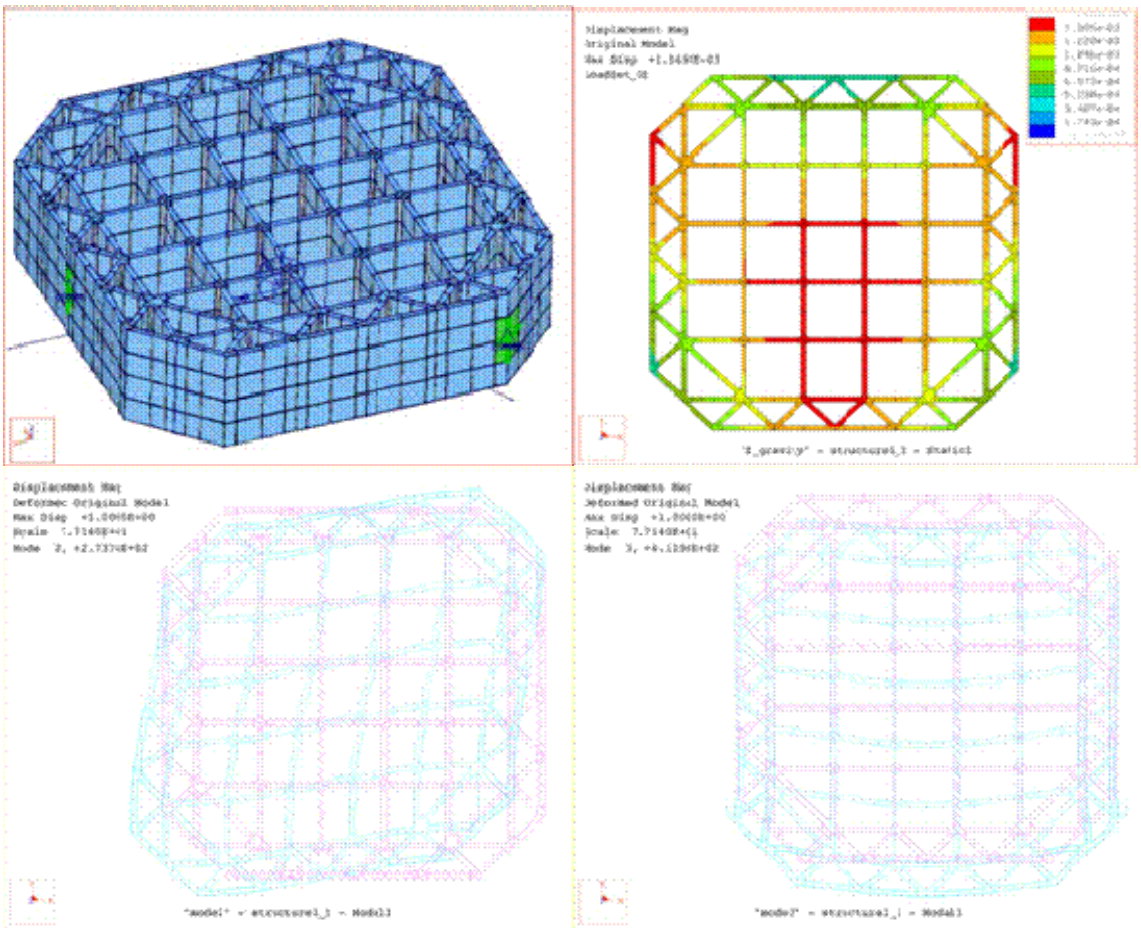


Figure 4.3.4-9 The finite element model (upper left) is constrained at three points, each with tangential and axial displacement constraints about a cylindrical coordinate system (i.e., kinematic constraint). The color fringe plot (upper right) shows displacement of the structure due to gravity along the camera axis (Z). Deviation from flatness (or sag) is derived from this displacement information. The two (lower) animation plots show mode shapes that the image stabilization system would tend to excite. These modal frequencies should be well above the desired system bandwidth for robust control.

Table 4.3.4-1 Key results from the finite element analysis using steel and scaled results for SC-30 and Invar-36.

Analysis result	Steel	SC-30	Invar-36
Mass (structure w/o payload)	210 kg	86.2 kg	216 kg
Z gravity sag over Ø 640 mm	0.40 μm	0.083 μm	0.59 μm
Torsion mode 1	227 Hz	497 Hz	187 Hz
X translation mode 2	274 Hz	600 Hz	226 Hz
Y translation mode 3	413 Hz	905 Hz	341 Hz
Mass (structure + 40 kg payload)	250 kg	126 kg	256 kg
Estimated Z gravity sag (Ø 640 mm)	0.48 μm	0.12 μm	0.70 μm

As previously discussed, the connections between the integrating structure and rafts provide flexural freedom to accommodate differences in thermal expansion. Otherwise each adversely affects the stability of the other. The same consideration must be given to the interface between

the integrating structure and the cryostat, which should be kinematic, that is, not over constrained. This is also a logical place to incorporate the image-stabilization mechanism.

4.3.4.5 Image-stabilization mechanism

Figure 4.3.4-10 shows a system of blade flexures that allows three planar degrees of freedom for image stabilization and provides three constraints against out-of-plane motion. This exact constraint system allows freedom for differences in thermal expansion between the housing and structure and sight misalignment of mounting surfaces. The flexures also serve as thermal resistance between the cold structure and the warm housing.

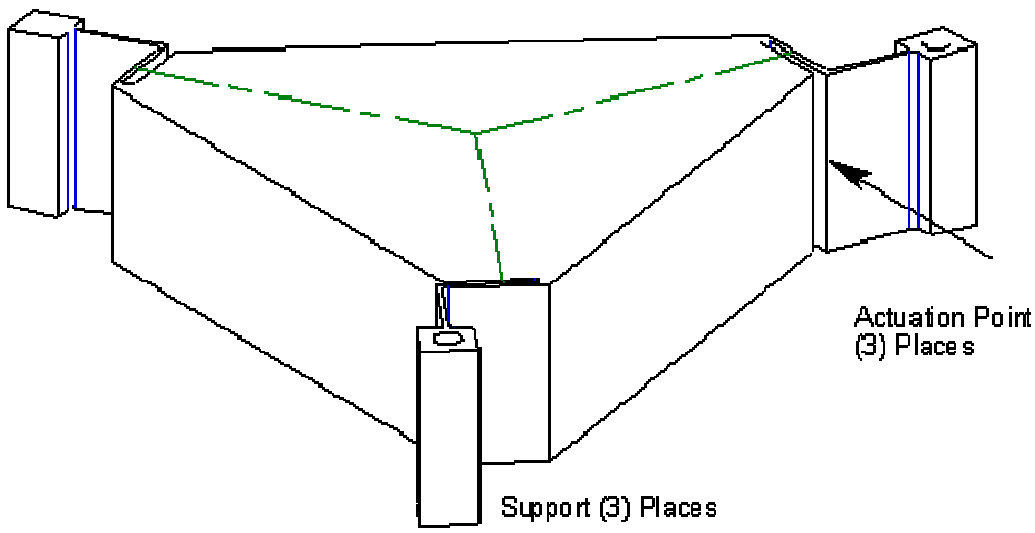


Figure 4.3.4-10: Example of a planar-motion flexure mechanism. Three folded-hinge flexures allow planar motion and each provide one out-of-plane constraint at the fold line. Together with three actuators, coordinated X-Y-θ_Z motion is possible.

4.3.4.6 Risk-reduction activities

Thorough engineering analysis will reduce risk to a minimum. There are still aspects pertaining to focal plane flatness where testing and prototyping will be prudent activities. Foremost is the effect of the low operating temperature on components and assemblies manufactured at room temperature. Full prototypes of sensor substrates and their assembly into a raft would be tested. Material options for the focal plane structure would be tested as simply as possible, for example, plates of the same characteristic size and shape. A subsection from the integrating structure, perhaps only one cell holding a single raft, would be built to test the connection/adjustment scheme for sensor packages. These tests will also expose any problems with precision assembly and allow time to develop solutions or revert to other strategies. Building full-scale mockups will be useful to investigate space and access issues particularly for assembly operations. Since silicon carbide is such an attractive material for the structure, different design/manufacturing options will be explored.

4.3.5 Mechanical Structures and Mechanisms

This section describes mechanical structures and mechanisms that integrate key camera components and subsystems into the full camera assembly. The optical design shown in Figure 4.3.7-1 and system requirements provide the starting point for mechanical design. Packaging and

space constraints have been considered in developing an initial conceptual design that appears in Figure 4.3.5-1. It provides a first cut at allocating and sharing space among the components and subsystems. Detailed descriptions begin with the inner cryostat housing and generally work outward towards the camera housing.

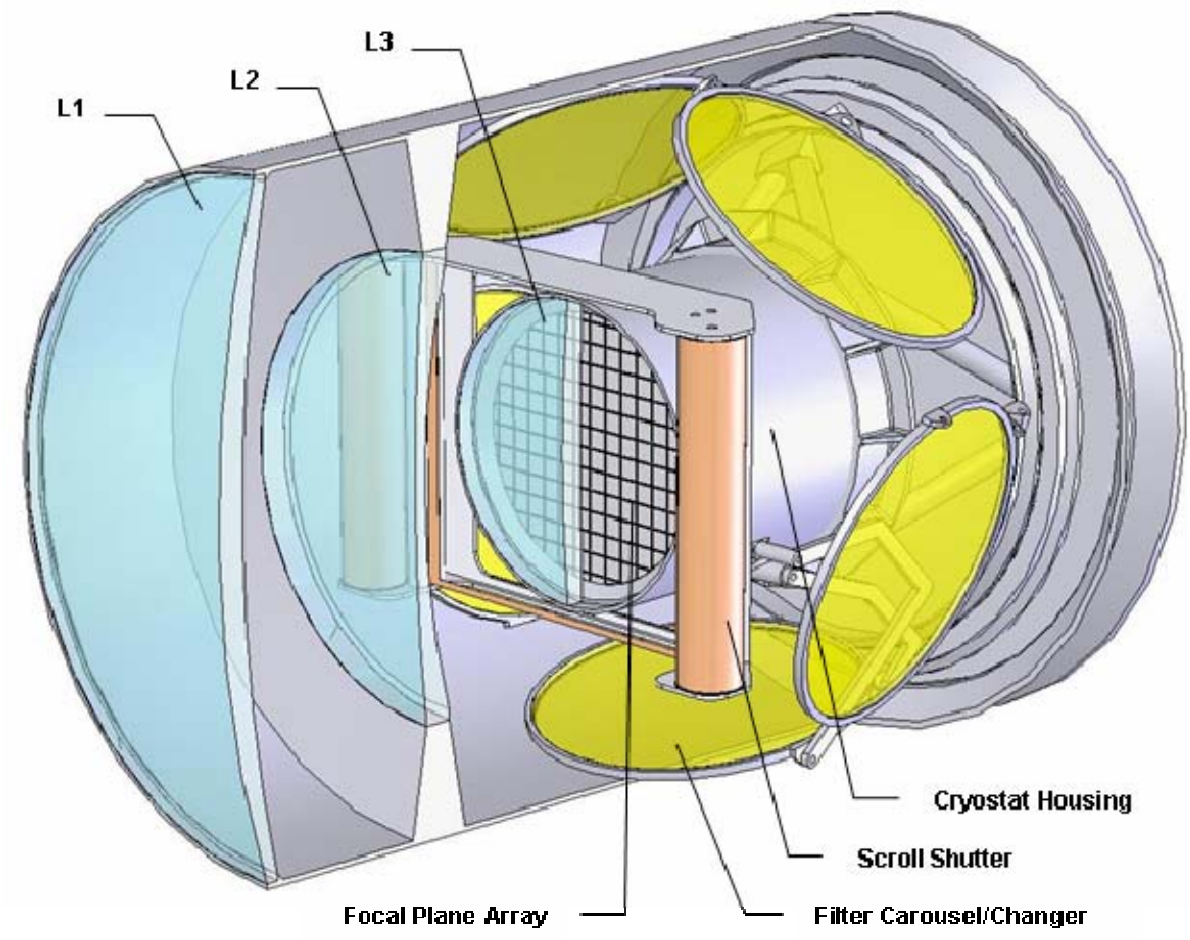


Figure 4.3.5-1: Concept-level design of the LSST camera.

4.3.5.1 Cryostat Housing

The inner cryostat housing contains the focal plane array with its image-stabilization actuators, front-end and back-end electronics, elements of the thermal management system, and all electrical, optical, fluid and mechanical feedthroughs. Optic L3 forms the window on the front end of the cryostat housing. A rear flange to the inner housing provides primary access to components inside.

4.3.5.1.1 Requirements and constraints

1. As the container for the focal plane array, it shares in fulfilling requirements for the focal plane array pertaining to precise location, temperature control and cleanliness.
2. Subsystems in the immediate vicinity of the cryostat housing are the shutter mechanism and the filter change mechanism. These subsystems share space and hence place space constraints on each other.

4.3.5.1.2 Thermal isolation of the focal plane array

The focal plane array will be contained in vacuum to minimize its environmental heat load. While vacuum eliminates natural convection, other modes of heat transfer are still present. Multi-layer insulation applied to the cryostat housing all but eliminates radiation heat transfer except where it cannot be such as the vacuum window (L3). Supporting the focal plane array on flexures naturally minimizes conduction through structures. Titanium, Invar 36 and stainless steel are the best metals having low conductivity compared to elastic modulus and strength. Kevlar® fiber has an incredibly low thermal conductivity, an order of magnitude lower than most plastics. With similar strength and stiffness to metals and a working temperature down to 73 K, it would appear to be the ideal choice given a suitable matrix material. This leaves radiation heat transfer with optic L3 as the dominant heat load on the focal plane array. The variation in heat load is minimized by holding L3 and the cryostat housing at a constant temperature, probably near the mean operating temperature of the telescope.

4.3.5.1.3 L3 as a vacuum window

Mechanical stress in the vacuum window has been considered in the optical design. Currently L3 is thick enough to safely carry the atmospheric pressure load at sea level. Further it is possible to achieve even greater safety factor by applying radial edge pressure, which both reduces the bending moment for a plano-convex optic and sets up compressive stress that counters tensile stress due to bending. Experience with large lenses as vacuum windows comes from ICF laser systems such as the National Ignition Facility (NIF) at Lawrence Livermore National Laboratory (LLNL). These systems have the added complication that the lenses become damaged with use, yet they can be operated safely for many years by monitoring damage sites in lenses with diagnostics.

A safe design may be determine using Equation 4.3.5–1 to find the allowable tensile stress that will propagate just one crack, thus avoiding implosion by a controlled leak down.[1] It depends on the diameter of the optic at the support and a material factor determined from destructive tests, fused silica in this case. For L3, the result is 4.53 MPa (657 psi). However, for a crack to propagate there first must be a flaw whose size may be determined using Equation 4.3.5–2, the Griffith fracture criteria. The result for fused silica is 8.7 mm diameter, assuming a crack with the shape and size of a half-penny that penetrates the surface. For an optic without flaws, this design stress has a safety factor in excess of 10:1.

$$\sigma_t = \frac{K_f}{\sqrt{d}} \qquad K_f = 3.79\ \mathrm{MPa}\cdot\mathrm{m}^{1/2} \qquad (4.3.5\text{–}1)$$

$$a = \frac{1}{\pi}\left(\frac{K_{1c}}{\sigma_t}\right)^2 \qquad K_{1c} = 0.748\ \mathrm{MPa}\cdot\mathrm{m}^{1/2} \qquad (4.3.5\text{–}2)$$

A finite element model of L3 on a realistic support was developed to determine the stress under load. As Figure 4.3.5-2 shows, atmospheric pressure at sea level produces a maximum design stress of 4.32 MPa (626 psi), less than the allowable stress mentioned above. Lower pressure at the telescope site, of order 70%, will provide even greater design margin. In addition, this plano-convex optic affords us the opportunity to greatly reduce tensile stress by applying radial edge pressure (effectively compressive stress) with the optic mount. This also decreases the deflection of the optic, which otherwise may require compensation in the optical design.

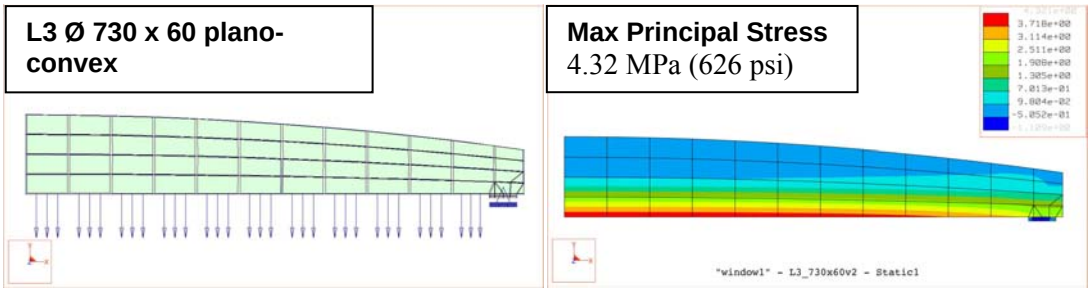


Figure 4.3.5-2 Finite element model (left) and plot of maximum principal stress (right). Positive stress is tensile in the normal sign convention used here. The plot looks unsymmetrical because the upper compressive stress shows up in a plot of minimum principal stress, which is the mirror image of this one.

4.3.5.1.4 Alignment and stability

Atmospheric pressure on L3, its mount and the cryostat housing will cause significant displacements upon pump down compared to alignment tolerances. Finite element modeling should predict magnitudes to 90% accuracy, but means of physically or optically verifying alignments must be considered in the design phase. In a similar way, bringing the focal plane down to operating temperature will cause displacements as well. Characterizing these displacements as well as focal plane flatness may be done optically more easily using a planar vacuum window in place of L3. In addition, it will be useful to physically measure the position of the focal plane (in three degrees of freedom) from outside the evacuated cryostat housing. One simple way to achieve this is to measure (or mount) through the housing directly along the three constraint lines of the flexure mounts that support and define the position of the focal plane array. The front flange of the housing, where the flexure mounts attach, will elastically deform as a diaphragm due to the vacuum load on L3 and the flange, carrying the focal plane with it. If the measurements were not made along the constraint lines, a displacement error would result due to the offset times the angular deflection of the flange. A second-order error due to angular deflection remains but is typically small in comparison.

Variations in barometric pressure, temperature and orientation of the telescope affect the stability of alignment. The approach to barometric pressure will be to reduce sensitivity to a sufficient level. Temperature control of the housing and L3 is already planned to minimize changing heat load to the focal plane. These approaches do not work well for telescope orientation; rather one of matched displacements is more viable. Mounts for L1, L2, L3 and the focal plane array may have more or less compliance engineered in certain directions to maintain a common axis and fixed spacing regardless of the gravity vector. For example, the static deflection of a 10 Hz mass-spring oscillator, typical of the telescope structure, is 2.5 mm. It drops to 25 μm for 100 Hz but it would be difficult to achieve such a high first mode frequency for a structure the size of the camera.

4.3.5.2 Shutter Mechanism

The shutter mechanism, like on any camera, defines the length of time that sensors are exposed to an image. An early feasibility study recommended a scroll shutter as the most practical solution to achieve LSST requirements and constraints. More recent thinking reached the same general conclusion but revealed several design alternatives. In the description that follows, a baseline design is presented with mention of design alternatives as appropriate.

4.3.5.2.1 Requirements and constraints

1. The minimum required exposure time for one closed-open-closed sequence is 1 s.
2. The maximum expected exposure time is 10 s and is not a design driver.
3. The variation in exposure time across the focal plane array should not exceed one part in 10^3.
4. The variation in time stamp across the focal plane array should not exceed 1 s.
5. The aperture must be sufficient in size when open so as not to obscure the focal plane array.
6. When closed, no light should reach the focal plane array.
7. The mechanism should operate without producing contaminants to maintain the cleanliness level on the optics.
8. The operational life should be 10^6 cycles per year for 10 years.
9. The frequency of maintenance should be less than once per year.
10. Any force disturbance to the telescope should be negligible compared to wind loading, for example.
11. The placement of the shutter sheet will be between L3 and the filter to minimize its size. Subsystems in the immediate vicinity are the cryostat housing and the filter change mechanism. These subsystems share space and hence place space constraints on each other.

4.3.5.2.2 Equal exposure time

A rectangular aperture that moves across the array controls the exposure time for all pixels. Equal exposure time requires the opening edge and the closing edge to have the same motion profile in time. This is achieved in the baseline design using one long sheet with a central rectangular aperture of constant width. For compact storage, the sheet attaches to and wraps around two rollers to form a scroll, as Figure 4.3.5-3 shows. Figure 4.3.5-4 shows the motion profile for a constant-width aperture. Notice there are two closed positions so the sheet advances in the same direction from closed to open to closed. The sheet reverses direction for the next cycle. Figure 4.3.5-5 shows an alternative version that uses two sheets whose positions are independently controlled, thus creating a variable-width aperture. This would be of great benefit if the minimum exposure time were much shorter than currently required. Both approaches require a tensioning scheme to roll up the sheet(s), but it is more complicated for the variable-width aperture.

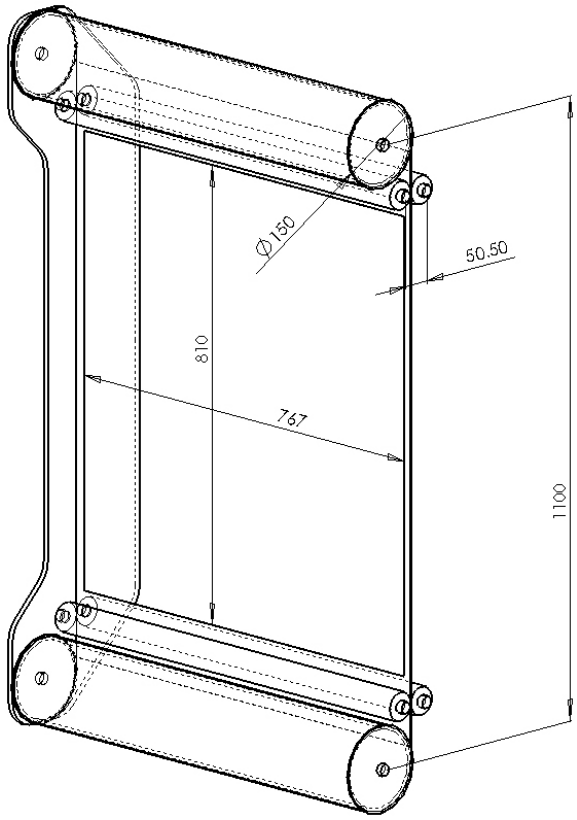


Figure 4.3.5-3 Proposed design for the shutter, incorporating both the one-sheet and the two-sheet alternatives. All dimensions in mm. The side frame has been removed for full view of the rollers.

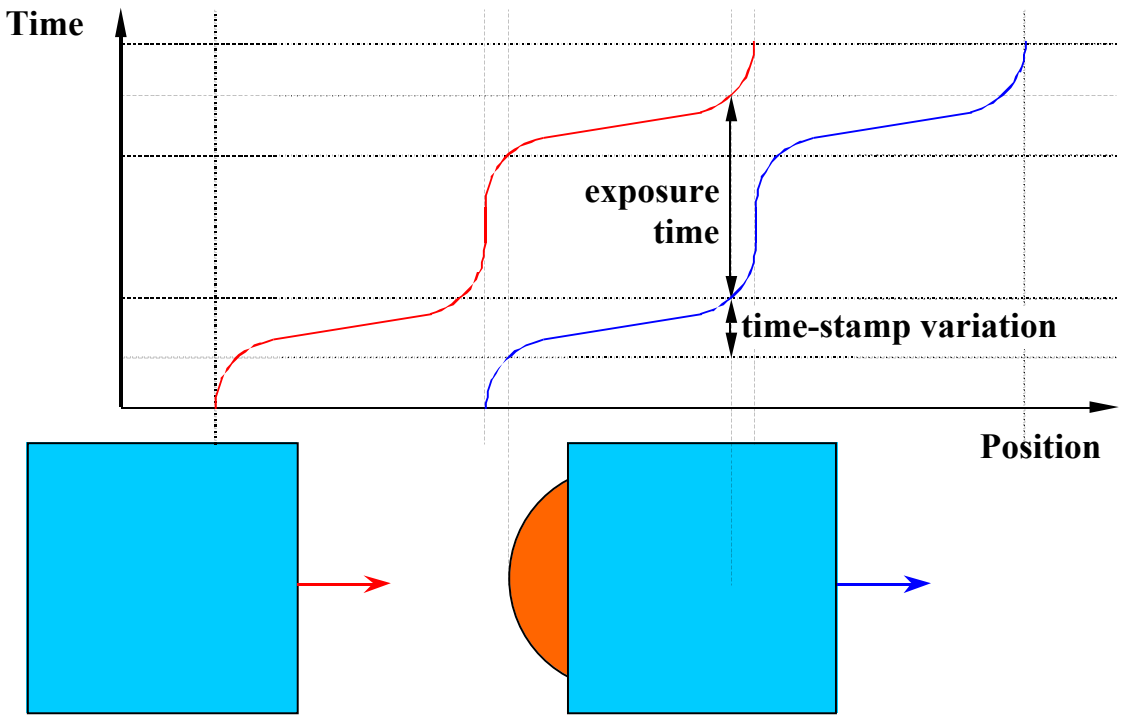


Figure 4.3.5-4: Time vs. position of a constant-width aperture through a closed-open-closed sequence. The curves correspond to the leading and trailing edges of the aperture. The next sequence would occur in the opposite direction.

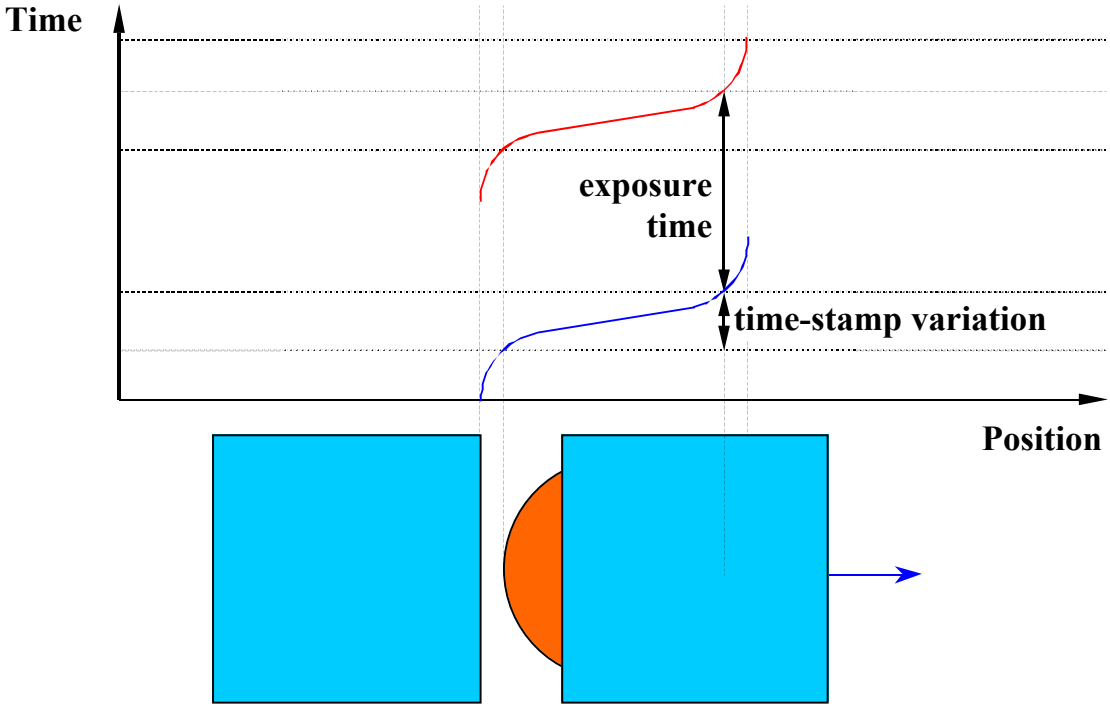


Figure 4.3.5-5: Time vs. position of a variable-width aperture through a close-open-close sequence. The curves correspond to the leading and trailing edges of the aperture. The next sequence would occur in the opposite direction.

Controlling the exposure time to millisecond precision is equivalent to controlling position of the aperture edges to several tenths of millimeters. Standard feedforward and feedback control techniques should easily maintain acceptable following error (i.e., the controller component of error) for this well-defined motion profile. Placing position feedback sensors directly on the sheet would minimize the total error but this may present implementation/reliability problems. Placing sensors on the rollers is easy and reliable and the average of the two accounts for the effective roller radius changing as the sheet rolls up on itself. This effect is small for a thin metal sheet. In addition, a metal sheet is opaque, does not outgas, can be processed for high fatigue strength and low wear, and is stable and rigid in plane.

4.3.5.2.3 Constraining the sheet

Proper constraint of the sheet will be important to the performance and life of the shutter. It is useful to consider in-plane and out-of-plane constraints separately. The sheet is effectively rigid in plane (including any portion tightly wrapped around a roller) so over constraint must be avoided. The large aperture in a single sheet complicates the issue because the sheet is weak in shear but not freely compliant. If the rollers constrain shear in the sheet, then their alignment must be precise to manage stress in the thin sections. If the rollers allow freedom in shear and the shear load is low enough, then the mechanism aligns itself. The latter approach is accomplished by giving each roller an in-plane rotational degree of freedom. There is some design freedom on where to place each pivot but ideally it would be on the centerline of the sheet and near the center of mass of that roller.

A similar situation occurs when applying tension to the sheet. One of several solutions used in anti-backlash transmissions may be employed to compliantly counter rotate the two scroll rollers. Both rollers are driven simultaneously either with a transmission that has an internal degree of freedom (similar to an automotive differential) or with two motors controlled in the same way. Typically a spring in the transmission applies preload to the degree of freedom, which tensions the sheet without loading the drive motor. The approach using two drive motors continuously generates heat to tension the sheet, or the sheet goes slack when the power is off. Figure 4.3.5-6 illustrates a simple tensioning device that uses one motor to drive both rollers with metal timing belts. The motor would mount with freedom in the direction of the preload force. This mechanism naturally accounts for changing effective roller radii thus allowing a single position feedback sensor to be used at the motor.

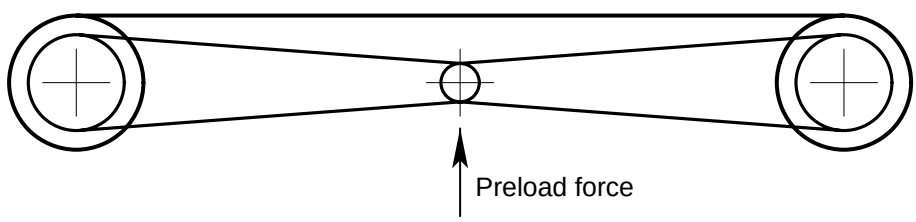


Figure 4.3.5-6 Mechanical tension mechanism for the one-sheet shutter. A motor (not shown) drives both rollers with equal torque while the preload force causes equal and opposite torque to tension the sheet. The geometry of this mechanism keeps these two functions orthogonal to one another.

Since the sheet is very flexible out of plane, full constraint is not practical, nor is it needed. The rollers constrain two ends of the unrolled sheet, and the sides between rollers are constrained loosely within polymer-lined tracks. It is important that the tracks not end too close to the rollers or be too tight since the effective radius varies with travel position. The edges of the aperture also need some constraint because there is a tendency for them to bow due to tension applied along the sides of the sheet. A simple solution is to place a beam along the edge. It could be a separate part

attached to the sheet or it could be formed in the sheet as Figure 4.3.5-7 shows. The beam must be shorter than the aperture and on the side shown so that it rolls up cleanly.

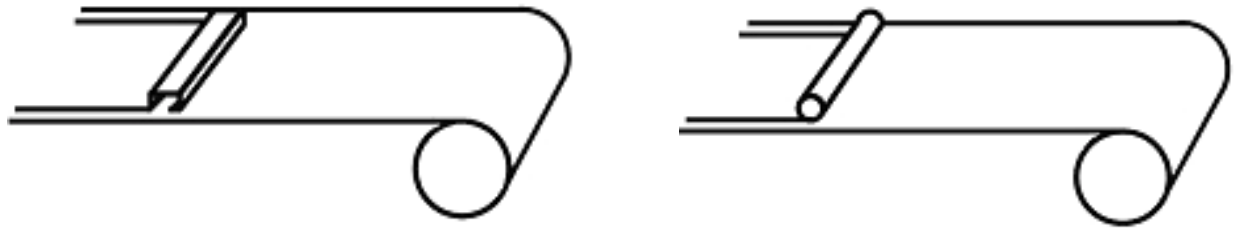


Figure 4.3.5-7: Possible edge stiffening beams formed on the sheet. Further though is needed on whether a formed beam is sufficiently sharp for the aperture. Instead, a separate formed beam could be spot welded to the sheet leaving a sharp edge.

The variable-width aperture with two sheets is simpler from a constraint perspective; one roller per sheet constrains all three in-plane degrees of freedom. The difficulty is in the tensioning device since it should not constrain any in-plane degree of freedom. Figure 4.3.5-8 shows an extra set of narrow pinch rollers per sheet for this purpose. The pinch rollers would drive the sheet through friction. The disks' axial compliance releases one in-plane constraint by allowing a slip angle to exist between the planes of the disks and the path of the sheet. To a lesser degree the same phenomenon occurs in the driving direction. Unfortunately this makes a mechanical tensioning device rather complicated to prevent the gradual loss of preload, although several workable solutions have been identified.

Figure 4.3.5-8: Two-sheet design using two pinch rollers to tension each sheet on its main roller.

Issues associated with out-of-plane constraint are no different for two sheets than those discussed previously for one sheet, except that the tendency to bow from tension will not exist with two sheets. A beam across the edge will still prevent excessive gravity sag of the thin sheet. Since each sheet operates independently from the other, the edge stiffener never has to pass between the pinch rollers.

4.3.5.2.4 Risk-reduction activities

Fatigue and wear must be evaluated on any mechanism expected to operate many millions of cycles and still maintain precise operation. Fatigue life of a sheet in simple bending around a roller is straightforward to engineer. Stress concentration in the corners of the aperture requires more detailed study using finite element analysis, and physical testing will be necessary to be confident of a robust design. Simple wear models can be used to base selection of materials and treatments, but ultimately a full-scale prototype will be built and life tested under operating conditions, particularly in dry nitrogen at expected site temperatures and different orientations. This will allow investigation of different materials and expose wear and contamination problems under realistic operating conditions.

4.3.5.3 Filter Change Mechanism

The filter change mechanism stores a set of color filters and inserts any one as needed for the particular observational task. The camera concept design illustrated in Figure 4.3.5-1 shows that five color filters fit within the camera housing on a carousel that rotates about the camera axis. The carousel presents the appropriate filter to the lower station where the exchange can be made through the space between the shutter rollers. A simple swing arm cannot move the filter along a path that misses all physical constraints, namely L2, the edge of the shutter, the camera housing and the cryostat housing. A four-bar linkage (illustrated in Figure 4.3.5-1) can generate a suitable path but the links may become unwieldy and obtrusive. If this becomes a problem as the design evolves, then unless a better implementation of the four-bar linkage can be devised, an approach using a curved track and followers can be implemented because the motion profile is more flexible and the mechanism is more compact. A wheeled cart (transporter) then conveys the filter along the tracks between the carousel and use position. A number of engineering details must be worked out to make this notional concept a functional device. Planned approaches to the more significant ones are presented in following sections.

4.3.5.3.1 Requirements and constraints

1. Planned operations of LSST require up to four color filters per night. Five filters are required long term with seven being a goal.
2. The time goal to exchange filters on the nightly scale is one minute. Thus the filter changer should operate at any orientation of the telescope. Changing the complement of filters, if needed, could be accomplished during the day.
3. The location of the filter should be as close to L3 as practical to minimize its size.
4. The filter changer should not cause additional obscuration over that allowed in the optical design.
5. The mechanism should operate without producing contaminants to maintain the cleanliness level on the optics.
6. Subsystems in the immediate vicinity are the cryostat housing, the shutter, L2 and the camera housing. These subsystems share space and hence place space constraints on each other.
7. The mechanism must be reliable and have a long (~years) service life.

4.3.5.3.2 Carousel bearing and drive

The cryostat housing occupies the central region of the carousel where bearings would normally go. This leads to a rather large annular bearing and drive. There are a number of solutions and off-the-shelf hardware available from the automation industry. For example, curved track segments can butt together to form a full annulus. However, a continuous annular track is not too large to manufacture and it could become a structural part of the rotating carousel supported on steel wheels. One or more wheels may be motorized to drive the carousel through friction since it is unbalanced only in five stationary positions that are easily locked. To minimize cycle time, the carousel should move the shortest angular distance to the next filter. This means the carousel should rotate in either direction as needed and over time could make many revolutions. Getting utilities to the rotating carousel would be difficult so the filter latches will be remotely powered only at the loading station, probably through a mechanical coupling. It will be important to design/select components for clean operation and to use covers to capture particles before they reach the optics.

4.3.5.3.3 Transporter track and cart

If the four-bar linkage arrangement becomes unwieldy, a transporter track and wheeled cart arrangement could replace it. The type of track used for the transporter may be similar to the carousel but with a more complicated shape. However the track(s) will be stationary and the wheeled cart will carry the filter. The motor will go on the cart and drive through friction, a geared track or a toothed belt. Utilities may be delivered simply over the finite length with an umbilical cord. The accuracy needed to place the filter in the optical system is not particularly demanding but it may be more that the cart can do alone considering the changing orientation of the telescope. Instead, the cart will push the filter bezel into additional fixed constraints, although over-constraint should be avoided. A similar approach may be used at the carousel position to minimize the number of latches to activate.

4.3.5.3.4 Fail-safe latches

The latches and other mechanisms that hold filters must not rely on power to stay engaged. Power should only be used to release latches. Latches can be spring loaded to engage but this requires strong springs and large release forces. A toggle mechanism (a.k.a. over-center mechanism) is useful in this respect as it can generate large clamping forces from lighter springs. More conventional latches as used for doors are also acceptable. Power to release latches may be pneumatic (compressed nitrogen) or electric.

4.3.5.3.5 Risk-reduction activities

A full-scale prototype of the transporter complete with latches will be built primarily to test functionality but also it will expose any wear and contamination problems under realistic operating conditions.

4.3.5.4 Optic mounts

The optic mounts provide the interface between L1, L2, L3 and filter optics and the structures or mechanisms that support them. These interfaces may include adjustment capability for alignment purposes. It is likely that vendors will be asked to supply the optics already installed and qualified in the mounts, as several vendors have proposed. Of the several benefits this provides, foremost perhaps is leveraging technical experience and facilities that otherwise would require duplication. It may also be advantageous for the vendor to test all the camera optics as a system. The RFQ will include designs of the mounts and other camera hardware that the vendor needs to manufacture and test these opto-mechanical assemblies.

4.3.5.4.1 Requirements and constraints

1. L1 and L3 mounts must be gas tight. L1 contains dry nitrogen within the camera housing and L3 is the vacuum window for the cryostat housing.
2. L1 and to a lesser extent L2 must operate over a fairly wide temperature range on the order of 30 ºC. For a steel mount and the fused silica L1 optic, the differential radial expansion would be 275 μm. The other optics operate in a temperature controlled zone that is approximately 20º C lower than assembly and testing at room temperature.
3. Alignment tolerances must be maintained for changing conditions at the telescope, including orientation, temperature and barometric pressure.
4. As a precaution, the optic should be demountable such as to allow refurbishment.

4.3.5.4.2 Filter bezel

The filter substrates differ from one another in their central thickness, ranging from 13.5 to 22 mm. The convex spherical radius is the same for all filters and is placed the same inside the camera, thus it is a logical surface for the bezel to register. The bezel can accommodate different

filter thicknesses using shims, for example, to space off the clamp ring. Outwardly to the filter changer, all bezels are the same except for a means to encode the identification of each. A simple method for mounting optics is between two o-rings, one on the bezel and the other on the clamp ring. The compliance of the o-rings accommodates tolerance in the parts and differential thermal expansion. Large-diameter o-rings are normally made from extruded stock so the cross section is well controlled. Further there is no need to join the ends since this mount does not have to form a seal.

4.3.5.4.3 L3 vacuum mount

The use of L3 as a vacuum window was discussed in a previous section on the cryostat housing, which is the structural element of the L3 mount. The optic is sufficiently thick to withstand stress for a simple annular support just outside the clear aperture. A thin elastomer gasket between the optic and the housing carries the vacuum load while an o-ring in the clamp ring maintains preload when the chamber is vented. The thickness-to-width ratio of the gasket has a strong effect of the degree of compression under approximately 26 kN (5900 lb) of vacuum load. Friction at the gasket keeps the optic centered. There will be an o-ring seal between the housing and clamp ring so that the gasket is not the primary seal. The cavity outside the gasket may be separately evacuated to provide a double seal or vented to the chamber.

The use of radial edge pressure was suggested as a way to reduce tensile stress in the optic. While this appears unnecessary, it is useful to mention how it could be implemented as a backup option. Conceptually one could achieve this effect by making the annular support conical as on a stopper. The half angle of the cone would be of order 10º and the axial compliance of the gasket interface would be approximately 30 times greater than if flat. Further this does not define tilt of the optic very well. It would be better to support the optic as before and add a metal hoop to the outer diameter. Epoxy under hydrostatic pressure would be applied between the optic and hoop then allowed to cure. In this case the hoop could act as the mount and directly interface to the cryostat housing without need for a compliant gasket.

4.3.5.4.4 L1-L2 subassembly

In terms of optical correction, L1 and L2 are the most significant optics in the camera. It stands to reason that their alignment to one another is most critical and forms logical datums to use in aligning the focal plane assembly (including L3). In addition, the size and location of L1 and L2 on the front end of the camera (away from the shutter and filter changer) make an aligned and tested subassembly a reasonable package for an optics vendor to produce. Further, it would be reasonable for an optics vendor to test the L1-L2 subassembly with L3 and a dummy filter mounted temporarily in the interferometer, thus avoiding complications from the shutter, filter changer and focal plane assembly. A repeatable kinematic interface allows the front end of the camera (L1 and L2) to be easily removed from and reinstalled on the camera housing to provide ample access to components inside such as the filter carousel. The interface may also contain the adjustments needed to align the focal plane assembly. The front end plays a role in the thermal system too; acting as a dual-pain window between ambient temperature outside and constant temperature inside the camera housing.

Presently the L1 and L2 substrate designs have a narrow annular flat that may be used for mounting. It would be convenient to bond to this flat with a number of radial-motion flexures, perhaps 48 on L1 and 12 on L2, cut into the optic mount. Several options exist to demount a bonded optic. A relatively thin epoxy bond of order 0.1 mm can be cut with a slitting saw or decomposed by inductively heating the metal pad on the flexure. A thicker RTV rubber bond can be cut with a wire, similar to a cheese cutter. The flexure mounts can be made with detachable pads that permanently stay with the optic once bonded. The gas seal for L1 is made on the outer convex surface with a highly compliant seal, such as a tubular o-ring or u-cup, in a captive ring.

The labyrinth offered by the captive ring on L2 is sufficient to keep separate the two thermal zones.

The mechanism for focus adjustment is necessarily part of the L1-L2 subassembly. There are two reasonable constraint arrangements given the basic geometry of the L1-L2 subassembly. One arrangement uses three linear actuators in the axial direction and three links tangent to a circle about the axis. The length of axial travel compared to the overall size is small enough to favor flexure bearings for the links. The actuators may be custom built from commercial components, for example, a ball screw driven with a stepper motor through a harmonic drive, or purchased complete from a vendor. The actuators must have reasonably accurate positioning from a home position to establish tilt and initial piston of the optic, for example, using the null position of LVDT sensors. Final piston is determined from wavefront reconstruction. The second arrangement uses one linear actuator and five or six links. The actuator may drive through a linkage that supports L2 symmetrically if an off-center connection causes too much tilt error. The symmetry of six links, two axially displaced circles each with three tangent links, is desirable for the same reason. Finite element analysis will be used to evaluate these options, opting for the simplest solution that works.

4.3.5.4.5 Risk-reduction activities

The consequence of L3 failing catastrophically is so great that the safety margin should be verified through full-scale testing. This would occur in several stages beginning first with a mock optic in aluminum. This allows verification of the finite element model for both stress and deflection. Compression of the gasket would also be measured and if necessary design modifications would be made. The second stage tests would use flat optical glass with similar properties as the fused silica optic. The glass would be loaded beyond the design stress, perhaps two times, to establish a reasonable proof test for the real optic. Then the tensile surface would be intentionally damaged to simulate a flaw and tested up to the design stress. This test would repeat inflicting more damage until the glass finally fractured. If the empirical model is correct, the glass should crack causing the pressure to slowly leak away. A final proof test would be made on the L3 substrate(s) probably intermediate in manufacturing.

4.3.5.5 Camera housing

The camera housing is the main structural element of the camera tying together the cryostat housing, L1-L2 subassembly and filter changer. It must be stable to optical tolerances, which is aided by the temperature controlled nitrogen gas it contains. An outside layer of insulation minimizes heat transfer to air in the light path around it. Side access panels will be provided where service can be done reasonably from outside. It may be possible to remove the shutter and individual filters through side access; otherwise access will be from the front after removing the L1-L2 subassembly at the kinematic interface. The cryostat housing may also have a kinematic interface to the camera housing and probably install from the front, although removing it should be infrequent since access to the rear flange of the cryostat housing should be straightforward from the rear of the camera housing. The camera housing is supported from its back annular surface by the rotation stage and hexapod positioner to be discussed next.

4.3.5.6 Rotation stage and hexapod positioner

A six-axis positioning system situated between the camera and the telescope maintains alignment of the camera axis, adjusts for focus and tracks rotation of the sky. A hexapod is a parallel kinematic mechanism capable of moving in six degrees of freedom. However it cannot provide enough rotation about the optical z axis. For this reason a large-angle bearing is also provided.

4.3.5.6.1 Requirements and constraints

1. Provide controlled actuation of the camera with respect to the telescope in six degrees of freedom.
2. The range of motion required for four degrees of freedom, x-y translation and rotation about x-y, is driven primarily by deflection of the telescope and camera structures due to changing orientation and temperature. These ranges could amount to several millimeters and several milliradians. In addition to deflection, z translation is used to adjust the focus for different color filters. This range could also amount to several millimeters. These five degrees of freedom are stationary during an exposure. Rotation about z has two widely different range requirements. The widest range is also stationary during an exposure, ±90º. The other must track during an exposure to keep the detector array rotating with the image of the sky. This range of motion depends on the length of exposure and the proximity to the singularity of the telescope's altitude-azimuth mount. The rotation of the Earth over a 10-second exposure is 0.7 mrad. The needed range could be an order of magnitude larger, say 7 mrad.
3. The resolution required for three degrees of freedom, x-y rotation and z translation, must be sufficient to achieve best focus as determined by wavefront reconstruction within ±1 micron. The tracking resolution for z rotation must be sub-pixel or a few microns at the edge of the focal plane primarily for smooth motion. Any inaccuracy in velocity will be corrected by image-stabilization actuators at the focal plane. The resolution required for x-y translation will be determined from a sensitivity study of the final optical design, particularly a shift of the camera optical axis with respect to the telescope optical axis.
4. The z-axis error motion during tracking should be no greater than the resolution requirements.
5. The accuracy requirement over full travel may be relaxed somewhat from the resolution requirement because periodic feedback will come from wavefront reconstruction. A reasonable goal is 3-5 times the resolution.

4.3.6 Thermal System

4.3.6.1 Introduction

An ideal camera thermal system would maintain all camera components at desired temperatures and operational conditions without adversely influencing the camera or telescope performance. The thermal system would prevent the exchange of heat or dynamic perturbations within the telescope structure or air in the field of view. Likewise all heat transfer would be balanced and the ideal operating thermal conditions would be maintained at no cost to the camera functions.

In the development process the ideal often times falls prey to the practical. The conceptual design for the thermal system presented herein represents preliminary results of studies and analysis intended to identify the actual driving requirements and achievable operating conditions of the LSST Camera.

The basic requirement of the Camera Thermal System (CTS) is to allow the solid-state detectors (either CCD or CMOS) that form the Focal Plane Array (FPA) to operate at ~ 1730K to reduce the contribution of thermal noise to the electronic signals while in parallel optimizing quantum efficiency (QE). The thermal gradient across the sensors must also be adequately reduced (±0.30K) to maintain uniform quantum efficiency in each pixel. For efficient camera operation, thermal stability of the FPA is necessary to maintain optical sensor performance as well as to eliminate the growth of mechanical distortions to the FPA, once it is aligned. The additional functions of the thermal system are to provide a particulate free environment for the

sensor surfaces, heat extraction from both the front end and back end electronics systems, and anti-fog protection for the three optical elements. In satisfying these requirements, the CTS must have minimal impact on the ambient environment of the telescope itself.

4.3.6.2 General Description

The Camera Thermal System (CTS) is shown schematically in Figure 4.3.6-1. It is primarily integrated within the LSST camera body but transfer lines for the connections to refrigeration and vacuum system components are expected to penetrate the camera body and be extended through the telescope structure to the observatory hall. Control lines will also follow a similar path.

The camera is composed of a cryostat containing the FPA and its readout electronics and an outer camera body which houses the cryostat, shutter, L2 optic, and the five filters and filter exchange mechanism. The cryostat is closed by the L3 optic and the outer camera body by the L1 optic. Figure 4.3.6-2 defines the primary thermal and vacuum zones within the camera. The cryostat is operated at ~10-6 Torr vacuum while the region between the cryostat and outer body will contain an inert gas (eg: dry nitrogen will be assumed) regulated to be slightly above atmospheric pressure. The temperature of the N2 gas will be controlled to maintain a stable thermal envelope. The other important role of the inert gas is to prevent degradation of the filter coatings. The L3 optic may require active thermal control to minimize heat transfer to the shutter and surrounding components. An external nitrogen flow curtain over the L3 optic is an option for controlling this heat exchange.

The cryostat contains a complex build of components and subassemblies shown for the reader's assistance in Figure 4.3.6-3, Figure 4.3.6-4 and Figure 4.3.6-5. To satisfy the thermal requirements, the primary heat sources that must be addressed in the CTS are the on-chip sensor readout electronics, radiation through L3 onto the FPA, the front end and back end electronics and the warm inner surfaces of the cryostat body. In addition there will be a modest number of structural supports, penetrations and feed-thru connections providing conductive paths to the outer camera body and N2 environment.

The conceptual design attempts to separate thermal control of the sensor plane and mechanical stability issues. Within the cryostat, the mechanical structures and supports are transversely symmetric to the optical axis of the camera allowing us to establish near isotherms along this axis. The sensor support, the raft structure (including the front-end electronics) and the integrating structure, all of which may impact the mechanical alignment of the FPA, are maintained within near-isotherms to reduce differential motions inducing stresses. Convective and radiative heat transfer between these components (and as well with the camera body) is minimized by the use of multilayer insulation (MLI) and by maintaining the cryostat under a modest vacuum. The radiation heat load through L3 onto the exposed surface of the FPA cannot however be shielded and therefore must be actively removed. The small heat load of the transistors along the two edges of each sensor must similarly be removed. Accordingly, the cooling of the sensors is controlled by first isolating them from conduction to the integrating structure and then by routing flexible conducting straps from the backside of the sensor support element, through the integrating structure and terminating them onto a liquid cooled plate. This heat sink establishes an isotherm near the center of the camera. The temperature at the sensor can then be trimmed by the use of a feedback system of heaters on each strap and temperature sensors on each CCD or CMOS detector element. The cold plate also mechanically supports the front end electronics cage. By adopting this cooling strategy, negligible loads are transmitted at any time to the sensors through the integrating structure or the front end electronics.

In the present design, the camera electronics module (CEM) located at the bottom of the cryostat also resides in the cryostat vacuum. Heat from these back end electronics is removed by

a second circuit of liquid, cooling a second heat-sink plate. The CEM is connected to the rafts by flexible cables penetrating both cold plates.

To insure that the FPA remains optically clean, the cryostat is divided into three distinct vacuum regions that are quasi-isolated, allowing us to minimize the migration of contaminants by molecular flow onto the cold sensor surface during cool down, warm up and while in operation.

Most of the vacuum system components as well as the liquid refrigeration system will be located remotely from the camera assembly, and off the telescope to reduce vibration and thermal sources. Some passive vacuum pumping and monitoring capability will likely be provided within the camera body.

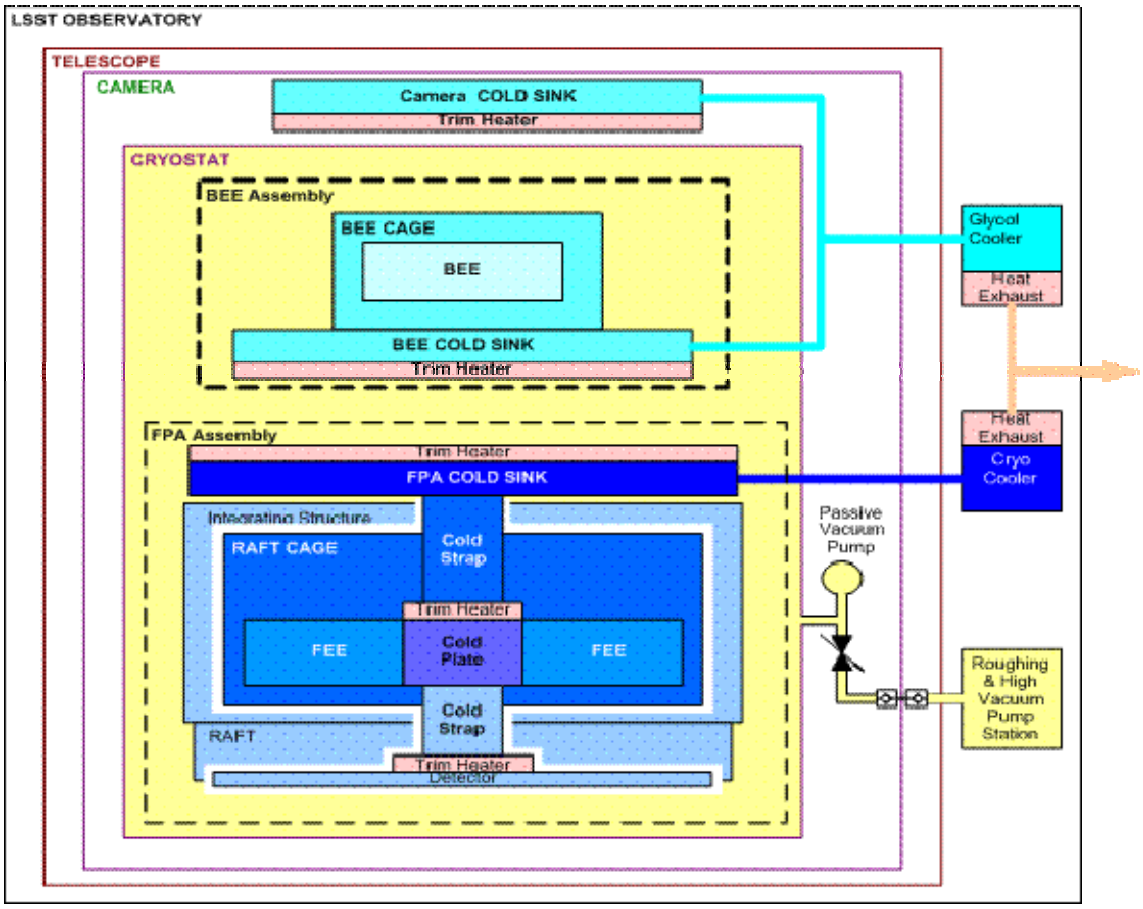


Figure 4.3.6-1 Schematic of the Camera Thermal and Vacuum System

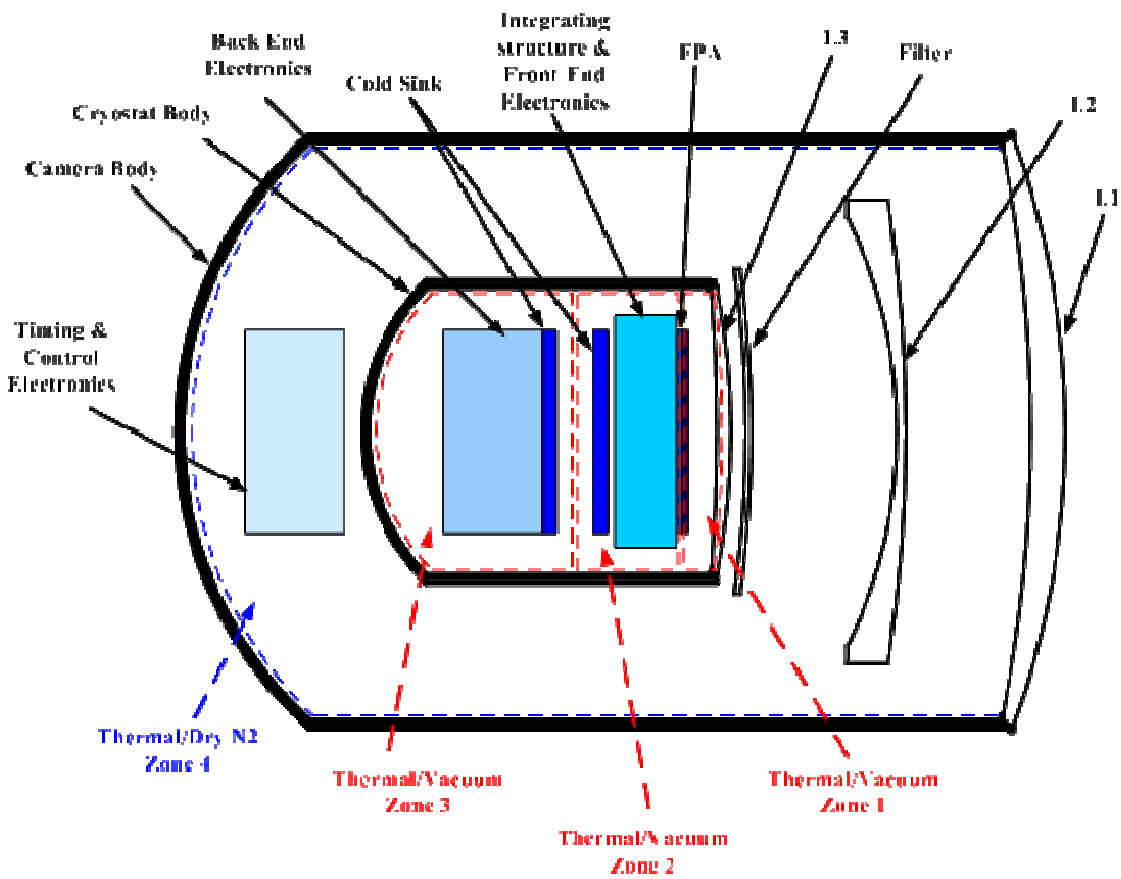


Figure 4.3.6-2 Camera Thermal / Vacuum System Zone Definitions

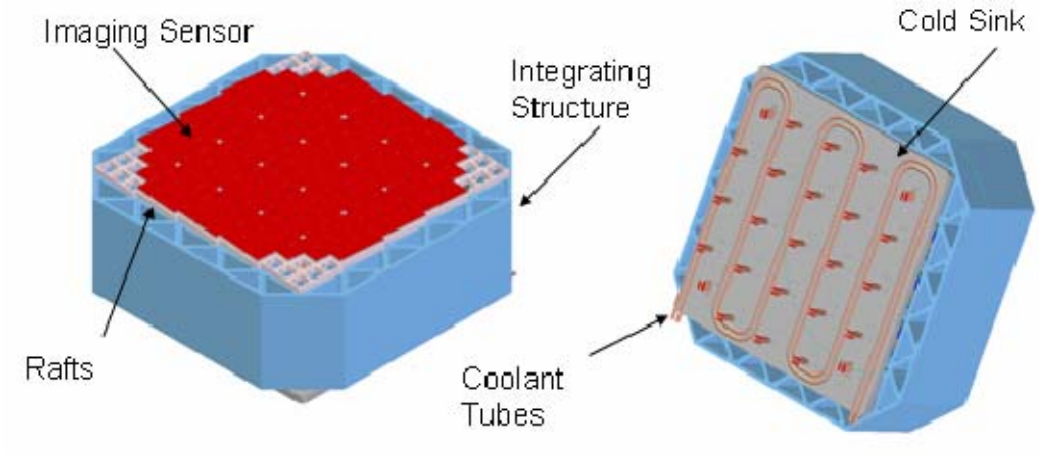


Figure 4.3.6-3 FPA Assembly — top & bottom

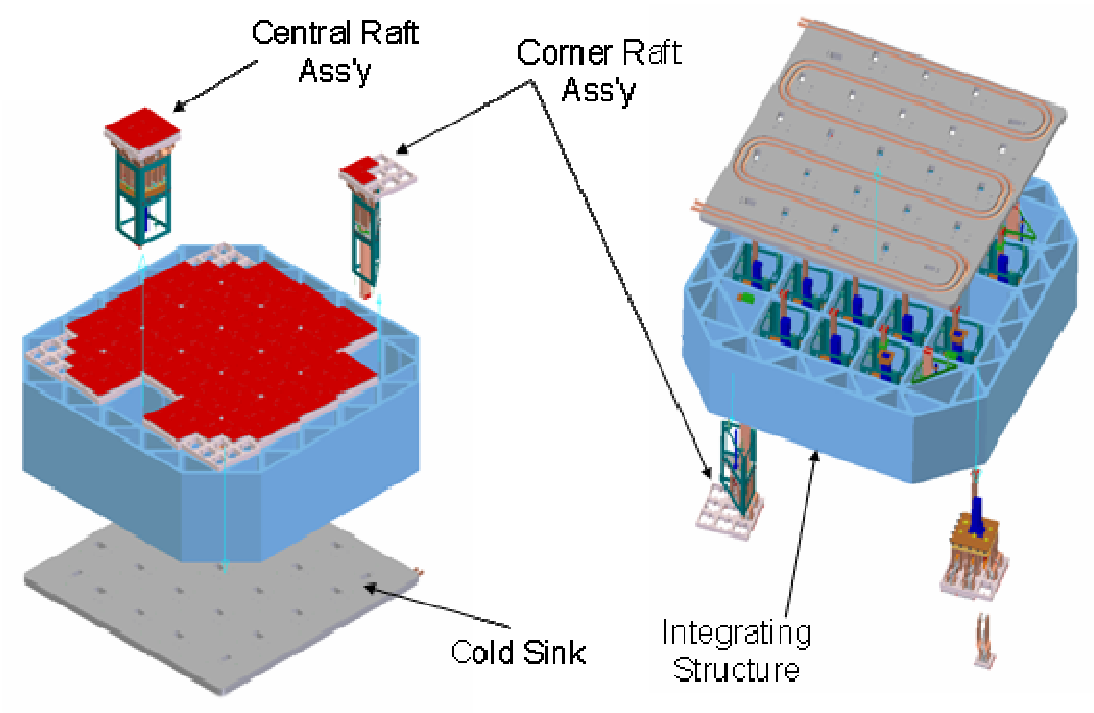


Figure 4.3.6-4 FPA Assembly with Break-out

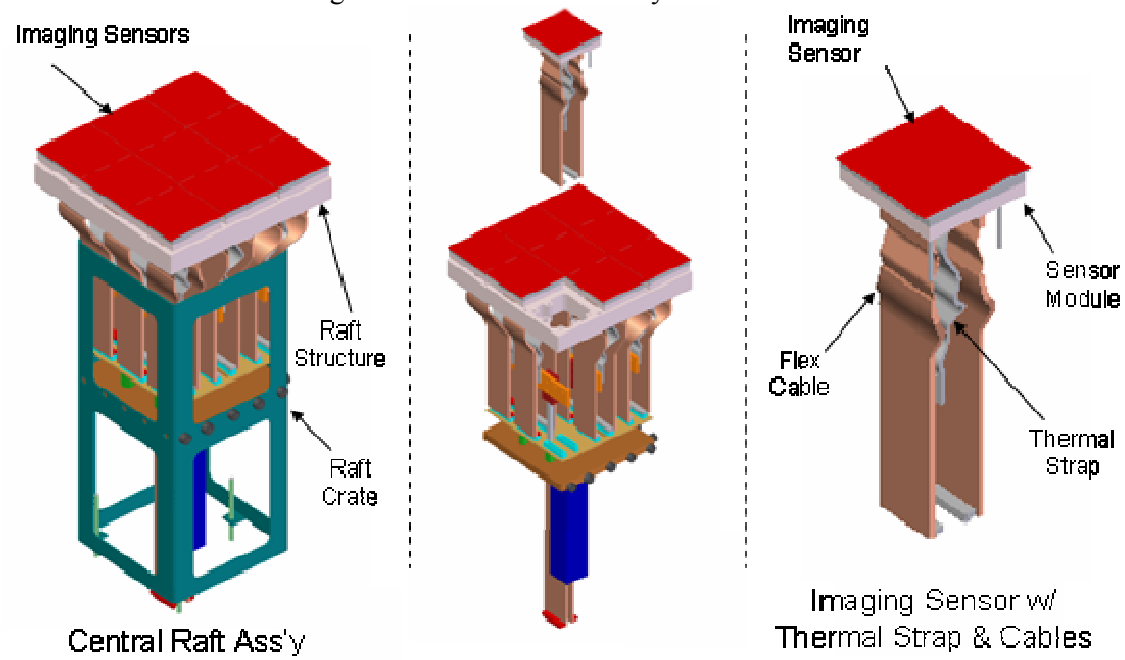


Figure 4.3.6-5 Raft Module with Break-outs

4.3.6.3 Thermal Requirements

The primary functional requirements of the camera thermal system are to maintain set temperatures within the instrument and to remove heat from it in a fully controlled manner, minimizing impact on the telescope operations. Camera thermal load and control conditions are identified and will eventually establish the overall performance requirements of the thermal system and its components.

Camera thermal requirements are presented in the following subsections and tables. These have been derived from the science requirements and drive the preliminary engineering concepts. Wherever possible the source of requirements have been noted.

4.3.6.3.1 Heat Extraction Within The Cryostat

Table 4.3.6-1 (Cryostat Heat Extraction) has been compiled to present the overall cooling requirements of elements within the cryostat itself of the overall camera system. The inner cryostat operates in a modest vacuum, so no convective transfer is present during operation. A conservative design margin has been included as a reserve so as to allow us to continue conceptualization of the thermal system and design at reduced engineering and programmatic risks.

Table 4.3.6-1 Cryostat Heat Extraction Summary.

Heat Source or Transfer Method	FPA Cold Plate (Watts)	BEE Cold Plate (Watts)
Internal Generation	650	350
Conductive	40	40
Radiation	140	10
Control Heaters	70	None
Total Heat Load	900	400
Reserve (100%)	900	400
Design Budget Cooling System	**1800 Watts**	**800 Watts**

The estimated heat loads and heat flow conditions of conceptual camera design have been compiled and are summarized in Table 4.3.6-2 thru Table 4.3.6-4.

The LSST camera's cryostat heat loads have been categorized as follows:

Internally Generated Heat Loads Any heat load within the cryostat produced as a result of the operation of electronics, sensor temperature control heaters, friction, or work. The generated heat appears to be dominated by front-end (FEE) and back-end (BEE) electronics components with a small component from transistors on the CCD sensors. These estimates of power consumption include: 1.5 Watts/sensor, FEE at 3 watts/sensor, and BEE at 1.75 Watts/sensor. These values are derived from engineering estimates of component performance, and the scaling of similar systems to this conceptual design. For a FPA based on CMOS sensors, these estimates would change somewhat.

Estimates of heat generated in position control or actuation of the FPA for tracking and possible remote adjustment mechanisms for individual elements (eg: rafts) of the FPA have not yet been compiled. These heat loads are anticipated to be low and well within the conservative margin of the electronics power estimates.

Table 4.3.6-2 - FPA Thermal Loads and Heat Flow Estimate.

Component	Generated (Watts)	Radiated (Watts)	Conduction (Watts)
Focal Plane Array			
Detector Module			
L3 to Det.		140	
Det. Power	20		
Det. To Det. Mount			160
Det. Mount to Raft			±1
Det. Mount to E-Cable			±1
Det. Mount to T-Strap			160
Trim Heater to T-Strap	35		
T-Strap to Raft Cold plate			195
Front-End Electronics			
E-Cable to FEE Board			±1
FEE Board	630		
FEE Board to FEE T-Bus			630
FEE Board to FEE/BEE Cable			5
FEE T- Bus to Cold Plate			635
Raft Cage			
IS to MLI			0.25
MLI to Cage			0.25
Cage to Cold Plate			±.25
Cold Plate Trim Heater	35		
Cold Plate to Cage T-Bus			870
Cage to Cold Sink			±.25
T-Bus to Cold Sink			870
Integrating Structure(IS)			
Raft to Int. Structure			±1
Cryostat Body to IS-MLI			0.25
IS-MLI to IS			0.25
Cryostat Body to IS Supports			0.5
IS Supports to IS			0.5
IS Actuators	?		
Cryostat Body to IS actuators			0.5
IS Actuators to IS			0.5
IS to Cold Sink			2.25
Cold Sink (CS)			
Cryostat to CS Supports			15
CS Supports to CS			15
Cryostat to CS-MLI			10
CS-MLI to CS			10
CS to Cryo Fluid			900

Table 4.3.6-3 – BEE Assembly Thermal Loads and Heat Flow Estimate

Component	Generated (Watts)	Radiated (Watts)	Conduction (Watts)
BEE Assembly			
BEE Enclosure			
Back End Electronics (BEE)	350		
BEE to BEE Cables			5
BEE to BEE T-Bus			350
BEE T-Bus to BEE Cold Sink			350
CryoStat Body to BEE MLI			5
BEE MLI to BEE Enclosure			5
BEE Enclosure to BEE Cold Sink			5
BEE Cables to Cryostat			5
BEE Cold Sink(CS) Assembly			
Cryostat to BEE CS Supports			40
BEE CS Supports to CS			40
Cryostat to BEE CS-MLI			5
BEE CS-MLI to BEE CS			5
BEE CS to Glycol Fluid			400

Table 4.3.6-4 – Camera Enclosure Thermal Loads and Heat Flow Estimate

Component	Generated (Watts)	Radiated (Watts)	Conduction (Watts)	Convection (Watts)
Camera Enclosure				
Ext. Environment to L3				-10
L3 to Purge Gas				-5
L3 to Camera Body			-5	
Ext. Envir. to Camera Body				-25
Camera Body to Int. Envir.				-25
Camera Body to Camera Mount			-5	
Camera Body to Cryostat Actuators			5	
Cryostat Actuators to Cryostat			5	
CryoStat Actuators	0			
Purge Gas to Cryostat Body				71
Shutter Assembly				
Shutter Mech. to Cryostat Body			5	
Shutter Drive	100			
Shutter Drive to Mechanism			100	
Shutter Drive to Purge Gas				50
Shutter Mechanism to Purge Gas				50
Filter Assembly				
Camera Body to Filter Mount			0	
Filter mount to Filter			0	
Filter to Purge Gas				0
L2 Assembly				

Component	Generated (Watts)	Radiated (Watts)	Conduction (Watts)	Convection (Watts)
Camera Body to L2 Mount			0	
L2 Mount to L2			0	
L2 to Purge Gas				0
Envir. Conditioning System				
Circulating fans	100			
Circulating Fans to Purge Gas				100
Purge Gas to HX				350
HX to Glycol Fluid				350
Purge Gas to Glycol Fluid				5
Camera body to Glycol Fluid Feed Thru				5
Ext. Envir. to Glycol Fluid lines				5
Vacuum System				
Vacuum Pumps (NEG)	5			
Vacuum Ion Gauges	10			
Vacuum pumps/ports to Purge Gas				15
Camera Body to Vacuum Ports			0	
Cryo Fluid feed-thru's				
Camera Body to Cryo Fluid Feed thru's			1	
Purge Gas to Cryo Fluid lines				0.5
Ext. Environment to Cryo Fluid lines				5
Camera Electronics Module				
Camera Electronics	300			
Purge Gas to Camera Elect. Module				300
Camera Body to CEM Cables			-5	
CEM Cables To CEM			-5	

Conductive Heat Load Any heat transferred to the component by means of conduction. For example, structural mounting of components will produce conductive flow path between the component and it supporting parent.

These estimates encompass the heat flow through conductive connections of the internal cryostat components and cryostat outer or warm wall. Amongst these would be the mounting for the integrating structure, and mounts and services for the BEE. Conservative estimates place heat loads at less than 4 Watts per connection. It is expected that 8 to 10 connections each will be required for the FPA and BEE.

Radiative Heat Load Any radiant heat exchange between components. In the cryostat, a primary source of heat load is the radiation exchange between the cold detector surfaces of the FPA and the warm L3 optic. Preliminary thermal radiation loads have been calculated at ~70 Watts heat gain to the surface of the detector array. Calculations were based on conservative emissivities for both the detector surfaces (assuming that an anti-reflection coating has been applied) and L3. A 1000K temperature difference was assumed between the detector array surface and L3.

Radiation/conduction heat exchange between FEE, BEE, rafts, integrating structure, and cold plates is controlled by design through the use of shielding blankets composed of multilayer insulation (MLI). A preliminary estimate of the heat exchange between these components was made assuming the properties of commonly available MLI, deployed within the cryostat in blankets whose thicknesses were chosen to fit conservatively within the space available in the conceptual mechanical design of the integrating structure and other regions within the cryostat. With MLI in place, the heat load from these radiative sources is estimated to be reduced below 1 Watt for the FPA and to be ~10 Watts for the BEE.

Control Heater Load The thermal control scheme for operating the detectors making up the focal plane array assumes the use of trim heaters at either the raft and/or individual detectors levels, and at the FPA cold plate. On detector temperature sensors combined with a thermal-mechanical model will be used for control of the trim heaters. As this heat must be removed through the cold plate, the parameter has been set to limit the heat load of the trim heaters to approximately 10% of generated heat load.

Trim heaters may not be necessary on the FEE since their thermal control is not as critical and the conductive pathways may be mechanically trimmed during assembly, to adjust the operating temperature of the FEE. There should adequate cooling of the FEE and BEE thereby minimizing the need for externally applied heat gain.

Reserve At the present stage of conceptual mechanical and electronics design of the camera, a 100% reserve capacity has been budgeted as a reasonable value for risk management of the program.

4.3.6.3.2 Temperature Control and Environmental Requirements

The camera temperature control and environmental requirements are summarized in Table 4.3.6-5 and Table 4.3.6-6. The LSST camera's thermal control requirements were derived from operational, science, and engineering requirements.

Table 4.3.6-5 Requirements For Temperature Control Within The Cryostat

Parameter		Units	Detectors	Front-End Electronics (FEE)	Back-End Electronics (BEE)
Control Range	T_{rng}	ºC	-80 to -110	-81 to -120	0 to -20
Setpoint	T_{set}	ºC	± 1	± 5	± 5
Knowledge	T_{ave}	ºC	± 0.1	± 2	± 2
Uniformity	ΔT_{un}	ºC	± 0.3	± 0.5 within board	
Drift, Total (from setpoint)	ΔT_{dft}	ºC	± 0.15 over 12 hr		
Variation (detector-to-detector)	ΔT_{d2d}	ºC	± 1		

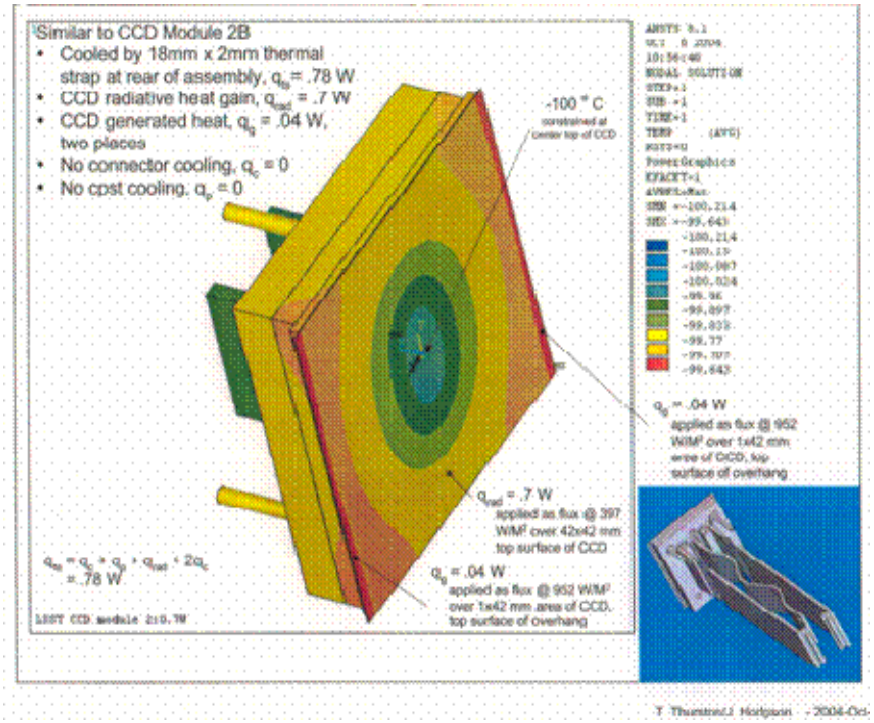
Table 4.3.6-6 Requirements For Temperature Control: Integrating Structure, Camera Body, and Camera enclosure excluding Cryostat

Parameter		**Units**	**Integrating Structure**	**Camera**	**Camera Enclosure Excluding Cryostat**
Control Range	T_{rng}	ºC	-80 to -110	20 to -10	20 to -10
Set-point	T_{set}	ºC	±1of detector set-point	T_{atm} to +5	T_{atm} +5
Knowledge	T_{ave}	ºC	± 0.3	± 1	± 1
Uniformity	ΔT_{un}	ºC	± 1	± 2	± 2
Stability	dT/dt	ºC/min	± 0.05		
Drift Rate	dT/dt	ºC/hr	± 0.5		
Drift, Total (from set-point)	ΔT_{dft}	ºC	± 1		
Heat transfer to Air	Q_{c2a}	Watts		< 50	

These sources are described in more detail in the following paragraphs:

Detectors The operating parameters for the detectors are derived from the science performance requirements as outlined in two reports: the 'LSST Detector Specifications' and the 'Study of Silicon Sensor Thickness Optimization for LSST'. As noted in these reports, the operating temperature for the detectors is not yet established. It will be optimized around low electronic noise and high quantum efficiency. The detector temperature is anticipated to be in the range of 153 to 193 °K. For the control system it is prudent to design the thermal system to accommodate control set-points within this range. These parameters have been established to allow correlation of detector performance using typical calibration and observing procedures.

Validation of the proposed thermal and control concept has been ongoing. Figure 4.3.6-6 shows for example the results of a detailed thermal analysis of a particular concept under consideration for the construction of an individual detector module with cooling provided by a single thermal strap. This modeling applied the predicted thermal loading for radiation, conductive and generated heat loads, as previously discussed. These results show ~0.3º K temperature variation across the exposed surface of a detector, meeting the requirement on the gradient. The predicted thermal distortions of the module were insignificant, less that 0.1µm.



Single Central Thermal Strap Concept:

~0.3 ^{0}K temp. profile at sensor surface.

Figure 4.3.6-6 Detector Module (DDC Module 2B) Thermal Analysis Using A Single Central Strap

With the ascribed parameters, detector performance is intended to be quite stable and thereby producing a weak coupling between telescope thermal conditions and the camera's performance. We believe however that it will be necessary to have an active system to control detector temperature at this level.

Front End Electronics The requirements for thermal range and control are primarily driven by the detector heat extraction concept. As part of the thermal concept, the FEE (which lie in close proximity to the detectors) are operated at colder temperatures than the detectors. This minimizes heat exchange between the detectors and prevents heat flow to the detector modules. The nominal operating temperature of the FEE is estimated to be as much as 10 °C below the nominal operation temperature of the detectors. This approximation is based on a cooling concept whereby the cold sink is operated a minimum temperature of -140°C. We believe that passive control of front end electronic board temperatures should be sufficient to satisfy these requirements.

Back End Electronics The BEE are not required to operate near the detector plane temperature. As the current design concept has the BEE inside the cryostat, it is however prudent operate the electronics at or below the camera ambient temperature. The conceptual cold plate for the BEE heat extraction is expected to operate in the range of -30 to -40 °C, giving adequate headroom to achieve the desired BEE temperature range with reasonable expectation. Passive control of back end electronic board temperatures should be sufficient to satisfy these requirements.

Set-point, Uniformity, and Drift Requirements These requirements are primarily driven by the QE and noise performance of the detector. Maintenance of the temperature set-point, drift and uniformity band will control the uniformity of the quantum efficiency across the FPA to within 0.5% and within 0.125% across any individual detector. The tight drift parameters are necessary to control the variation in detector QE to within .063% throughout a telescope observing shift.

It is not yet known if active control will be required at the individual detector level or at the raft level. Obviously the complexity would be greatly reduced if it can be controlled at the raft level.

Knowledge Requirements This requirement stems from both control and calibrations issues. For control, a feedback sensitivity of 10% of control range should be adequate for control. For calibration, knowledge at this level should allow QE calibrations at better than 0.05%.

Integrating Structure The Integrating Structure (IS) thermal requirements are driven primarily by engineering requirements derived from the thermal control conceptual design and the structural stability requirements of the FPA. The thermal system concept minimizes heat flow through the IS and thereby minimizes thermal distortion of the focal plane.

Set-point control of the IS has been derived from the minimal thermal distortion requirements. Analysis of conceptual IS mechanical designs indicate the presence of small thermal distortions. One example is shown in Figure 4.3.6-7. Scaling this model to greater thickness indicates thermal distortions of the IS should be less that 0.25 μm for a heat extraction of ~4 Watts (producing a ~4º K temperature distribution). The temperature, uniformity and drift rates should assure thermal distortions do not exceed 10% (± 0.1 μm) of the IS's budgeted flatness.

The knowledge requirements placed on the integrating structure are driven by standard system health monitoring requirements for the camera operations.

Camera Thermal requirements of the camera body, mounts and internals are driven by the telescope seeing requirements as impacted by thermal distortions introduced into the atmosphere within the telescope field of view. Also, condensation on or within the camera drive the thermal control requirements. Based on typical thermal requirements of other large telescopes, the overall camera requirement for thermal load introduced into the field of view is expected to be < 50 Watts. The knowledge and uniformity requirements placed on the camera body are driven by standard system health monitoring for camera operations.

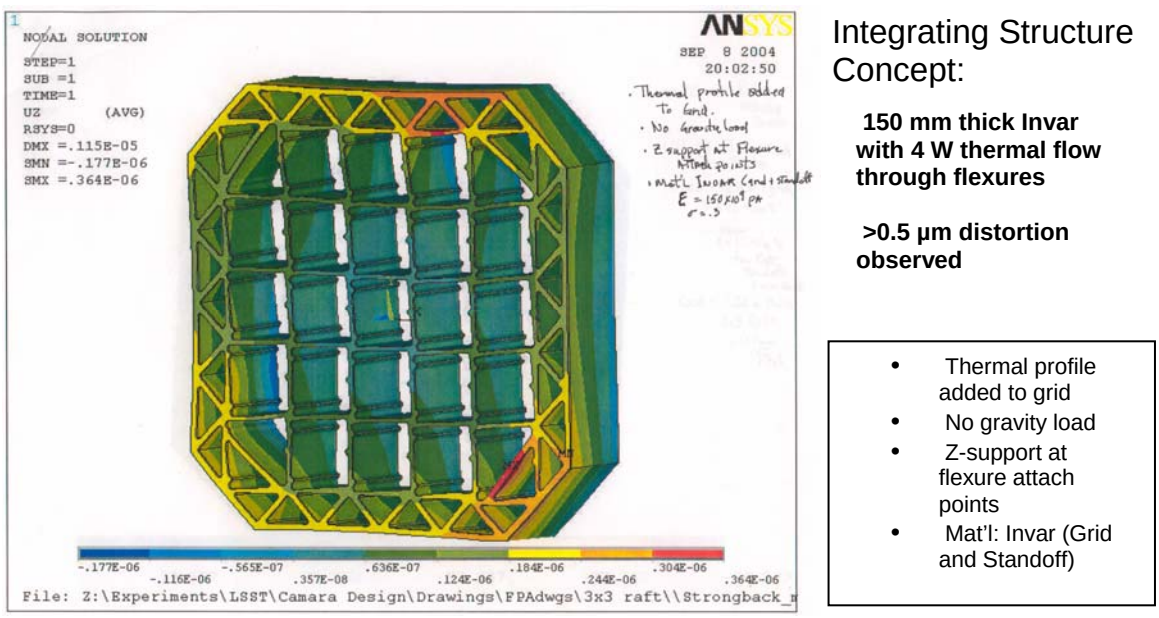
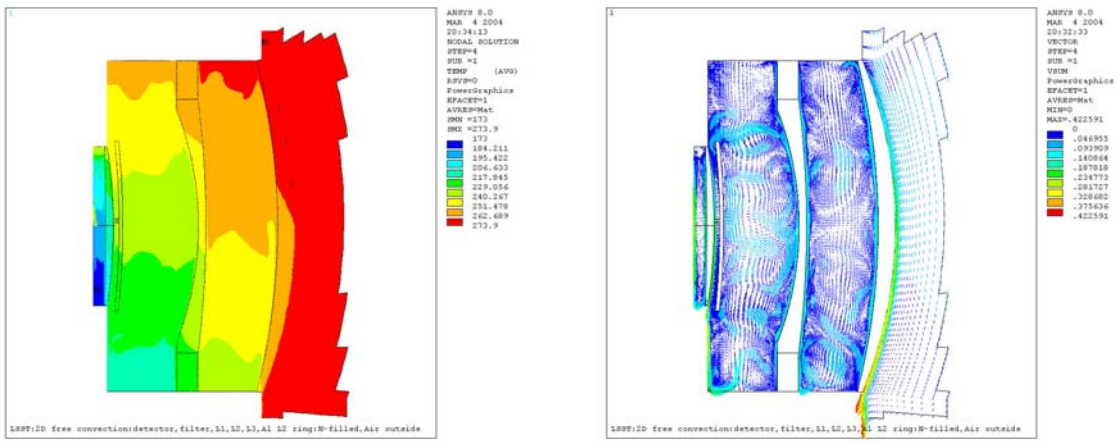


Figure 4.3.6-7 Example of Integrating Structure Thermal Distortion Analysis

Camera Enclosure The purpose of the internal environment requirements of the camera are fourfold; (1) to minimize temperature differences that facilitate heat transfer through the walls of the camera body, (2) to stabilize thermal gradients in the optical elements, and (3) to minimize or eliminate condensation conditions both on the interior and the exterior of the camera, and (4)preventing the slow degradation of optical element coatings.

Environmental conditioning and flow distribution of the gas inside the camera will dominate the control of purposes (1) and (2). Initial purging of the camera internal environment with a dry gas coupled with a responsive conditioning system will satisfy condensation requirements (3). Utilizing a closed conditioning and recirculation system without continuous dehydration will satisfy purpose (4).

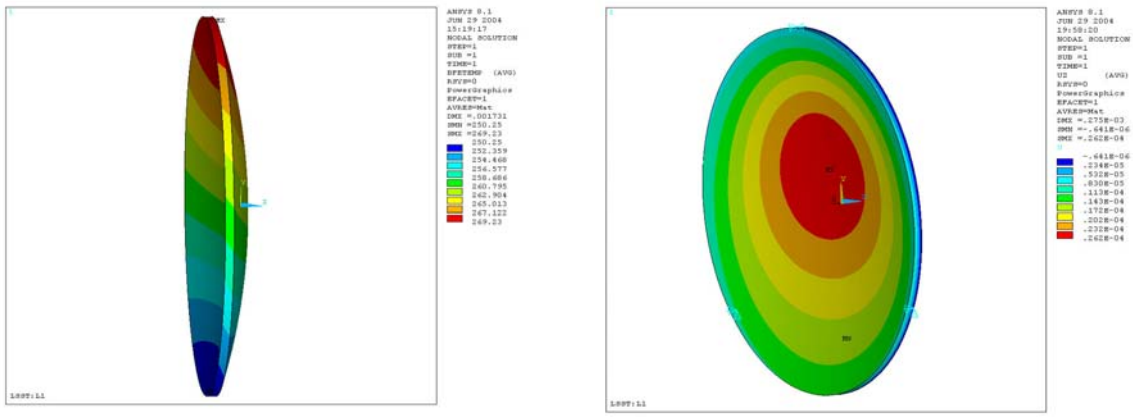
A flow distribution system will most likely be employed to control thermal uniformity within the camera environment as well as across the optical elements. Preliminary work performed to investigate the need for vacuum insulation of the cryostat not only indicated a strong support for a vacuum cryostat, it also makes a strong argument for using forced convection across L3 and an actively controlled temperature environment inside the camera. Figure 4.3.6-8 shows the impact of natural convection on the thermal distribution within the camera environments. Figure 4.3.6-9 show both a thermal distortion and temperature distribution in L1 associated with a camera environment that allows free convection. Controlling the camera environment will also minimize optical distortion in the optical elements due to thermal distortion.



(a) **(b)**

Figure 4.3.6-8 Thermal/Convection Analysis of Camera (a) Thermal profile in L1, L2, L3, Filter, and Dry N_2 gas (b) Convective flow pattern in camera.

Note: This analysis is shown to demonstrate convective effect of N_2 environment of camera; it does not include influence of vacuum cryostat environment. The range of temperature and flow patterns is exaggerated under these conditions.



(a) **(b)**

Figure 4.3.6-9 Thermal/Convection Analysis of Camera - L1 Termperature and Distortion. (a) Temperature profile of L1. (b) Thermal distortion pattern in camera.

Note: This analysis is shown to demonstrate convective effect of N_2 environment of camera; it does not include influence of vacuum cryostat environment. The range of temperature will be slightly exaggerated under these conditions.

Knowledge and uniformity requirements placed on the camera body are driven by standard system health monitoring for camera operations.

4.3.6.3.3 Optical Elements Thermal Requirements

The camera optical elements will be subjected to the requirements compiled in Table 4.3.6-7.

Table 4.3.6-7 Requirements For Temperature/Pressure Control: Optical Elements

Parameter		Units	L1	L2	L3
Control Range	T_{rng}	ºC	15 to -15	25 to -5	25 to -10
Uniformity	ΔT_{un}	ºC	± 5	± 5	± 5
Pressure Differential	Δp	Atm	±.1	±.1	+1.1, -.1

The temperature ranges of the optical elements are primarily driven by environment conditions external to the camera. Uniformity requirements are driven by the maximum allowable optical distortion and focusing requirements of the optics design. As can be seen in Figure 4.3.6-8 through Figure 4.3.6-11, the natural convection inside the camera body has the potential to produce larger than desired thermal distortion and temperature differentials in the optical elements.

Preliminary analysis of L3 distortions resulting from thermal gradients and differential pressures indicate that these requirements may be tolerable without employing extraordinary manufacturing or design features into the optical elements.

4.3.6.3.4 Vacuum & Contamination Requirements

The primary purpose of the vacuum system is to eliminate convective heat transfer in the cryostat and cryostat fluid transfer lines. Because it controls the environment within the cryostat, its design will have a significant impact on the operational contamination control.

Control of contaminants is required to maintain proper operation of components, particularly the imaging detectors. As the detectors are maintained at a lower temperature relative to their surroundings the potential for deposition of outgassed materials is a concern. Engineering controls and design features must be implemented to mitigate and prevent the detrimental effects of contaminants.

Contamination control during the camera assembly and maintenance is largely procedural and deals with the preparation, processing and handling of components. Engineering procedures and processes will be implemented for the processing of components prior to installation into the camera assembly. Likewise, the camera maintenance procedures will be equally focused on the minimization of introduced contamination.

Control of contaminates during camera operations is accomplished by design. The camera design isolates areas within the cryostat sensitive to contamination. As shown in Figure 4.3.6-2, the cryostat is divided into three thermal/vacuum zones. Zone 1, the most critical for contamination, contains the exposed surfaces of the imaging detectors. Creation of the zones is accomplished by placement of barriers on detector modules, raft modules, and integrating structure as indicated in Figure 4.3.6-10 and Figure 4.3.6-11. These barriers prevent and restrict molecular flow between the zones while the cryostat is under vacuum.

Vacuum porting and pumping components also restrict flow between zones. A concept for the vacuum system implementation is shown in Figure 4.3.6-12. It is being evaluated for such isolation of zones. This configuration also allows for removal of the active pumping systems during telescope operations. Disconnection of the active portion of the pumping system will require periodic regeneration of the passive pumping components. The design goal for reconnection of active vacuum system components for regeneration is semiannually

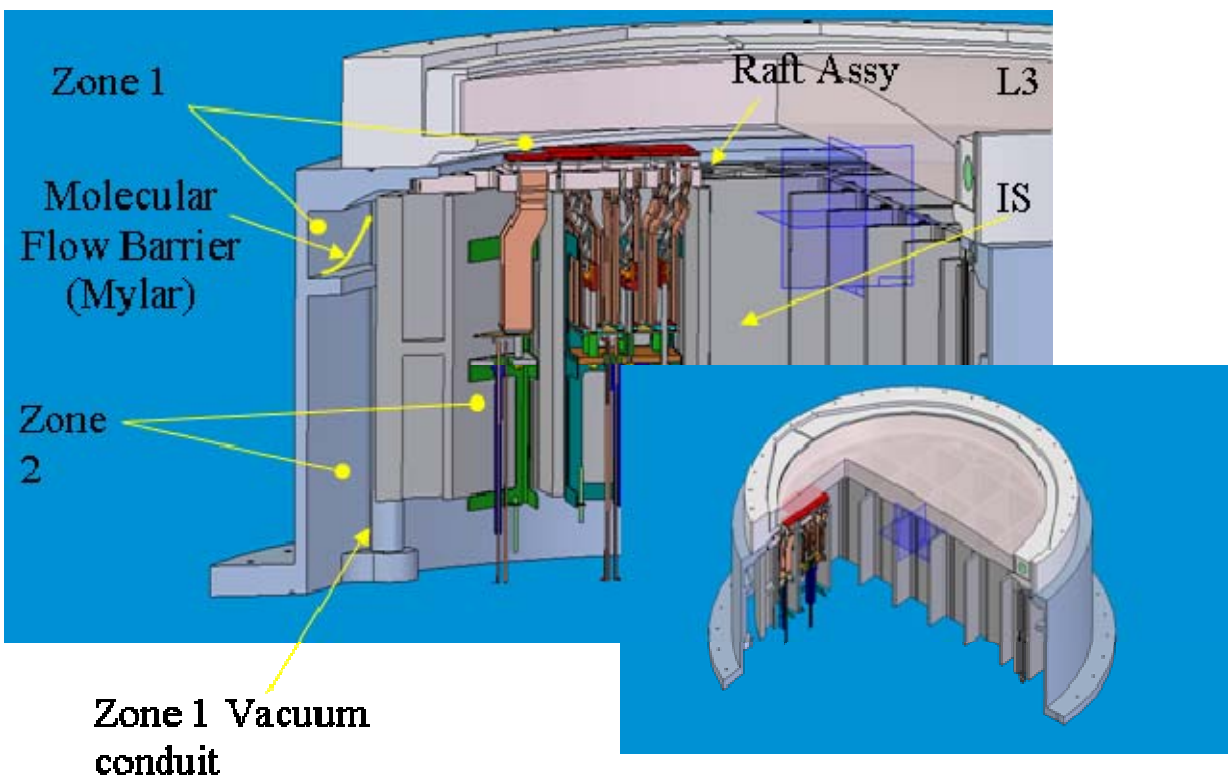


Figure 4.3.6-10 Cryostat Cut-a-way Indicating Vacuum Zone Barriers

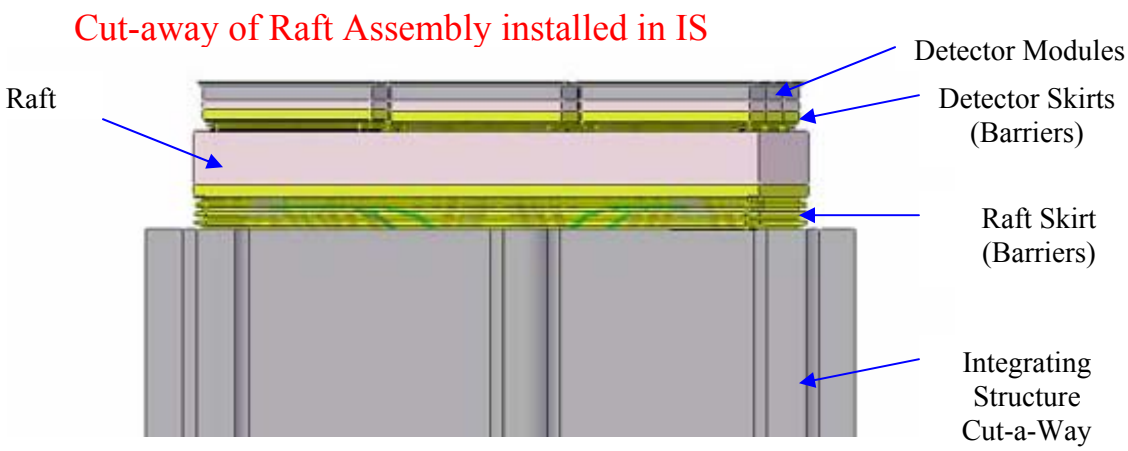


Figure 4.3.6-11 Raft Assembly Cut-a-way indicating Vacuum Zone Barriers

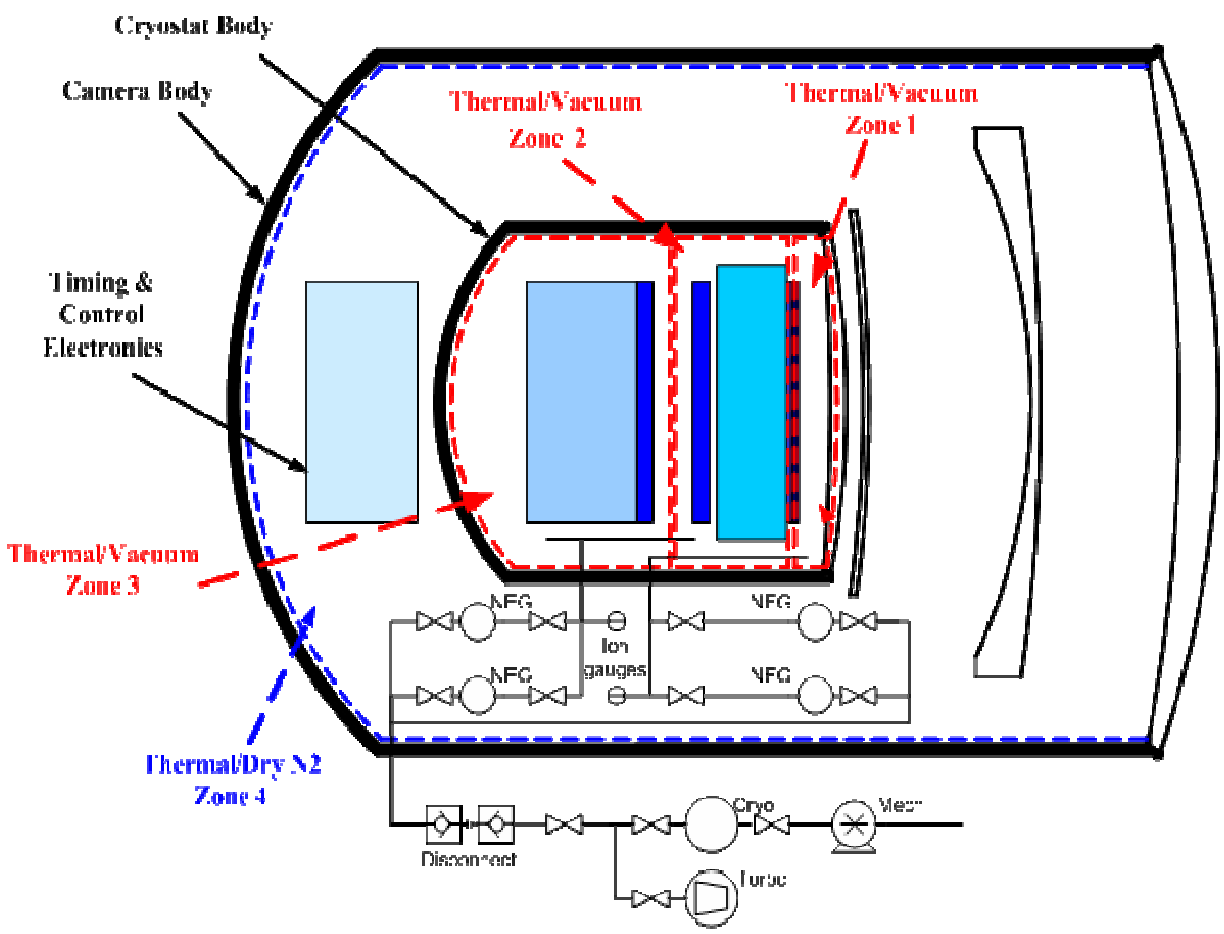


Figure 4.3.6-12 Conceptual Vacuum System Implementation

The basic requirements of the vacuum and contamination controls systems are in Table 4.3.6-8.

Table 4.3.6-8 Requirements for Vacuum & Contamination Control

Parameter		Units	Zone 1	Zone 2	Zone 3
Control Range	P_{Cryo}	Torr	$<10^{-7}$	$<10^{-6}$	$<10^{-6}$

4.3.6.4 Camera Thermal Interfaces

The modular design of the Camera and Inner Cryostat/FPA Assembly minimizes the overall camera and telescope interface. The primary interfaces associated with the thermal design are described in this section.

It is assumed that the camera will be mounted to the telescope structure via a hexapod positioning mount and a camera rotator. The basic parameters for positioning and rotating the camera module are ±90º rotation and ±1cm in piston, and lateral directions. The camera thermal interface must accommodate these actuations.

4.3.6.4.1 Structural

The principle structural interfaces of the thermal system are the cryostat cold plates to cryostat body, and the cooling fluid transfer lines to cryostat body, camera body, telescope mount, telescope structure and observatory.

The **cold plates to cryostat interfaces** support the FEE & BEE electronics and are must be thermally insulated to minimize conduction between the cryostat and cold plates. Structural mounts for the BEE and FEE must not only support the varying gravity loading imposed by the motion of the telescope but also provide an adequate conduction for extraction of heat generated or absorbed by the electronics. No relative motions are required at these interfaces.

The **cooling fluid transfer lines interfaces** cut most system boundaries of the observatory and telescope. Figure 4.3.6-1 indicates the system interfaces and related placement of the transfer lines. Two fluid systems will be employed to extract heat from the camera. To minimize dynamic loading from the cooling systems the base refrigerators for these systems will be remotely located from the telescope. Figure 4.3.6-13 displays the placement of components relative to the telescope, observatory and major interfaces.

All transfer lines will be rigidly mounted to the telescope mount, camera, and observatory hall. Insulated mounts and flex joints will be used to secure and support gravity and dynamic loads as well as thermal expansion of the lines. Penetrations are required to bring the cooling lines through the camera and cryostat bodies. Flexible adapters and seals will also be required at these mounting points.

At the rotating and actuating interfaces of the telescope and camera, both flexible coil and joint will be incorporated to accommodate the wide range of rotations of both the telescope and camera. Accommodating rotations for the vacuum jacketed cooling lines will be much more cumbersome than for the glycol cooling lines. Design studies will need to be performed to best optimize the structural implementation of the rotating interface components. Standard flexible joints will be employed to accommodate the linear actuated motions of the camera and cryostat.

4.3.6.4.2 Thermal Interfaces

The thermal interface to the FPA and inner cryostat electronics module cold plates is a convective fluid coupling built into the cold plates. This interface facilitates the closed loop flow of chilled fluids through the cold plate and returning warmed fluid to a refrigeration system (chiller) in the hall of the observatory. The cooling loop extends through the cryostat, camera body, camera rotator, telescope and telescope rotators to the chiller. The coolant lines for the FPA cooling system will be vacuum jacketed and insulated. The schematic of the chiller system is shown in Figure 4.3.6-13.

Both the glycol and cryo-fluid cooling is used to extract heat from the Camera. The cryo-cooling system only operates on the FPA and FEE whereas the glycol system serves to provide cooling both to the BEE and the camera internal convective environment. Control of the camera internal convective environment minimizes the heat exchange between the camera and the air in the camera field of view. This is a controlled interface utilizing both camera insulation and environmental control.

Care must be given to the design of the transfer lines that traverse the telescope field of view to minimize the obstructed view by keeping the geometry within in the shadow of the support structure of the secondary mirror. In addition to the view obstruction, the heat gain from the transfer must also be held to a minimum. A total heat gain or loss from the camera and support systems must be no more than 50 Watts. All other heat transfer from the chillers and transfer lines will be gathered in the observatory hall and exhausted vie the observatory thermal control system.

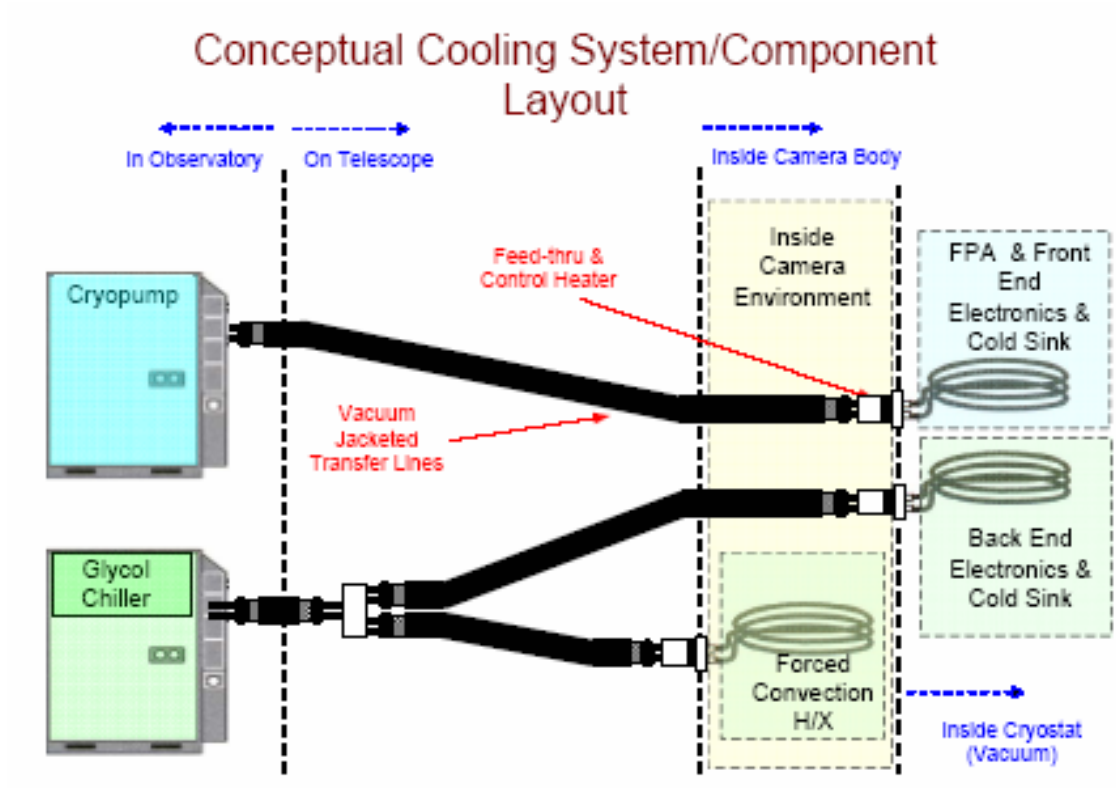


Figure 4.3.6-13 Chiller System

4.3.6.4.3 Electrical

The electrical interface consists of data, control signals, and housekeeping signal. Regardless of the signal type, all interfaces mate through connectors at the inner cryostat electronic module. The thermal conduction from the cold sections of the camera are being closely monitored so as not to introduce excessive thermal loads into the camera.

4.3.6.4.4 Vacuum

The vacuum system interfaces consists of pump-out ports on the cryostat and a passive pumping system on the exterior of the cryostat. The passive system will monitor vacuum zones within the cryostat. Vacuum connections to the exterior of the camera body are planned for vacuum maintenance. Initialization and maintenance of the vacuum system will be performed during telescope maintenance periods using a portable, active vacuum pumping system.

A separate vacuum system will be used to evacuate and create the thermal barrier for the vacuum jacked cooling lines. This system will be located with the chiller systems.

4.3.6.4.5 Cooling Fluids and Purge Gas

Special precautions and fluids will be engineered to eliminate any harmful effect that may result from the use of the cooling fluids and purge gasses. The selection of the compounds is yet to be determined but the design criteria for these will be highly selective so as to minimize risk to scientific performance and equipment.

4.3.6.5 Design Issues

The key thermal design considerations centered on maintenance of the thermal requirements of the FPA. These are:

- Geometric/Thermal uniformity and control of the imaging sensor.
- Geometric/Thermal uniformity and control of the integrating structure.
- Thermal isolation of sensor array, electronics, and integrating structure.

To ensure adequate thermal performance, conceptual design validation is required. The following summarizes the design issues requiring attention.

4.3.6.5.1 Imaging Sensor Geometric Thermal Uniformity

The ± 0.3 ºK uniformity of the thermal sensor requires a creative and well integrated sensor packaging and thermal system design. The scheme being used to achieving this requirement relies on thermal conduction of > 99% of the absorbed and generated heat to be extracted through flexible thermal straps. The temperature of the sensor is modulated and controlled using trim heaters in the thermal straps and temperature sensors on the sensor structure.

The design of the sensor mounting structure has two major functions. First it must provide a geometric stable platform for achieving the ± 5 μm flatness requirement of the sensor. The second function is rapidly extract absorbed and generated heat from the sensor so as to allow for uniform temperature control of the sensor.

4.3.6.5.2 Integrating Structure Thermal Uniformity

The integrating structure is required to maintain ± 1 μm under all orientation and thermal conditions. To assist in meeting this requirement, the thermal uniformity of the entire integration structure is tightly maintained to ± 1 ºC. This will to minimize geometric distortions associated with thermal expansion. Weak thermal connections between the raft assembly and the integration structure will allow the integrating structure to follow the average temperature of the raft assemblies. The thermal uniformity will be control via cooled thermal straps and heaters at the flexural mounts of the integrating structure.

In conjunction with the geometric stability design, thermal properties of joints, flexures, materials, radiation shields and conductors require validation to ensure final sizing and selection of the design. The need for an active thermal control system is yet to be determined.

4.3.6.5.3 Isolation of Thermal Critical Components

The thermal design of the FPA is based on the control of the heat transfer mechanisms. As the FPA will be contained in a vacuum cryostat, the convective heat transfer effect has been eliminated. Conduction and radiation heat transfer is largely control through design. Because of the variability of the thermal properties of the combined material and design geometries, the design will require validation by both analysis and test. Refined parameters for thermal straps, multilayer insulation, joint conductive, FPA radiation absorption, and power generation are needed to determine the extent of active vs. passive control to be used in the overall thermal design.

4.3.6.6 Design Validation

The overall thermal design will be validated through both analysis and test. Concepts for the insulation of critical components will be validated through test wherever thermal and functional properties are of low confidence. This particularly applies to custom designed thermal devices such as multilayer insulation, flexible thermal straps, joints, and complex geometries. Wherever necessary, engineering models will be constructed, tested to expected operational conditions and correlated to analytical models.

Engineering models will be designed and constructed for the integrating structure, sensor assemblies, rafts, conductive elements and thermal shielding. These will be investigated under simulated operating conditions. The investigations will validate design assumptions and demonstrate the viability of system concepts.

4.3.6.7 Integration and Test (I&T)

The CTS I&T activities verify at two levels the system performance. The CTS will initially be assembled with simulated loads and environmental parameters to verify the design performance. Once CTS components performance has been verified the components will be disassembled for integration into the FPA, inner cryostat, camera, telescope and observatory. At appropriate stages in the integration, tests will be design to verify the continued performance and proper installation of the thermal system.

4.3.6.7.1 Optical Sensor/FEE Assembly Thermal I&T

The optical sensor module assembly activities will most likely will undertaken by the sensor manufacture. An integrated test program will be designed to verify the thermal conductivity of each sensor and it's mating FEE and thermal connections. This will ensure the mated products meet the stringent scientific and geometric performance requirements.

With the anticipation that the sensor production will span many months, deliveries will be scheduled so as to allow electronic and raft assembly activities to run concurrently.

4.3.6.7.2 Raft Assembly Thermal I&T

The raft assembly brings together matched sensors module/electronics units, cables, thermal straps, electronic/thermal crate assembly, fixation components, raft frame, and flatness sensors. Each of will have been accepted according to verifying inspection and testing procedures.

The key element of raft assembly is the geometric assembly of the sensors modules, raft frame, fixation, and flatness sensors. Precise alignment of the raft frame, sensors, and fixation devices is require prior to integrating the sensor modules.

As seen in Figure 4.3.6-5, a sensor module with electronics, connectors and thermal straps will be integrated into the raft assembly. Precision alignment of the sensor module and connectivity will be verified prior to integrating additional sensors.

At the completion of the full raft assembly, a complete geometric, thermal and functional verification will be performed prior to being staged for integration of the FPA.

4.3.6.7.3 FPA Thermal I&T

Integration begins with the integration of flexural mounts, thermal sink, actuators and the flatness measurement system. Functional systems checks will be performed throughout the integrations of these components and systems.

As each raft assembly is integrated, thermal joints will be verified to ensure proper assembly into the FPA.

Prior to integration of the FPA into the camera inner cryostat, full functional and environmental verification testing will be performed under simulated camera conditions. Thermal system verification will be an integral part of the FPA test suite.

4.3.6.7.4 Inner Cryostat Thermal I&T

Integration involves assembly of the FPA module, inner cryostat electronics module (includes thermal cold plate), feed-throughs (vacuum, control & data), the L3 assembly and the Inner Cryostat. Functional systems checks will be performed at each step of the integration to verify proper assembly.

At the completion of assembly, a full functional and environmental verification testing will be performed under simulated camera conditions. Thermal system verification will be an integral part of the inner cryostat test suite.

4.3.6.7.5 Camera Module Thermal I&T Testing

Integration involves assembly of the inner cryostat assembly, camera body, shutter module, filter module, L1 and L2 modules, feed-throughs (vacuum, gas services, control, and data) and the camera environmental control system. Functional systems checks will be performed at each step of the integration to verify proper assembly.

At the completion of camera assembly the camera will be operated while connected to support systems including vacuum, fill & purge system, transfer lines, and chiller. Full functional and environmental verification testing will be performed under simulated camera conditions. Thermal system verification will be an integral part of the inner test suite.

4.3.7 Corrector Optics and Color Filters

4.3.7.1 Introduction

The LSST camera optics consist of 3 fused silica lenses with diameters of 1.6 m, 1.1 m and 0.73 m that correct for field aberrations, along with interchangeable filters with diameters of 0.78 m that give spectral coverage from the UV to near IR, in 5 broad bands. The three lenses along with 0.64 m diameter detector array are housed in a canister, ~ 1.6 m in diameter, the size of the largest lens, as shown in Figure 4.3.7-1.

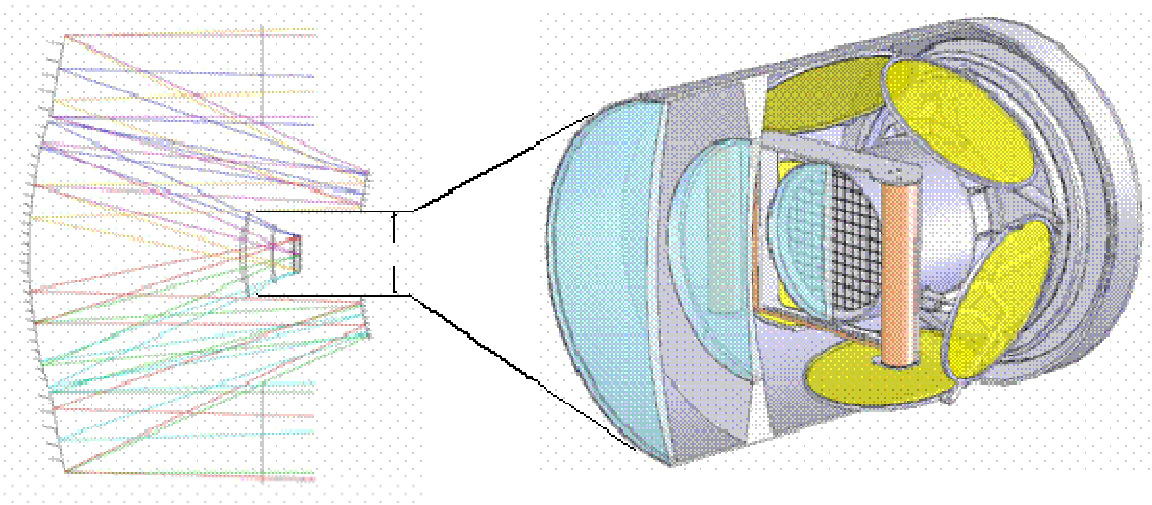


Figure 4.3.7-1: The optical design defines the size and placement of the optical elements and ray bundles define stay-out zones for support hardware. This concept-level design of the LSST camera shows three corrector optics, storage of five color filters, the focal plane array and other camera hardware.

A schematic of the LSST camera optics, including the rays bounding the light distribution incident on the central and peripheral field points, is shown in Figure 4.3.7-2. The largest lens, L1, is nominally 1.6 m in diameter. A larger lens, up to ~1.7 m in diameter, would be required in order to accommodate an atmospheric dispersion compensator (ADC). The current design of L1 calls for an edge thickness of ~3.3 cm and a center thickness of ~6.7 cm. The middle sized lens, L2, has a central thickness of 3.0 cm. The minimum space between L2 and the filters is 30 cm in order to accommodate the filter interchange. The smallest lens, L3, is also the vacuum barrier for the cryostat containing the detector array. There is 2.5 cm between the inner surface of L3 and the focal plane. The central thickness is specified in order to provide a significant safety margin for potential fracture of L3 due to the pressure differential. Empirical data shows that a thickness ratio of ~12 is adequate to provide this safety margin, which yields a thickness of 6.0 cm for the 73 cm diameter lens. The filters consist of multi-layer dielectric interference coatings deposited on fused silica substrates. The baseline design has the first surface of the filters concentric about the chief ray in order to keep the angles of the light rays passing through the filters as uniform as possible over the entire range of field positions. The central thickness and the curvature of the second surface are optimized for image quality. The minimum center thickness is 1.35 cm. Detailed parameters for the camera optics are given in Table 4.3.7-1.

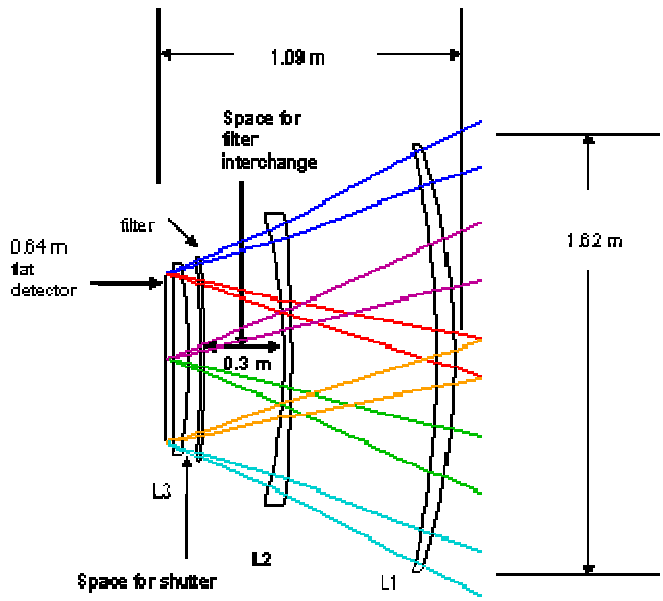


Figure 4.3.7-2 Schematic layout of current LSST camera optics and detector array.

Table 4.3.7-1 Parameter details for the three corrector lenses and (i band) filter.

Property	Units	L1	L2	L3	Filter
Aperture radius	mm	793	514	346	375
Freeboard	mm	17	36	19	15
Outer diameter	mm	1620	1100	730	780
S1 spherical radius	mm	2729	6294	2887	5986
S2 spherical radius	mm	–3722	-2038	plano	-5986
Sag of S1	mm	123.0	24.1	23.2	12.7
Sag of S2	mm	-89.2	-75.6	0.0	-12.7
Sag of centroid	mm	106.1	49.8	11.6	12.7
Center thickness	mm	66.9	30.0	60.0	15.3
Virtual edge thick.	mm	33.1	81.5	36.8	15.3
Actual edge thick.	mm	28.4	74.3	36.8	15.3
Approx. volume	m^3	0.104	0.053	0.020	0.007
Approx. mass	kg	228.1	116.1	44.60	16.1

An early dimensioned drawing of the LSST camera optics is shown in Figure 4.3.7-3. This dimensioned drawing, along with early versions of detailed mechanical drawings of each of the optics, as exemplified in Figure 4.3.7-4, have been sent to commercial vendors in order to obtain preliminary cost estimates for the optics.

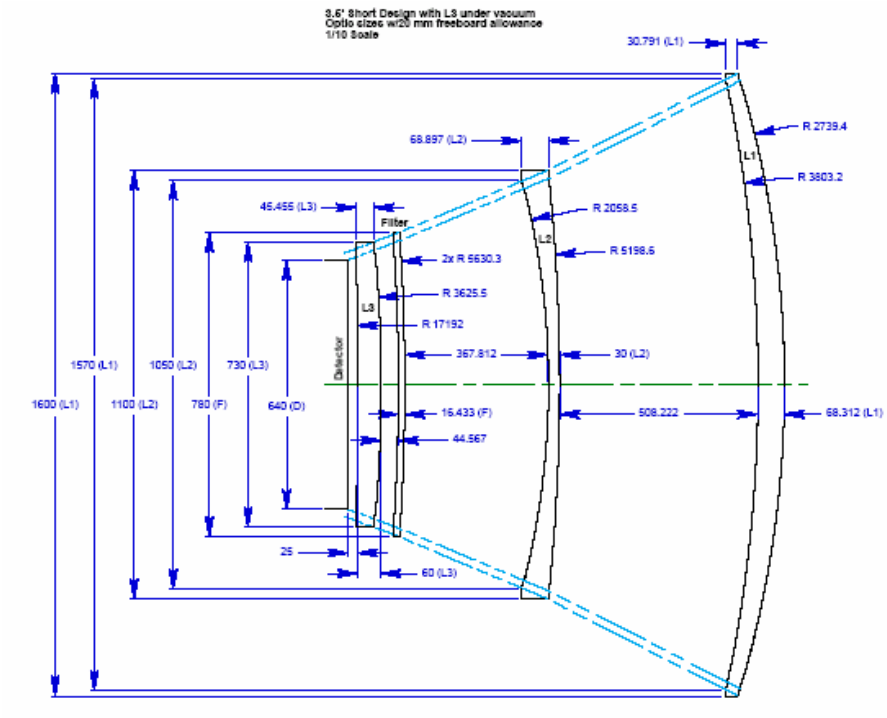


Figure 4.3.7-3 Dimensioned drawing of an early version LSST camera optics and detector array layout.

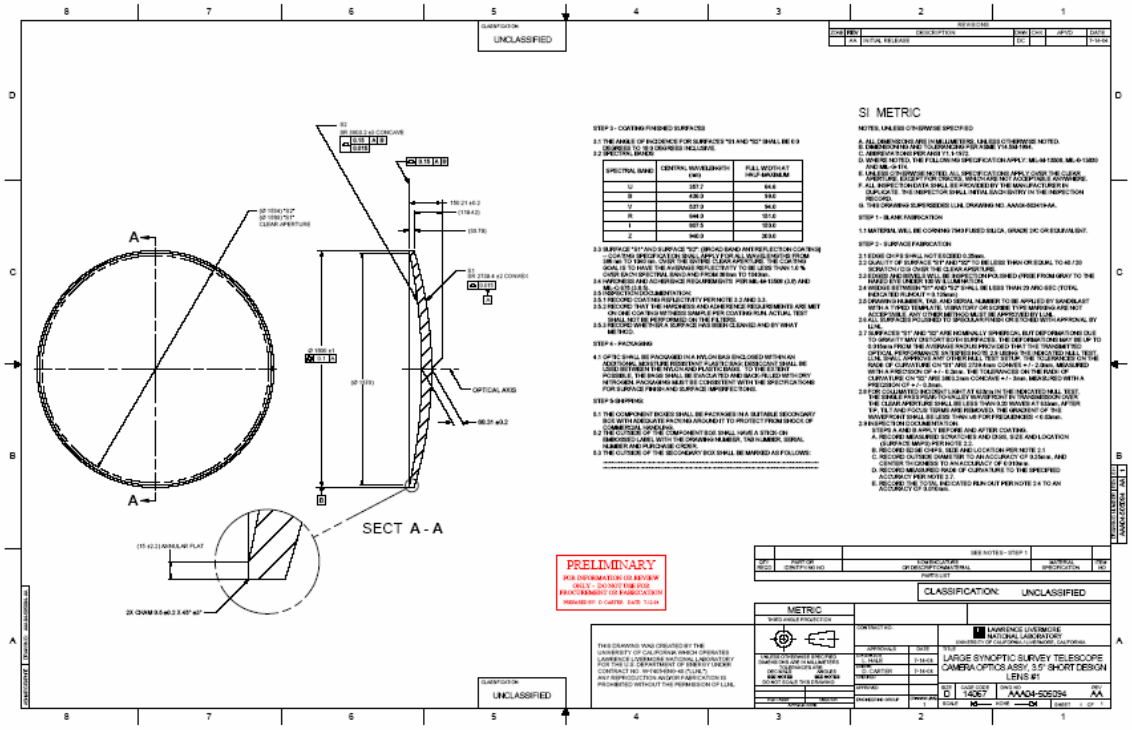


Figure 4.3.7-4 One example of a detailed mechanical drawings that were produced for early versions of each of the LSST camera optics.

4.3.7.2 Null Testing

A key aspect of the corrective camera optics that enables fabrication using industry standard techniques is that they have spherical surfaces and relatively simple null tests. Schematic configurations of the null tests, using a common 1.7 m spherical mirror with a 4 m radius of curvature, are shown in Figure 4.3.7-5. In addition the three-mirror telescope system delivers, without the camera corrector optics, a spherical wavefront on axis that greatly helps in initial assembly and alignment.

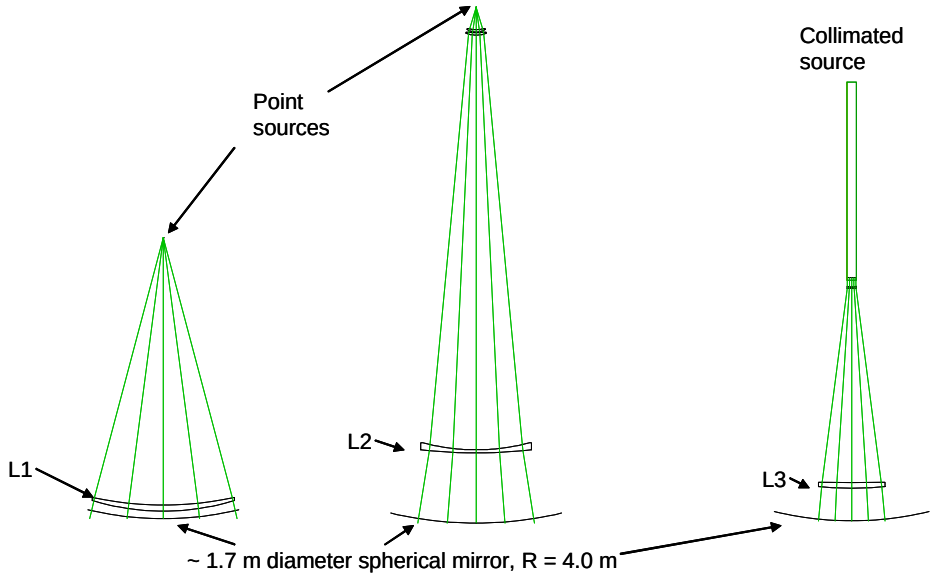


Figure 4.3.7-5 Simple null test configurations for the LSST corrector lenses using a common 1.7 m diameter spherical mirror with a 4 m radius of curvature.

4.3.7.3 Fabrication

The main challenge in the production of the LSST camera optics is the fabrication of large, thin lenses and filter substrates. In order to assess this risk, a team of LSST representatives has visited multiple commercial vendors, supplied these vendors with documentation on the baseline optics designs as described above, and initiated discussions with the vendors concerning the specifications, cost, schedule and technical risk of these optics. The preliminary feedback from all vendors indicates that commercial costs and schedules are consistent with LSST budgetary and program planning estimates. Furthermore, the responses from multiple commercial vendors demonstrate that a substantial industrial base exists for fabricating large, thin optics. An example of a 2 m class large, thin optic produced by U.S. industry is shown in Figure 4.3.7-6.

Figure 4.3.7-6 Photo of a large (2 m class) conformal window produced by commercial U.S. vendors.

Based on the initial assessment of commercial vendor capabilities, the following plan for completing the LSST optics procurement has been formulated.

1. Complete camera optics design study
2. Finalize baseline design
3. Finalize baseline null test procedure
4. Prepare RFQ for camera optics
5. Iterate design specifications with vendors
6. Define acceptance criteria and test procedures

In addition to these steps, we are also considering issuing an RFQ for the complete camera opto-mechanical assembly. This idea is based on feedback from multiple commercial vendors indicating interest in bidding on the complete assembly. A related consideration would be to issue contracts for opto-mechanical engineering studies to commercial vendors.

Once the lenses and filter substrates are fabricated, they must be coated. For the lenses, the coating is a relatively straightforward broad-band anti-reflection coating to minimize the optical loss through the system. The main challenge here is the size of the largest lenses. In the case of the filters, the coating is a relatively sophisticated multi-layer interference coating that is designed to transmit only light with wavelengths in a specified band, and to reject light at other wavelengths with a specified fidelity. The main challenge here is to deposit uniform coatings with the desired characteristics on the large, curved substrates.

4.3.7.4 Filter Set

The current LSST baseline design uses a filter set comprised of standard astronomical u, g, r, i, z, and Y bands. The goal of 1% relative photometry in LSST images defines the general feature of the LSST filter set and is detailed below.

Approximate FWHM transmission points are shown in Table 4.3.7-2.

Table 4.3.7-2 Baseline LSST filter band-pass FWHM points.

Filter	λ_1	λ_2
u	330	400
g	402	552
r	552	691
i	691	818
z	818	922
Y	970	1028

LSST Filter General Feature:

No gaps should exist between filter band-passes, except in the spectral region between 930-960 nm where it is desirable to exclude a variable water vapor absorption feature.

Filter band-passes should not overlap.

Band-pass throughput should be as high possible

The transition between stop and pass band should be less than 10% of the filter band-pass.

The final specification of the filter characteristics will be determined through detailed science modeling.

KG will add some more here…

4.3.8 Controls and DAQ

4.3.8.1 Design Approach and Logical Structure

The camera comprises the last stages of the optical system, the sensors to convert light to digitized images, and the hardware and software to transfer the images into the LSST data management system. The Camera Control System (CCS) will manage the activities of the various camera subsystems and coordinate those activities with the Observatory Control System (OCS, section 3.2.6). The CCS comprises a set of modules (nominally implemented in software) which are typically responsible for managing one camera subsystem. In this section, a logical (or symbolic) view of the CCS is presented followed by discussion of the various camera subsystems and their corresponding control modules. Finally a short discussion of the data acquisition system and the interface to data management is presented.

The approach we are following parallels the current OCS model described in Section 4.1.6. We assume that a single control module will manage a camera subsystem. Generally, a control module will be a long lived “server” process running on an embedded computer in the subsystem. Multiple control modules may run on a single computer or a module may be implemented in “firmware” on a subsystem. In any case, control modules must exchange messages and status data with a central control module (CCM). The main features of this approach are:

1. Control is distributed to the local subsystem level where possible and time critical loops are closed at the local level
2. The systems follow a "Master/Slave" strategy with one CCS module (CCM) acting as the master
3. Coordination will be achieved by the exchange of messages through the interfaces between the CCS and its subsystems.

A possible logical view of this architecture is shown in Figure 4.3.8-1 below:

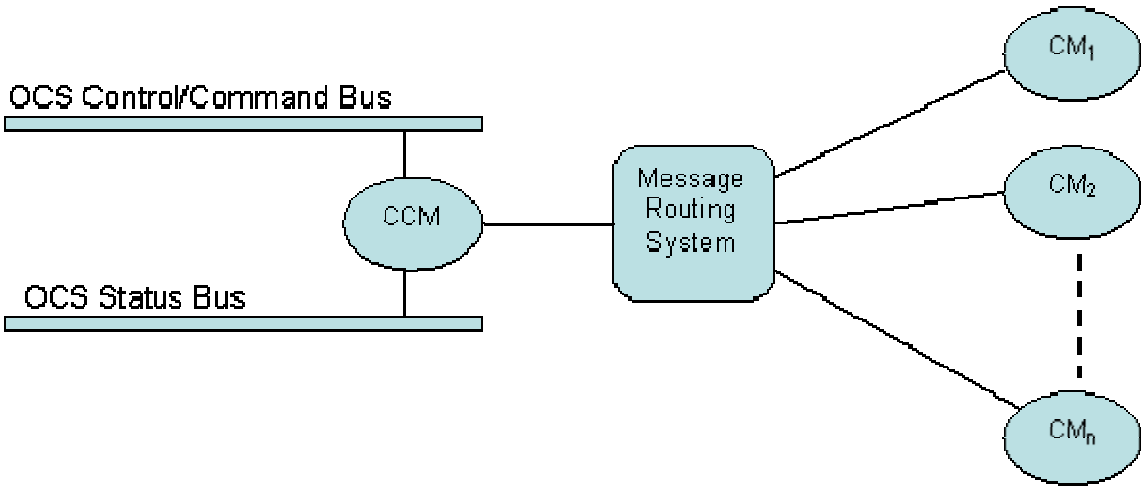


Figure 4.3.8-1 Logical view of Camera Control System modules and communications. The CCM is the master control module. Each of the control modules labeled CM_j is responsible for managing a single camera subsystem. The message routing system transports control and status data messages between the control modules.

As shown in the figure, the Central Control Module (CCM) process acts as the interface to the OCS, receiving commands as well as publishing camera status data and subscribing to selected observatory status data. Within the CCS, the CCM process must schedule tasks and oversee all camera activities. When a task is requested by the OCS (eg. a 10 second exposure), the OCS must sequence messages including appropriate handshaking, prerequisite condition and error checking to each subsystem control module. Additionally, the CCM must continue "housekeeping" duties such as monitoring subsystem status, the status of selected observatory parameters, and publishing camera status to the OCS status system. The "Message Routing System" in the figure represents the collection of hardware and software mechanisms which may be used to exchange messages such as Ethernet, USB, fifos, shared memory, etc.

4.3.8.2 Camera Subsystems and Control Modules

To apply the logical representation to control of actual hardware components of subsystems first consider the types of components that are involved. These subsystems will have motors (shutter, filter), encoders, pressure sensors, temperature sensors, limit switches, voltage monitors, relays and other electro/mechanical devices. Control of these components will require multi-axis motion controllers, digital i/o interfaces, multi-channel A/D interfaces and other specialized computer interfaces.

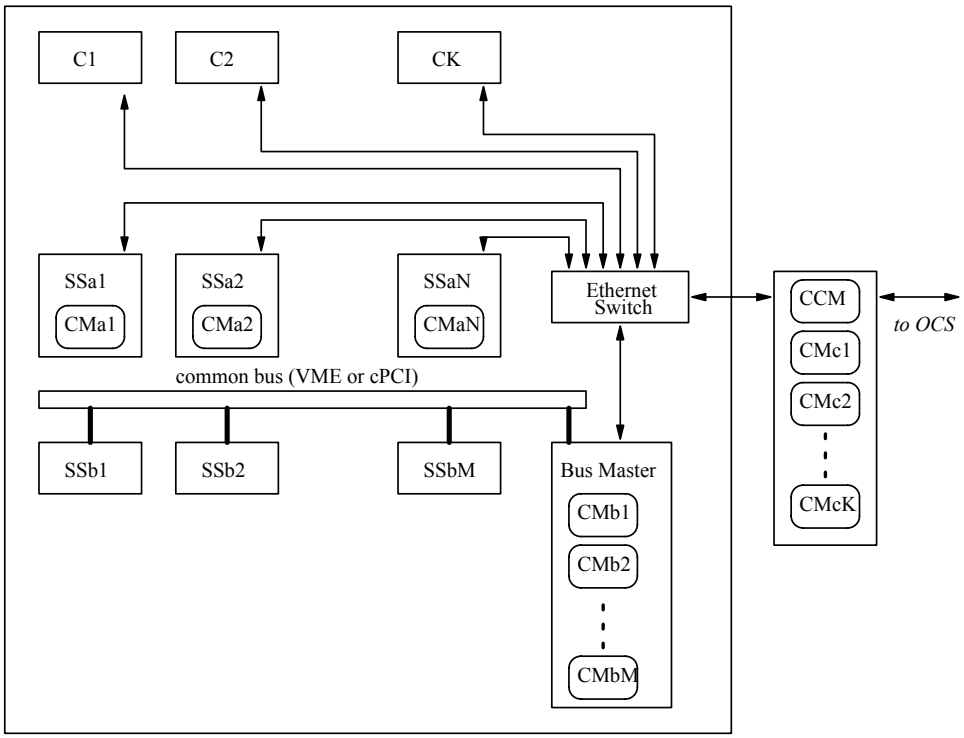


Figure 4.3.8-2 Model architecture for camera control. The top row shows Ethernet attached components, (C_j's), with control modules (CMc's) running on the same host as the CCM (right). The middle row shows control modules (CMa's) running on processors embedded in the subsystems labeled SSa_j. The bottom row illustrates subsystems (SSb's) on a common bus with control modules (CMb's) running on the bus master.

A model architecture for this system is shown in figure 3.3.8-2. Three arrangements between control modules and subsystems are illustrated. Some subsystems are assumed to be self controlled, with an embedded processor exchanging messages with the CCM via Ethernet. These systems are labeled SSa1…SSaN. That is, the control module software (CMa's) for the subsystem runs on an embedded processor in the subsystem. Several subsystems may share a single bus system such as VME or compact PCI with a single Ethernet connection to the bus master processor. These are labeled SSb1 through SSbM in the diagram. Several control module processes (CMb's) would then run on the bus master. Other subsystems may be controlled by digital i/o lines or motion controllers from the CCM host system through Ethernet attached components (C_j's). In that case, the control module software for the subsystem (CMc's) would run on the CCM host and the message exchange with the CCM would be internal to the CCM host. The actual implementation of this architecture is dependent on the final design of the individual subsystems. The architecture is intended reduce subsystem dependencies, and to simplify system integration, upgrades and maintenance of the subsystems.

Each of the currently defined subsystems is listed below.

4.3.8.2.1 Science Array Subsystem (SAS)

This subsystem comprises the science image sensors (201 CCDs), the front end electronics, the back end electronics, and a collection of electronics boards on a common bus along with the power supplies, cabling, fiber outputs, etc. needed to operate the CCDs in the science array. We assume the 201 CCDs fully populate 21 focal plane rafts with 9 CCDs each and partially populate 4 rafts with 3 CCDs each (figure 3.3.3-1). A dedicated state machine based controller issues timing and control signals to Camera Electronics. Specifically, it controls all signals required by

both Front End Cards and Back End Cards. These include all CCD clocks, Signal Processing ASIC timing/control signals, A to D conversion control signals on the Back End, and so forth. Note that sensor data is normally not returned to the SAC, but is routed out of the camera on multiple Data Fibers. The control module for this subsystem would run on an embedded processor providing an interface between the CCM and the dedicated state machine based controller.

4.3.8.2.2 Science Array Data Acquisition Subsystem (SDS)

This subsystem comprises 25 computers (one per raft, not including spares), each hosting a fiber interface card to receive the image data generated by a raft. These computers would buffer the image data in memory and transfer the data into the Data Management stream. Some functions of data management, (eg. initial data quality assessment), may also be carried out on these hosts.

4.3.8.2.3 Wavefront Sensor Subsystem (WFS)

This subsystem comprises the wavefront sensors, the inner Dewar electronics, and a collection of electronics boards on a common bus along with the power supplies, cabling, etc. needed to operate the sensors in the wave front array. This system may share its common bus and electronics architecture with the SAS. The reason for this is that the wavefront sensing element may well be smaller versions of the same silicon technology as the Science Array Sensors.

4.3.8.2.4 Wavefront Sensor Data Acquisition Subsystem (WDS)

This subsystem comprises several computers, each hosting a fiber interface card to receive wavefront sensor data. These data would then be transmitted to both the wavefront sensor analysis system (part of the telescope control system (TCS)) and to the DM system. The latencies involved are still being evaluated, but clearly the transfer to the TCS will require low latency.

4.3.8.2.5 Guide Sensor Subsystem (GSS)

This subsystem comprises a set of guide sensors and associated electronics used to monitor the motion of stellar psf centroids in the LSST focal plane. While the architecture, number, and size of these sensors are TBD, a possible implementation would be frame-transfer CCDs arranged around the outside of the science array with sufficient coverage to guarantee several guide stars in any field. This system will operate at ~100 frames/sec and will likely not share components with either SAS or WFS.

4.3.8.2.6 Guide Sensor Analysis Subsystem (GAS)

This subsystem is a computer or set of computers which receive and analyze the guide sensor image stream to generate (x, y, θz) offset parameters appropriate for use by the focal plane actuation unit and telescope tracking systems. This subsystem would likely reside off the telescope receiving data from the GSS over optical fiber. High frequency offset parameters would be used to drive the focal plane actuation unit while secular terms would be forwarded to the telescope tracking system.

4.3.8.2.7 Focal Plane Actuation Unit (FPU)

This unit comprises a motion controller and any other sensors needed to control the position of the focal plane based on input (x, y, θz) offset parameters. Section 3.3.4.5 describes a possible implementation of the image stabilization system.

4.3.8.2.8 Thermal Zone 1 Subsystem (T1S)

This subsystem comprises the devices and controllers necessary for monitoring and control of the temperature of the focal plane sensors and front end electronics. This zone must be maintained at approximately -100C to high accuracy (to be specified). Its function is to remove heat from the focal plane due to radiation through L1 as well as heat dissipated in the CCDs and Front End Cards. (See Figure 3.3.3)

4.3.8.2.9 Thermal Zone 2 Subsystem (T2S)

This subsystem comprises the devices and controllers necessary for thermal monitoring and control of the Back End Electronics attached to the bottom plate of the Inner Dewar. The temperature will be in the neighborhood of -20C but does not require high accuracy or stability. Its function is to remove heat generated by the Back End Electronics.

4.3.8.2.10 Thermal Zone 3 Unit (T3U)

This unit monitors and controls temperature in the Timing/Control Crate. Specifically, it removes heat generated by electronics residing in the Crate and monitors crate temperature.

4.3.8.2.11 Thermal Zone 4 Unit (T4U)

This unit monitors and controls temperature of the atmosphere, likely dry air or dry nitrogen, contained in the outer Dewar.

4.3.8.2.12 Thermal Zone 5 Unit (T5U)

This unit monitors and controls temperature in the External Services Crate. Specifically, it removes heat generated by electronics residing in the Crate and monitors crate temperature. (Figure 3.3.3-1)

4.3.8.2.13 Raft Alignment Subsystem (RAS)

This subsystem comprises actuators in the Raft Kinematic Mounts and the associated devices and controllers needed to apply signals and/or voltage levels which set or maintain the alignment of the rafts.

4.3.8.2.14 Shutter Controller Unit (SCU)

This unit is responsible for controlling and monitoring the shutter position and motion. The shutter position as a function of time will be monitored by encoders and included in observational meta-data.

4.3.8.2.15 Filter Controller Subsystem (FCS)

This subsystem is responsible for control and monitoring of the Filter Exchange Mechanism.

4.3.8.2.16 Vacuum Control Subsystem (VCS)

This subsystem comprises vacuum monitoring sensors, pumps, and other vacuum components.

4.3.8.2.17 L2 Controller Unit (L2U)

Responsible for motion control of L2 for compensation of different filters

4.3.8.2.18 Power and Signal Controller Unit (PSU)

This unit will control power (AC and DC) to the other subsystems and may provide other signal interfaces for digital and analog i/o (eg. voltage monitoring).

4.3.8.3 DAQ and interface to Data Management

The LSST data management system Infrastructure is divided in to 3 segments: mountain, base, and archive. A full discussion of the mountain/base/archive data management architecture appears in section 3.4.5. In short, the mountain/base segments will be operated as a tightly coupled system via LSST owned high speed links between them. The bulk of data quality assessment and real-time data analysis will take place in the base system. The archive center(s) will support the full data analysis (and re-analyses), archiving of images and data products, and data services to the user community.

The LSST Observatory on-site computing system must contain sufficient functionality to continue operating the observatory in the event of a network outage lasting several days. The main computing functions, data analysis, archiving and data base management will be located at the "base camp" and archive sites. The 3 sites will be connected via high speed networking. Additionally, on-site system must provide support for engineering and maintenance modes (eg. calibration, system characterization, etc.). Note that the on-site computing system would not provide real-time alerts which require nearly the entire analysis architecture; hence if the high speed network to the base is down, alerts will not be generated. The functions of the mountain computing facilities include:

1. Data Acquisition Nodes providing the interface between the camera data stream and the data management network
2. Storage: 1-3 nights raw data (20-60 TB)
3. Data quality assessment backup system to be used during commissioning and to allow critical DQA operations during network outages. This may be a full analysis of a small percentage of data or a set of basic measurements on all data.
4. Observatory control & status system (including engineering support)
5. Camera control systems
6. Telescope control systems.

The data from the Science array will be transmitted over fiber optic cables with one fiber per "raft" of 9 CCDs. Additional camera data from the wavefront sensors and optionally from the guidance system will also be transmitted directly to data acquisition nodes over optical fiber. The wavefront sensor data must be made available to the telescope control system (TCS) in real-time as well as being archived by the DM system. The science array data will be transferred to the DM chain for archiving, analysis, and quality assessment. Some portion of the quality assessment will be feed-back in real-time for use by the OCS in running the observatory. Some fraction of these real-time QA data must be generated during observatory operation with latencies less than the exposure cadence to allow partial assessment of data quality before the telescope is re-pointed. This and the latency requirements for astronomical alerts determine the transfer rate of data from the camera to the DAQ nodes. Assuming that the data can be transferred in 2 seconds from the camera to the DAQ nodes we require a transfer rate of 3 Gbps per node. Each DAQ node would receive 288 channels of data (1MB/channel) which would initially be stored in a local memory cache. The DAQ nodes would then inject the stored data into the data management stream. For the nominal observing cadence of 2 images every 30 seconds, the data stream between the mountain and base sites will be approximately 4 Gbps. This rate is consistent with the expected network capabilities available when LSST begins operations. Further discussion of the LSST data stream appears in section 3.4.5.

4.3.9 Camera Integration and Testing

In this section we describe the overall integration and testing sequence for fabrication of the LSST camera, incorporating the components and mechanisms as described in detail in the above sections. The ordering of this section generally follows the order in which assembly would be

performed, not necessarily the order in which testing is conducted. The camera strawman design is modular enough so that preparation of major sub-assemblies can be performed in parallel, tested and then integrated together with no single parts on the critical path for completion of the overall assembly.

As described in detail above, the camera is a nested cylinder geometry, with an outer camera enclosure vessel that has as one end which mechanically holds the first refractive lens L1 and the other end the main camera access flange with electrical and mechanical feedthroughs.. Service access ports on the main body vessel sides will be available to allow routine maintenance of the inner dewar for components not requiring inner camera removal. The inner dewar will need to be removed for replacement of modular parts upon failure: Signal chain ASICs, Front-end boards and data cards (all within the front-end electronics), sensor rafts, back-end electronic boards and wavefront sensors. Failure of less modular components will also require removal of the inner dewar but with potentially extended downtime: For example, focal plane articulation actuators, thermal sensors and temperature regulation devices will require the whole camera to be removed from the telewcope for servicing. Failure of components within the camera vessel but outside of the inner dewar are generally modular, i.e., the shutter, the filter exchange mechanism, the filter selection "revolver" and fine focus adjust actuators on L2), but access may be significantly impeded without removal of the inner dewar unless designed with these stipulations in mind.

Our current assumptions are that the inner dewar will be mounted to the inside surface of the main camera access flange via a hexapod assembly which will facilitate exactly constrained 3-axis translation plus 3-axis rotation of the inner dewar optical elements (L3 and the focal plane array) . The process of removing the inner dewar from the camera vessel will be extremely delicate (given the spacing tolerances between optical surfaces). The hexapod may therefore be used to retract the inner dewar (and L3 surface) away from the shutter and filter exchange mechanisms prior to extraction, and similarly, it will be used to extend the inner dewar to operational position after insertion. No cabling will pass directly from the inner dewar or the main camera access flange to mechanisms, monitors or control devices inside of the main vessel but outside of the inner dewar: This way, the main camera access flange and inner dewar can be extracted as a whole immediately after performing the necessary cable disconnections on the outside of the camera assembly and breaking the access flange seal.

4.3.9.1 Integration and testing of the sub-assemblies

The inner dewar is a cryostat which contains the mechanical, thermal and electronic and silicon detector components. It will be assembled such that it will be mounted as a unit onto the main camera body access flange by a hexapod mechanism.

The cryostat operates at very high vacuum (~10^{-8} Torr), made necessary by the operating requirements of the focal plane environment. The third lens (L3) provides the vacuum barrier at one end of thecryostat, while a flexible bellows, a residual gas analyzer (RGA) and vacuum valving will be on the vacuum side between the cryostat interior and the main access flange of the camera body. Titanium sublimation pumps and ion pumps, situated inside of the cryostat, will be shielded from the optical instrumentation by optical labyrinths to minimize light leakage. The proximity to the focal plane of the large surface area of electronics within the cryostat requires running the electronics at lower temperature then ambient. To facilitate this, the cryostat is partitioned by plates that delimit three different "vacuum zones".

4.3.9.1.1 Vacuum Zone 1 (Camera Lens L3 Vacuum Barrier to Sensor Surfaces)

The most sensitive volume has positive pressure feeds to control flow direction during venting. The lowest temperature in this zone will be the operating temperature of the cooling system. The detector temperature will also approach this temperature (~-100C).

4.3.9.1.2 Vacuum Zone 2 (Integrating Structure interior)

This region contains all of the front end electronics, their cables and ASICs, the raft backside surfaces and the integrating structure. It also contains the cold plate used in drawing heat from zone 1. The lowest temperature in this zone will be the cold plate, followed by the front-end electronics (including signal chain ASICs, front-end boards and raft backplane electronics), which are held colder than the sensor surfaces by design. Vacuum zones 1 and 2 make up the thermal zone 1.

4.3.9.1.3 Vacuum Zone 3 (Back-end Electronics, Dewar door and Feedthroughs)

With the cable count reduced substantially between the front-end and back-end electronics, these cable bundles corresponding to individual rafts pass between zones 2 and 3. Zone 3 contains the back-end electronics (pixel digitization, timing, data buffering and optical fiber drivers) as well as another cold plate. Cabling passing between these zones will connect via high density feedthrough connectors on a bulkhead. The back-end electronics are located close to the end of the cryostat supported on the outside by a hexapod structure. Vacuum zone 3 corresponds to thermal zone 2, and operates at a substantially higher temperature than thermal zone 1. It is therefore necessary for vacuum zone 3 to be effectively isolated from the two other vacuum zones within the same inner dewar, to contain contamination of the colder surfaces, particularly the focal plane of sensor arrays, by outgassing of materials from warmer surfaces.

4.3.9.2 Electrical Feeds into Inner Dewar

A variety of mechanical and electrical ports pass through the wall of the cryostat. Generally there are six or seven categories of these ports. Science data will be transmitted via optical fiber cable bundles. Digitized, buffered engineering housekeeping status data will be formatted into packet structure, possibly carried by internet protocol and fed to the outside by either optical fiber or twisted pair. Real time commanding triggers of the data acquisition sequence will be transmitted by twisted pair conductors. Real time focal plane articulation signal will be in the form of analog (servo) DC voltage. Heater element currents will be transmitted through the cryostat wall directly by current feeds. Real time data quality feedback on time scales shorter than individual exposures (the fine guide sensor) will be transmitted either by an intensified analog signal (NTSC) or by a high speed digital link. Wavefront sensors on the other hand, may be read out synchronously with science exposure frames, and may share the same optical fiber lines

4.3.9.3 Focal Plane Articulation (either Integrating Structure, Raft, or Sensor basis)

Wind buffeting of the telescope structure (primarily by wind impulses on the secondary mirror) and the resulting image motion on short time scales will be measured by the fine guide sensor imaging devices (read out at ~10-100Hz rate). The composite drift signal from image analysis will be minimized by two dimensional articulation of the focal plane, with ~100ms combined (image analysis plus articulation command) latency and maximum travel of +/- 100 microns. Maximum acceleration to the focal plane will therefore be on the order of ~0.1g. The current baseline assumes that the focal plane will be articulated by actuating the entire integrating structure; however, given the mass of the assembled integrating structure (~100 kg) and potential for driving oscillations within the camera and/or telescope, alternative articulation schemes (individual rafts or individual sensor packages) are also under consideration. See Section XX for details.

Under the current baseline, actuation of the integrating structure (which carries the focal plane) is achieved within the attaching mechanism that joins it to the housing via points on the inside wall of the cryostat. Electrical connections for this actuation will pass from the outside through

the cryostat wall, so that assembly and disassembly will be simplified. Furthermore, the section of the cryostat that houses the integrating structure will resemble a ring attachment that mates and seals with adjacent sections, i.e., the section that carries the L3 lens and the section that provides the cold plate that cools the focal plane.

4.3.9.4 Focal Plane Assembly

The assembly of the focal plane is planned to have several levels of integration and is assembled in a hierarchical manner onto the integrating structure using 25 identical CCD "rafts" containing 9 identical sensor packages each. The geometric integrity of the focal plane will be maintained by one of several approaches under study. One approach requires the integrating structure and the mounting of the rafts to maintain the flat geometry over the range of expected static and dynamic loads, and also limit thermal variations to within these tolerances. This will be possible only through thorough integrating structure design, appropriate choice of materials and raft mounting design. The alternative approach is to implement an *in situ* focal plane geometry verification facility into the focal plane assembly, by which any changes in raft position and orientation can be sensed on the scale of the required tolerance. Focal plane sag and thermal distortion would be calibrated for the variety of expected environmental conditions using the *in situ* alignment verification, and using the calibration information, the three actuators on each raft will be used to maintain the flat focal plane geometry during operation. A practical, safer compromise will be to design for a stable thermo-mechanical focal plane, but to also implement the *in situ* alignment instruments to provide focal plane geometry feedback, primarily during the integration and testing, and again when focal plane servicing is necessary. Different *in situ* verification systems are under consideration, and these include capacitive edge sensors (between raft and raft, between raft and integrating structure) and also a grid of orthogonal laser beams (each forming an optical straightedge) whose positions with respect to each raft's internal coordinate system would be read out using thin beam splitters paired with imaging detectors (4 per raft). These internal alignment systems would be considered carefully as research and development activities as viable built-in geometry verification facilities.

See Section 4.3.3 for details of these R&D efforts

4.3.9.4.1 Rafts (Vacuum Zone 2)

The rafts, with the 3x3 array of sensors, are approximately 12cm on a side. With cabling, electronics and thermal straps contained in a "raft crate", the third dimension of the raft assembly is roughly 20cm. The sensor end of the raft will reside in vacuum zone 1, whereas the remainder resides in vacuum zone 2. The cabling, straps and electronics will be fed into the open "bays" of the integrating structure from the camera aperture side and the raft crates will be secured along the sides of the integrating structure's "bays". The attachment points of the raft backplane will be secured to the integrating structure's interface points.

4.3.9.4.2 Kinematic Attachment and Actuation Mechanisms

The rafts are individually actuated in three degrees of freedom (providing piston, tilt and tilt) while imparting negligible stress to the raft backplane (which would in turn distort the sensor surfaces). Flextured interface schemes, both for the piezoelectric stage and for the attachments between stage and raft backplane, and between stage and attachment sheme, provide practical solutions (TBC). Actuation and feedback readout of these mechanisms must be functional at the operational temperature of the camera focal plane.

4.3.9.4.3 Sensor Packages

The sensor packages, provided by the vendor to required flatness specifications, will need to be positioned to co-planar configuration with the other sensor packages residing on each raft. In order to not stress the sensors or their substrates, a kinematic mounting will be required with adjustments to the attachment points made by polishing those interfaces down to the correct values. The sensor packages will then be attached to the raft backplane by this stress-free interface. Each sensor will be attached using two identical high density connectors and a thermal strap, each of which extend behind the focal plane to the raft "crates" that house the front-end electronics and thermal sinks belonging to each raft.

KG will update this draft. 5-24-05

4.4 Data Management

4.4.1 Overall Description

The Data Management System (DMS) has these main responsibilities:

- Process the incoming stream of images generated by the Camera System during observing to generate and archive the Level 1 data products specified in this document.
- Provide real-time information on data quality to the Observatory Control System (OCS) during observing,
- Reprocess archived data products as required to incorporate pipeline improvements and correct errors, and to generate the Level 2 data products specified in this document.
- Provide a VO-compliant interface that makes publicly available all generated data products.

The successful design, implementation, and operation of the DMS is clearly a major challenge, for a variety of reasons:

- Its responsibilities extend over a wide range of timescales, from roughly ten seconds between images, to over a decade for generation of some Level 2 data products, and perhaps to multiple decades for curation of LSST data products.
- It collects large volumes of data at high rates
- To produce its data products, it must perform extensive computations on the data at high throughput and low latency
- It is a distributed system, produced by a distributed team
- High reliability and availability are required

In this document we present our approach to overcoming these challenges. We have high confidence that the DMS can be created and operated successfully, with an affordable budget, in the anticipated LSST baseline program schedule.

Section 4.4 is organized as follows. The remainder of Section 4.4.1 presents a concise overview of the DMS. As discussed in operational model in Section 4.4.1.2, the DMS will be a distributed system with at least one node co-located with the telescope, and another at a major computing/data center in a location with high bandwidth to the global public network.

The DMS has three major external interfaces: to the Camera System; to the OCS; and to the outside world via VO compliant interfaces. The LSST support for community science support, provided by the VO compliant interfaces, is discussed in Section 4.4.1.3. The Camera and OCS interfaces are discussed in Section 4.4.1.4.

As a means of managing complexity, the architecture of the DMS is layered, with an underlying Infrastructure layer supporting Middleware and Application layers above it. This aspect is discussed in Section 4.4.1.5.

The remainder of Section 4.4 presents the Data Products in 4.4.2, the Application Layer in 4.4.3, the Middleware Layer in 4.4.4, and the Infrastructure Layer in 4.4.5.

4.4.1.1 Operational Model and Mountain, Base, and Archive Nodes

The operational model for LSST data management is described in this section. While the accuracy of this model is dependant on additional analysis that will occur during the entire R&D period, it represents the current “most likely” scenario and serves to provide a departure point for further investigation.

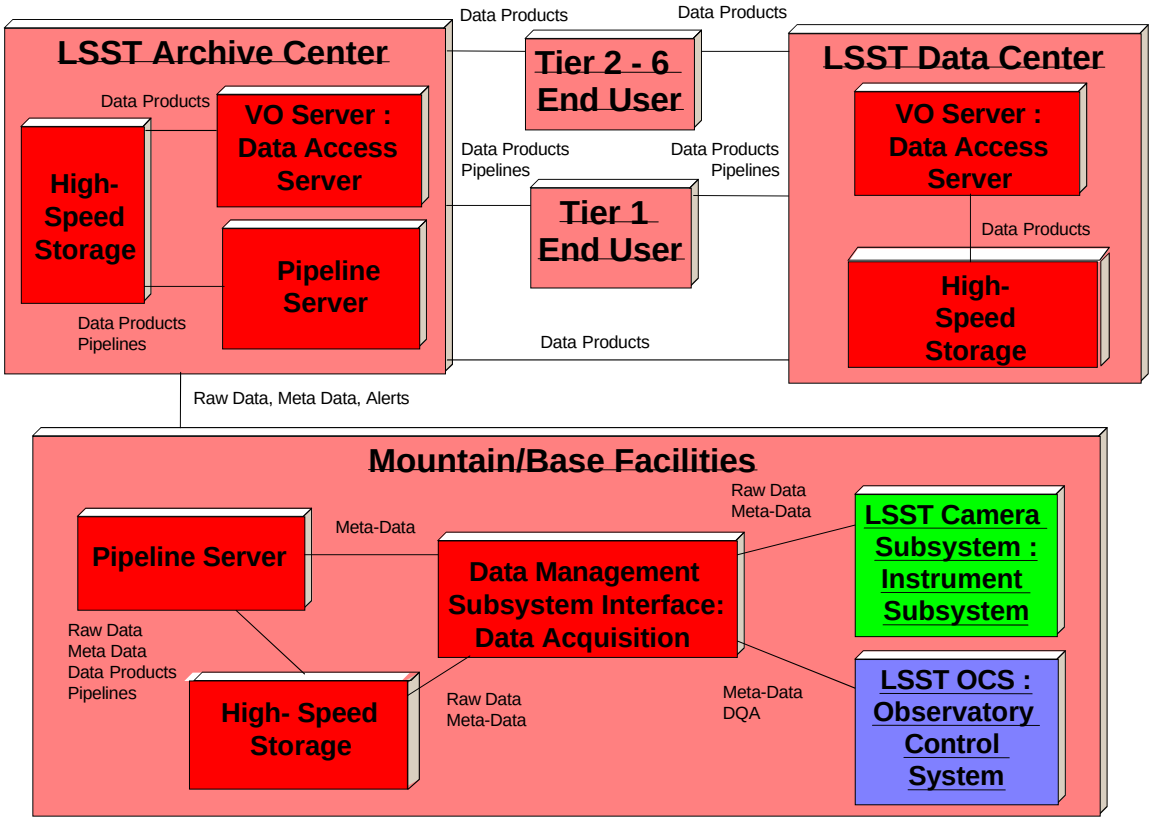


Figure 4.4.1-1 LSST DMS Operational Concept

4.4.1.1.1 Keeping Up with the Data

A number of factors scale the throughput requirements of the data management pipelines, including the observatory operational cadence, the output of the focal plane (number of pixels times pixel depth in bits), the expansion of 16-bit to 32-bit floating points during processing, the volume of derived intermediate and final data products produced as pipeline output. In addition, the real-time transient alerting requirements drive the pipeline latency requirements.

Given an average 10-hour night of observing the current focal plane size and cadence baseline translates into approximately 15 terabytes per night of raw data, not including daytime calibration frames. Given approximately 320 nights per year of observing, this nets to a total of approximately 4.4 petabytes per year of raw data, uncompressed. Typical lossless compression of image data of this nature yields about 50% reduction or 2.2 petabytes per year of compressed raw data.

All data must move through the data management system on timescales defined in part by the science requirements, but more importantly by the requirement that the system never get behind in processing. This latter requirement stems from the large data volume and high data rate; to get behind would delay the delivery of timely data products in an increasing wave through the entire survey period, or require major investments in additional processing and storage capacity to catch up. Neither alternative is deemed acceptable.

4.4.1.1.2 LSST Data Management Facilities

As indicated in Figure 4.4.1-1, the operational model begins at the data acquisition interface between the camera and telescope subsystems and the data management subsystem, and flows through to the end users. On the way, it moves through three types of managed facilities

supporting data management, as well as end user sites that may conduct science using LSST data or pipeline resources on their computing infrastructure.

The specific allocation of processing, archiving, and end user access responsibilities to these three facilities directly defines the amount of computing, storage, and network bandwidth needed in each. There is a distinct tradeoff to be made, with each variation in allocation driving the resultant cost, reliability, and availability of the entire system. This tradeoff must anticipate the technology that will exist in during LSST Construction and Operations, and is discussed in detail in sections 4.4.5. This will be the subject of research and modeling throughout the R&D phase, as described in section 5.4. Specific assumptions favoring a compute-intensive architecture (versus storage or network intensive) have been made in order to describe the baseline operational model.

The Mountain/Base Facility is composed of the mountaintop telescope site, where data acquisition must interface to the other LSST subsystems, and the "base facility" site, where rapid-turnaround processing will occur for data quality assessment and real-time alerts. While these two sites will be physically separated by some distance depending on the site selected for the LSST observatory, it is envisioned that these sites will be inter-connected with high speed, reliable communications under LSST control to enable managing them as one integrated processing facility.

The Archive Center is a super-computing class data center with high reliability and availability. This is where the data will undergo complete processing and re-processing and permanent storage. It is also the main repository feeding the distribution of LSST data to the community. It is anticipated that this facility will be located at an existing NSF supercomputing facility. Given the LSST data's anticipated high value to science and the decade-long survey period to collect it, best practice data management operations suggest that at least one additional Archive Center be commissioned for disaster recovery and possibly for load balancing.

One or more LSST Data Center sites for broad user access are envisioned, according to a tiered access model, as described in section 4.4.1.3. These centers provide replication of all or a subset of the LSST data and are chosen in order to optimize community access.

4.4.1.1.3 Daily Operations

In a typical observing night, the LSST will collect data throughout the night. As it is collected, the data will be transferred, along with the associated meta-data, from the mountain to the base in real time. Initial data quality assessment will occur at the mountain and base sites, on the individual detector output images, within 10 seconds of sensor readout. This assessment will be fed back to the Observatory Control System for telescope and camera operational control.

No Data Left Behind

It is assumed that even relatively "poor quality" data will be processed via the Image Processing Pipeline at the base site and ultimately archived, as we make no presumptions about the evolution of scientific algorithms that may permit this data to contribute to our knowledge. All quality assessment data will be archived with the associated images to enable later filtering for different applications.

Real-time Transient Alerting

Standard references, and co-added images and related meta-data from previous nights, as well as the most up to date version of object catalogs, will be hosted at the base site and updated daily. This will enable the detection and association pipelines to run here and produce real-time transient alerts.

Alerts will be generated within the SRD-specified latency requirement, currently envisioned as under one minute for the most time-sensitive alerts. Alerts will be dispatched to all LSST alert

subscribers from the base. It is anticipated that the IVOA VOEvent interface will be the primary mechanism for alert notification, and that alerts will be dispatched via multiple communications paths (e.g. internet2, public internet, dedicated circuits) to ensure no single point of failure. Refer to http://www.ivoa.net/twiki/bin/view/IVOA/ for more information regarding VOEvent.

Optimized data transmission

While the pipelines are running, the data will also begin transmission to the Archive Center in the United States, possibly several thousand miles away. Only the raw image data and associated meta-data will be sent during the observing night and on into the next day. The transmission will occur such that an entire night's observing can be transferred before the next night begins. The entire pipeline processing will be repeated at the Archive Center on the raw data. A subset of the archive center output (e.g. image templates for subtraction) will be sent back to the base, but this data volume is relatively low.

This approach has two distinct advantages:

- It minimizes the long-haul communications bandwidth from the base to the Archive Center, which is currently projected to be a larger cost driver than the computing and storage capabilities needed to repeat processing.
- It allows a consistency check to be made between the initial processing at the base and the repeat processing at the Archive Center to assist in detecting possible errors in transmission or out-of-date versions of data products or modules used in processing.

The link between the Mountain/Base and the Archive Center is envisioned as a leased fiber optic link. An alternative method of transmission is to save the data on removable storage and ship it via overnight courier. This method is deemed less desirable, since it involves waiting for the entire data set to be saved, shipped, and loaded before any further processing or access can be started. However, this method is almost certainly one of the backup methods that will be utilized in emergency scenarios. These alternatives are also a subject of research and tradeoff analysis during the R&D phase.

Quality Assessment Driven Data Product Release

While new data products will be produced with each night's processing, past experience with large surveys indicates that end users and their applications need a certain degree of stability and quality in data products to effectively utilize them. In order to accomplish this, the LSST will release updates to specific data products on different time intervals and conditions.

For example, basic data integrity checks and image quality assessment are all that is required to release raw image data, and it is envisioned that this data will be made available on a daily update.

On the other hand, objects in the Deep Object Catalog require multiple exposures and extensive processing to detect, classify, and characterize. Even then, the classification and characterization are probabilistic in nature. As such, a longer time interval or a probability threshold may be employed to determine when this catalog should be updated.

Additional quality assessment of all SRD-defined metrics (sky coverage, image quality/depth, cadence, photometric accuracy, astrometric accuracy, pipeline latency, etc.) will be performed each day, at the image, catalog, application, and mission levels, to determine if the quality criteria for release have been achieved. This assessment will analyze the state of the raw data and data products in terms of basic usability metrics, defect counts and rates, mission completeness metrics, and current version usage statistics, in an automated fashion.

It is envisioned that this raw assessment will be fed into a rules-based system that will make recommendations regarding update readiness and provide rationale. This analysis will then be used by the LSST program director to determine if an update is appropriate and of what artifacts.

In all cases, raw images will be retained and archived. The system will also provide for recovery of previous versions of derived products, within the limits of cost efficiency. In order to optimize storage, the versioning approach will likely not always be a full snapshot, but rather an incremental version including the provenance information needed to recover the artifact from a previous baseline version. This requirement drives several of the research areas described in Section 5.4.

Finally, as previously mentioned, the raw data will be processed twice, once in the base facility and once at the archive center, providing some checks on certain types of errors. It should be noted that while not in the baseline model, a third processing run would not only allow for error detection, but would also permit a "majority wins" approach to handling detected errors.

4.4.1.1.4 Re-Processing Operations

Over the operational life of the LSST, improved algorithms and new science will appear, driving the need to re-process substantial portions of the LSST data. The Archive Center will be sized to permit continuous re-processing of the LSST data at a rate permitting the entire data set to be re-processed once per year. This re-processing will occur in parallel with normal daily processing. This implies that the capacity will change with time, since the accumulated data volume increases with time.

4.4.1.1.5 Emergency Operations

Although the baseline operational model is scoped to provide minimal downtime, any given segment can and almost certainly will develop problems during the lifetime of the project. Backup measures must be identified to provide for all of the functionality described above. The solutions for storage and processing are well known and common practice in the large scale computing communities, ranging from fault tolerant storage such as RAID systems to redundancy in storage and processing units with automatic failover.

Similarly, the network links must have failover strategies to provide for effective data transport in the case of unexpected downtime. As mentioned above, as the data flows out of the data acquisition system, it would be cached in storage sufficient to provide for one or more nights of observations. The baseline MTTR for the mountain to base link is <24 hours. In this scenario, a slower backup link (microwave or manual transport) would provide additional bandwidth for an interim period, both to minimize the backlog to be transferred and to provide additional "catch up" bandwidth once the main link is operational.

In the event of a major interruption in mountain to base facility communications that can't be repaired in <24 hours, manual movement of this data via transportable medium (i.e. disks or tapes) will provide for data transport to the base facility on less than 24 hour timescales. In both cases, only those transient events that have a shorter window than the recovery period would be affected, all other science would still be achieved in required windows.

Clearly this places additional constraints on the mountain "cache" system, but existing technologies (e.g., hot swap disks and USB style connections) already meet these requirements. Further, if LSST owns and operates the mountain to base facility network link, the additional measure of having trained staff able to effect repairs provides for minimal downtime in the case of major faults such as cut fibers.

Similar measures of in house repair are more difficult to put in place for the longer base facility-to-archive center network segment, given the cost and complexity of running what will almost certainly be a long international network connection. However, provision for manual transport of large storage units is also practical for this link, via FedEx or similar courier service. While the turnaround time might be >24 hours, leading to some backlog of data processing, the

archive center will provide for the necessary additional processing to rapidly recover from the backlog, using the re-processing capacity available.

4.4.1.1.6 Simulating and Testing the Operational Model

Both the nominal baseline operations and emergency operations will be simulated to provide both real world experience and optimization.

It is interesting to note that the baseline operational model described above depends on bandwidth and computing power requirements that are not completely known at this time. A number of factors, such as the bandwidth needed for community alerts, the exact size of the Archive-to-Mountain/Base transfers, and the efficiency of a leased network between the observatory and the United States, are accounted for by allowing a range of capacities. Thus, there are opportunities to optimize our model. In particular, we can address the question of whether it is really necessary to redo all of the real-time processing at the archive site, or transmit at least some of the derived intermediate data and data products.

To test and refine our operational model, we plan to set up a testing and simulation environment on the national TeraGrid facility, a federation of supercomputing facilities distributed across the U.S. and connected by a dedicated, high-speed network. In this distributed computing environment, we will deploy a simulated LSST grid in which different TeraGrid nodes will each operate as a different node type (i.e., mountaintop, base, archive center, and data center sites) of the LSST grid. In this environment, we can test different variations on our operational model, adjusting what processing gets processed where, and test directly the effects of the network on limiting through-put. Section 4.4.2.2 discusses how the simulated LSST grid will be developed and used during the R&D phase of this project. It is assumed that this capability will also form the basis of a useful tool for optimization during LSST operations.

4.4.1.2 Community Access and LSST Data Centers

The LSST is intended to be an open resource to a broad range of end users in the astronomy and physics communities. This section describes representative communities of users, as well as typical access patterns of the LSST data and computing resources for various science and outreach purposes.

In this section we also introduce the concept of access tiers, which are implemented as a way of equitably apportioning processing, storage, and communications resources among communities with very different needs and capabilities. The use of tiered access is well-established in the HEP community and we are emulating this model.

4.4.1.2.1 LSST End User Communities and Access Patterns

(Note to reviewers: This section is TBD until we conduct a Community Access requirements workshop with the Data Products Working Group, anticipate doing this April time frame. This section is probably less important to DOE, except insofar as it supports DE/DM researchers)

4.4.1.2.2 LSST Data Centers and Access Tiers

As previously described, access tiers are implemented as a way of equitably apportioning processing, storage, and communications resources among communities with very different needs and capabilities. In this section we explain each tier and give an example of a user community and application for each tier.

- Service Tier 1: Access to archive computing cluster and storage farm, high-speed (e.g. Teragrid/10 Gbps) access to cluster and storage

- Service Tier 2: High-speed access to storage, low-speed access (e.g. 1 Gbps or less) to computing resources
- Service Tier 3: Low-speed access to storage and computing
- Service Tier 4: Low-speed access to storage only
- Service Tier 5: Public access only, via VO interfaces
- Service Tier 6: Other

4.4.1.3 Interfaces with OCS, Camera, Telescope

The interfaces to OCS, Camera, and Telescope are summarized in section 4.4.1.5 below, and are discussed in terms of their capabilities and performance requirements in the System Engineering section. The LSST development process requires the use of formal interface specifications to define the interfaces. In data management, we will maximize the use of off-the-shelf technologies for implementation, as they contribute to stable interfaces and system integrity.

4.4.1.4 Layered Architecture

Modern data management systems employ a variety of strategies to optimize the effort needed to develop an initial capability and to facilitate the ability to adapt and extend to new requirements. One of the principal approaches that LSST will employ is to utilize a layered architecture that separates the computing, storage, and communications infrastructure from the applications, and further separates the applications into services shared by or common to all applications (middleware) versus functions that are specific to individual applications. This approach has been employed in a great many data management systems to positive effect, including many in the domains of aerospace, intelligence, and geophysics.

The first separation between infrastructure and middleware, has the following characteristics and benefits:

- Creates a "virtual" interface to the underlying computer, storage, and communications hardware and system software (where heterogeneous resources appear as a single type of resource) consistently accessed by the applications, thereby simplifying the applications.
- Permits evolution of the underlying infrastructure with minimal rework to the applications, allowing technological improvements in infrastructure speed and capability to be readily incorporated.
- Permits re-architecting the distributed aspects of the applications (e.g. parallel processes, replicated data, messaging paths) with minimal rework to the applications, allowing for adaptable optimization of resource utilization.

An example of this layering is the TeraGrid, a middeware-based grid that permits a distributed network of computing clusters to appear as one virtual computing resource to the application.

The second separation, between middleware and applications, is characterized by the following features and benefits:

- Creates a consistent, standard set of services, eliminating inconsistent approaches and improving reusability of these common functions across applications:
    - Data access
    - Process control
    - Error handling/recovery
    - Data and process packaging and deployment
    - Security and access control
    - User interface support
    - Maintenance and administrative support

- Permits heterogeneous service implementations to appear as a single type of service, consistently accessed by the applications, thereby simplifying the applications.
- Permits the establishment of integrated frameworks for application development, where only the modules and data structures specific to the new application need be created, reducing the cost of adding new functions.
- Provides the ability to leverage both open source and commercial middleware implementations, thereby amortizing the cost of implementing the services across a user base beyond the individual project.

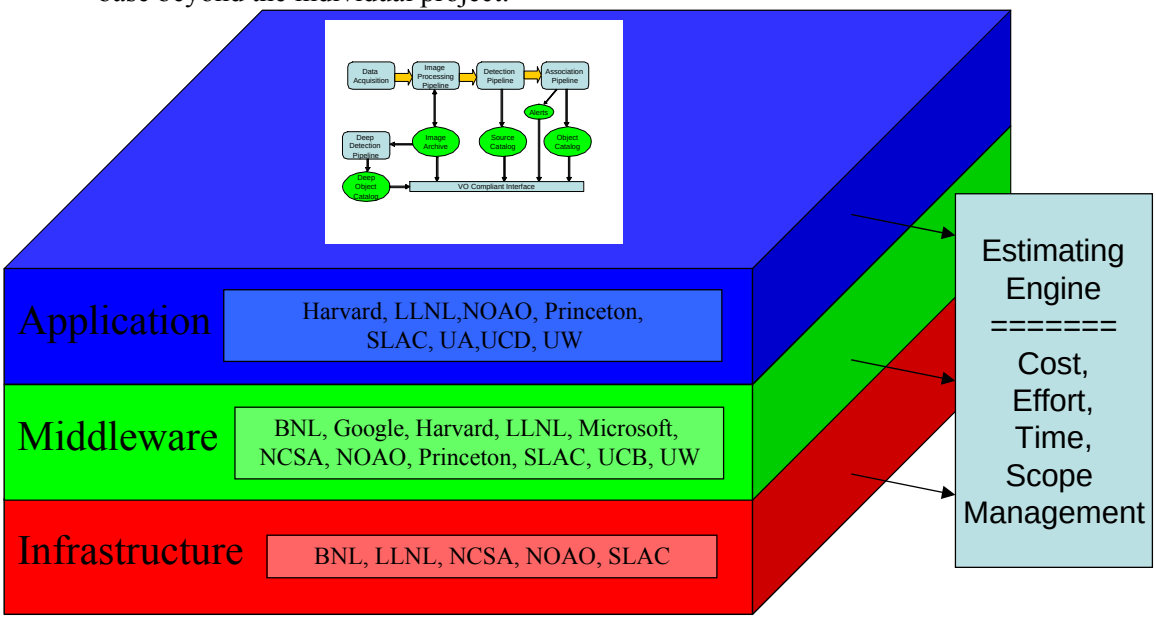


Figure 4.4.1-2 The LSST Data Management Layered Architecture

Figure 4.4.1-2 depicts the 3 layers of the architecture and indicates the LSST partner organizations with R&D technical responsibilities in each area. The LSST pipelines and data products as described in section 4.4.2 are located at the top in the Application Layer.

Note that each layer will be specified using Unified Modeling Language (UML) diagrams in accordance with the Iconix Process.

The diagram also indicates that as each layer is specified, the specifications are fed into the LSST Data Management Estimating methodology.

The remainder of this section describes each layer in summary fashion. The layers are discussed in detail in Sections 4.4.3, 4.4.4, and 4.4.5.

4.4.1.4.1 The Application Layer

The application layer contains the astronomy-specific pipeline modules and data products of the LSST. This layer contains all of the astronomy domain-specific software that enables the LSST to process the images and produce catalogs and alerts and to assess data quality. This layer will include most of the custom software that will be needed for the LSST DMS. The application layer is directly driven by the LSST science and derived system requirements, as described in the requirements flow-down traceability matrix in Section 3.2.

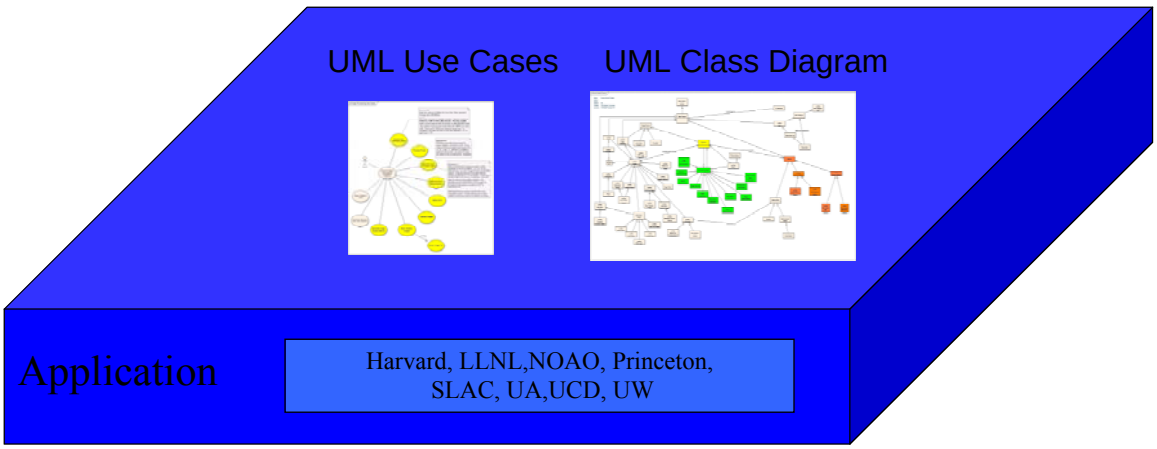


Figure 4.4.1-3 The Application Layer

Refer to Section 4.4.3 for further details regarding the Application layer.

4.4.1.4.2 The Middleware Layer

The Middleware layer is composed of components providing services (exposed as interfaces) that the applications invoke to access data, run processes, log status and errors, and other functions common to many applications. Given the broad scope of open and commercial work in web services, grid technologies, database and file systems, user interface libraries, and other middleware, the LSST will have a broad base of middleware from which to select and we presume this layer will largely be composed of off-the-shelf software.

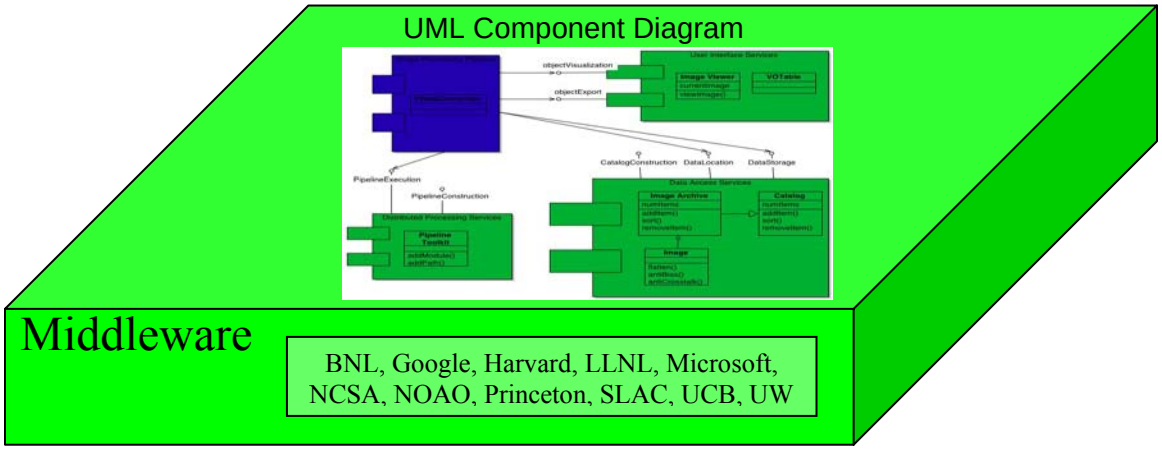


Figure 4.4.1-4 The Middleware Layer

Middleware services will be specified using the UML Component Diagram, which shows the packaging of the middleware and the interfaces to the services. Application layer components access the services via these interfaces. This access is further specified in UML Sequence Diagrams (also known as Interaction Diagrams).

The advent of web services and component technology middleware has opened up architectures in which at least some of the data or functions are being accessed via a remote interface, permitting portions of the application to reside in different locations than the middleware serving them. A proxy for the middleware is accessed locally by the application, which then

communicates to a stub that invokes the middleware service and returns the result to the proxy and finally to the application.

The middleware provides this remoting capability, typically in a way that the application need not concern itself exactly where the service or data being served actually resides. This is a fundamental property of robust client-server and multi-tiered architectures, and is important for scalability and extensibility. Indeed, location transparency is one of the primary benefits of middleware, although performance implications must still be carefully analyzed and managed during design.

Refer to Section 4.4.4 for further details regarding the Middleware layer.

4.4.1.4.3 The Infrastructure Layer

The infrastructure layer contains the computing, storage, and networking hardware and systems software that hosts the middleware and application layers.

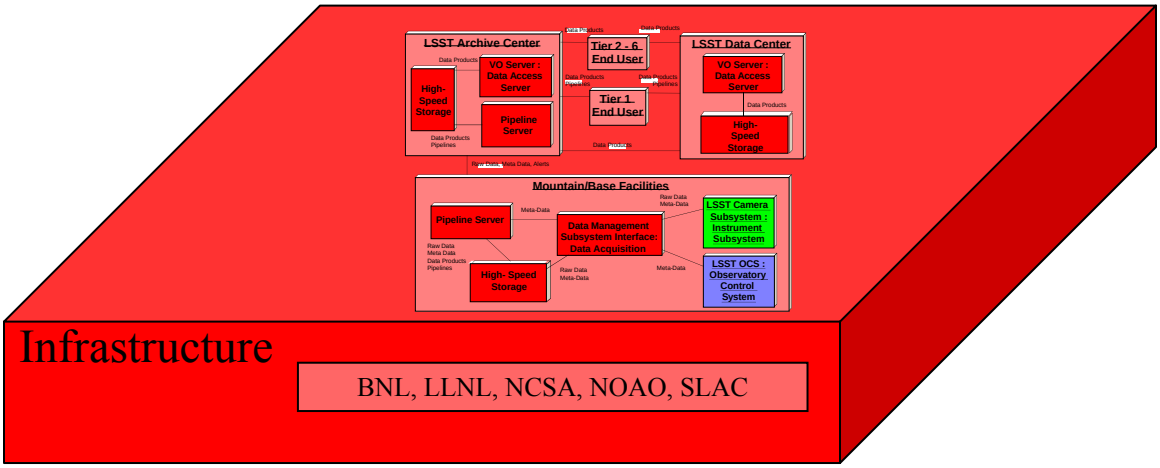


Figure 4.4.1-5 The Infrastructure Layer

We will specify the Infrastructure Layer with UML Deployment Diagrams. These diagrams depict the hosting of application and middleware components on the infrastructure resources. The diagrams also depict the type, quantity, and capacity of the resources, and the interconnections and communications protocols in the infrastructure.

This layer is presumed to be almost entirely off-the-shelf. While the LSST will certainly require very high-performance hardware and systems software by today's standards, current trends indicate this will be quite feasible in the required time frame at a reasonable percentage of the total LSST project cost. Analyzing these trends and optimizing the LSST DMS architecture relative to the trends are the subjects of considerable research activity, as described in Sections 5.4.2.2 and 5.4.2.3.

Refer to Section 4.4.5 for further details regarding the Infrastructure layer.

4.4.1.5 Development Standards

Certain development standards will apply to all LSST software, independent of layer. Those overall standards are described in this section.

4.4.1.5.1 Design Standards

All custom components will have design documentation in conformance with the Iconix process for the DMS, including the application of UML. Off-the-shelf components will provide

the same if available, or a suitable substitute enabling LSST Data Management developers to effectively integrate applications with the middleware and support the middleware in LSST usage.

4.4.1.5.2 Coding Standards

Custom components will be implemented in C++ or an LSST standard scripting language that will be specified (candidates include Python and Ruby). Off-the-shelf software that supports a C binding will be permitted as well.

All code will be required to meet coding standards, covering the following areas:

- Naming of classes, objects, and member functions
- Formatting of code text
- Inclusion of standard comment blocks that provide information on the purpose, usage, and implementation of all classes

The coding standards will be selected from the many openly available standards currently available.

Off-the-shelf components will provide the same if available, or suitable substitute enabling LSST Data Management developers to effectively integrate applications with the middleware and support the middleware in LSST usage.

4.4.1.5.3 Testability

All custom components must be supplied with test harnesses and test datasets that allow automated testing of their functionality:

- Unit test cases and data that exercise all methods in every class
- Normal execution and exception handling
- All parameters and both valid and invalid data inputs

In addition, all components will be subject to semi-automated integration and system testing:

- Integration test cases and data that exercise all use cases within a component, including all public interfaces of the component`
- System test cases and data that exercise all use cases that span multiple components, (Note that in the Iconix Process, system test cases are derived directly from use cases.)

Off-the-shelf components will provide the same if available, or suitable substitute enabling LSST Data Management developers to effectively integrate applications with the middleware and support the middleware in LSST usage.

4.4.1.5.4 Documentation

All custom components will provided with documentation

- Design documentation (PDR, CDR, FDR) in UML format and supplementary tables and graphics
- Code documentation (Doxygen)
- User manuals for developers and end users

Off-the-shelf components will provide the same if available, or suitable substitute enabling LSST Data Management developers to effectively integrate applications with the middleware and support the middleware in LSST usage.

4.4.2 Data Products

The LSST SRD associates a set of analysis methods with the LSST observing programs. To execute these methods the LSST DMS is required to generate a set of data products, the subject of this section. The data products are organized into two groups, distinguished by the cadence with which they are generated. Level One products are generated by pipeline processing the stream of data from the camera system during normal observing. Level One products are therefore being continuously generated and / or updated every observing night. This process is of necessity highly automated, and must proceed with absolutely minimal human interaction. Level Two products, including calibration images, co-added images, and the resulting catalogs, are generated on a much slower cadence, and their release will be driven by data quality assessments. Although many of the steps that generate Level Two products will be automated, they need not all be so, and significant human interaction may be required.

All DMS data products are tied to a UML domain model, as shown in Figure xx1. The domain model shows the attributes and family relationships of the data products, and is the first major step to their complete specification.

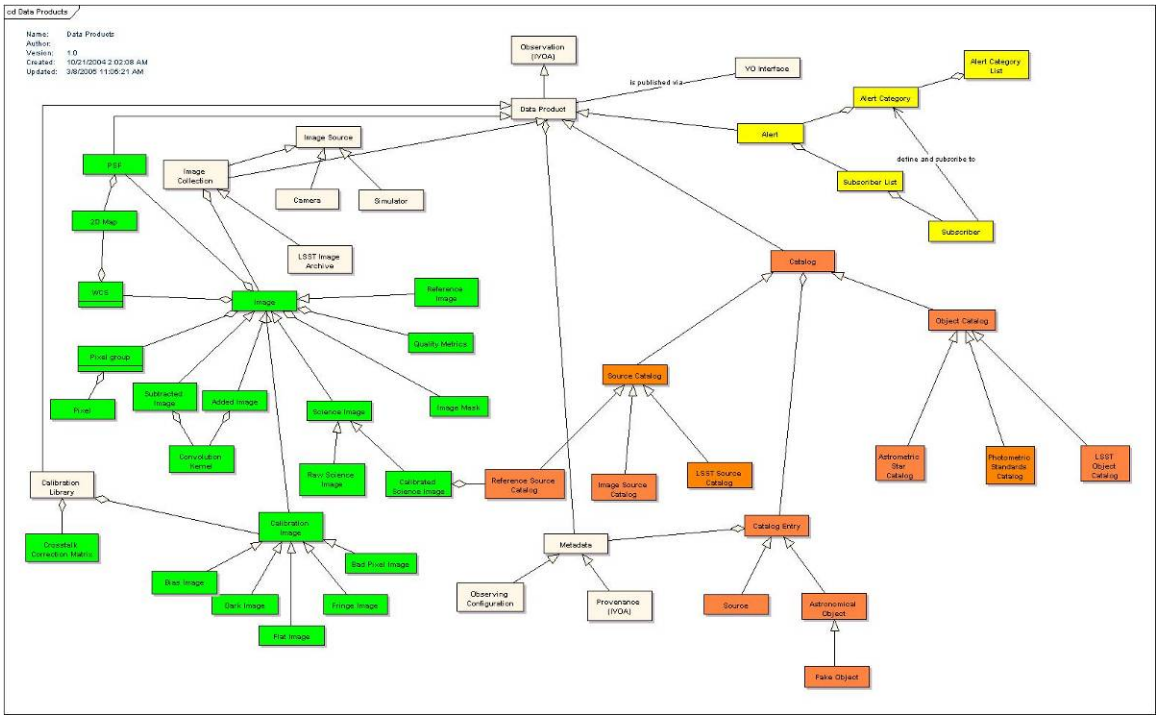


Figure 4.4.2-1 LSST Data Products Domain Model

Data products which are archived may have a separate internal format for pipeline use and an external format for archiving and public data access. This difference is noted below where applicable.

4.4.2.1 Level One Data Products

Level One data products are divided into Images, Catalogs, and Alerts. These are separately color coded in Figure 4.4.2-1.

4.4.2.1.1 Images

The internal and external formats of images will differ. The internal format will be chosen to maximize I/O performance. The external format will be chosen to facilitate transport to external systems and to capture logical associations within the data and metadata. One possible choice is multi-extension FITS (MEF).

4.4.2.1.1.1 Raw Science Image

The Camera System generates 7040 parallel streams of 16-bit pixel data during readout. The raw science image is a container for that pixel data, along with the metadata for the exposure made available by the Camera System and/or Observatory Control System. This metadata can be broken down in to science and engineering metadata, and includes:

- Site metadata (site seeing, transparency, weather)
- Telescope metadata (active optics state, environmental state)
- Camera metadata (wavefront sensors, environmental state)

Raw science images are the primary data product of the LSST, and will be archived.

4.4.2.1.1.2 Calibrated Science Image

The transformation of a Raw Science Image into a Calibrated Science Image is the task of the Image Processing Pipeline. The Calibrated Science Image has many properties not possessed by the Raw Science Image:

- All data read out from multiple readouts of a single ccd have been combined into a single image structure. This reduces the number of separate image structures from an exposure to 201 from 7040.
- Instrumental signature has been removed. This signature includes at least crosstalk, bias, flatfield response.
- Pixels with significantly nonlinear response characteristics have been added to the image's bad pixel mask
- Fringing from night sky emission has been removed, if required.
- Areas of the image adversely affected by artifacts such as satellite trails or ghost images from nearby bright sources have been identified, and a mask generated for them.
- The image is astrometrically calibrated, so that there is a precise mapping from pixel coordinates to ICRS sky coordinates
- The image is photometrically calibrated, so that there is a precise mapping from a source's intensity in data units and its astronomical AB magnitude for the filter used.
- The point spread function (PSF) has been accurately determined.
- Data quality has been assessed and recorded in the image metadata.

Generating the Calibrated Science Image involves a number of challenges, particularly in the areas of PSF determination, astrometric calibration, and photometric calibration. These issues are discussed in Section 3.4.3.

A Calibrated Science Image can be recreated at will from the raw science image and associated calibration data. It is a possible design option to not archive these images, instead recreating them "on-the-fly" as needed. This choice will be driven by analysis of the cost of processing versus bandwidth as well as scientific usability.

4.4.2.1.1.3 Subtracted Science Image

Subtracted Science Images are used to enable detection of transient sources. A Subtracted Science Image is created by the Image Processing Pipeline from two input images. One input image is always a Calibrated Science Image. The other input can be either a Calibrated Science

Image, or a Co-added Image. The subtraction process makes use of the PSF information, astrometric calibration, photometric calibration and mask information of the two images. The metadata of the resulting image logically includes the metadata of each input image, as well as information about the registration and PSF matching process.

A Subtracted Science Image can be recreated at will from the two input images. It is a possible design option to not archive these images, instead recreating them "on-the-fly" as needed.

4.4.2.1.2 Catalogs

Catalogs are data products derived from images. Logically, the structure of a catalog is that of one or more tables with a row for each entry. Each row possesses a number of attributes, arranged into columns. The relational structure of tables within catalogs will be complex, and is not yet determined. It may be practical, for example, to store catalogs as tables within a relational database (RDB). This decision will be based on information gained from D&D activities. All catalogs will be archived.

4.4.2.1.2.1 Source Catalog

Every Calibrated Science Image and Subtracted Image is processed by the Detection Pipeline. The outcome of this processing is a set of sources found in the image. Each source has a key that references the image metadata and the pipeline parameters used to derive it, and a set of attributes:

- Focal plane position
- Focal plane position error
- ICRS sky position
- ICRS sky position error
- Shape parameters, quantifying the way that the source shape differs from the local PSF.
- Instrumental fluxes (multiple algorithms)
- Instrumental flux uncertainties
- Photometric fluxes (from instrumental fluxes)
- Photometric flux uncertainties
- Set of flags showing any problems encountered in deriving the source properties.

4.4.2.1.2.2 Object Catalog

An entry in the source catalog is the result of processing a single image. Depending on the pipeline parameters chosen, such as detection thresholds, some fraction of source catalog entries will result from noise rather than an actual astronomical source. In contrast, the object catalog synthesizes the results from source detections over a time period. As a result, it

- Filters out noise sources
- Contains information on time dependent flux
- Contains information on object motion
- Associates source detections from multiple filters, resulting in color information
- Associates source detections from multiple epochs
- Contains current best values for object properties

The Association Pipeline is the creator of the Object Catalog. The catalog is logically more complex than the source catalog, because it must capture associations over time and between filters. Associations that involve rapidly moving solar system objects can in general be made only probabilistically, so a given source may have a significant probability of association with multiple other sources at a given time. Similar issues arise when changing seeing conditions cause sources to split into multiple sources and merge back again. The Object Catalog data

structure must be flexible enough to accommodate these changing assessments of the "truth" regarding the sky.

4.4.2.1.3 Alerts

An Alert is a message sent to a group of registered clients, informing them of the occurrence of a particular class of transient event. Different clients will subscribe to alerts for different types of transients, by specifying criteria (e.g., brightness, color, rise time) for events for which they want to receive alerts. The details of this data product are still incomplete, and are expected to conform to the evolving VOEvent standard. An Alert will be a small data structure including;

- Class of event detected
- Confidence of detection
- All associated information from the Object Catalog
- "Postage stamp" images centered on the event for all time points in the event sequence

Alerts are always archived to maintain a historical record and for false alert analysis.

4.4.2.2 Level Two Data Products

Level Two data products include both images and catalogs. They include data products for which extensive computation is required and those for which many observations are required for their production. While some Level Two products can be defined now, the set is assumed to be incomplete since some science programs will need to create data products beyond those currently listed.

4.4.2.2.1 Images

4.4.2.2.1.1 Calibration Image

The transformation from Raw Science Image to Calibrated Science Images requires a variety of Calibration Images that describe the instrumental signature of the telescope and detector. Their types include

- Dome flat
- Sky flat
- Bias frame
- Dark frame
- Fringe frame
- Illumination correction frame

These will be produced by the Calibration Pipeline, with a cadence that will be set largely by operational experience with the LSST. The cadence for these calibration images will vary widely, and will be determined by the calibration plan.

These data products are always archived, and are additionally important to monitor the health of observatory systems.

4.4.2.2.1.2 Co-Added Image

It is an LSST science requirement to produce images of the static sky for a variety of scientific analyses. Multiple images could be combined to achieve the deepest possible image, or to achieve the best PSF. This requirement is met by a Co-Added Image, which is produced by combining a set of images taken of the same region of the sky with the same filter that meet the required conditions. This task is periodically performed by the Image Processing Pipeline. As with the Subtracted Image, a Co-Added Image possesses all the attributes of a Calibrated Science

Image with the addition of links to metadata for all contributing images, and information about registration and PSF convolution.

Production of Co-Added Images is computationally expensive, and in view of the relatively small storage involved, they are always archived.

4.4.2.2.2 Deep Object Catalog

The Deep Object Catalog is the object catalog resulting from running the Deep Detection Pipeline on Co-Added Images. Object attributes in this catalog lack most time dependent information but have additional attributes to describe

- Detailed shape descriptions and photometry of extended objects
- Proper motions and parallaxes

The Deep Object Catalog is likely to be produced and released at a slow cadence, and new releases would be predicated on significant improvements in the precision or information content.

4.4.2.3 Level Three Data Products

Level Three data products are those that are derived from level two data products, usually requiring the use of LSST data across significant areas on the sky. These may include derived catalogs and/or the results of large queries. Examples include phase-folded light curves for periodic variables, catalogs of specific subsets of objects, catalogs of derived properties such as photometric redshifts, and catalogs of clusters. Level Three data products will be archived if they require significant resource investment or are commonly requested by users.

4.4.2.4 Core Science Data Products

These are the archivable science products from the core projects that the LSST mission is designed to address. These products are the result of running science-specific pipelines on level one and two data products.

4.4.2.4.1 Weak Lensing Products

These would include shear maps, mass maps, …

4.4.2.4.2 Supernova Program Products

These include light curves, photometric redshifts, classifications, …

4.4.2.4.3 Clusters and Large Scale Structure Products

These include cluster catalogs (including cluster photo-z's), cluster luminosity functions, ….

4.4.3 Application Layer

The application layer of the DMS is partitioned into top-level components, as shown in Figure 4.4.3-1. Each is responsible for producing a subset of the LSST Data Products described in Section 3.4.2:

The Image Processing Pipeline (IPP) produces all Level One and Level Two Images

The Detection Pipeline (DP) produces the Source Catalog

The Association Pipeline (AP) produces the Object Catalog and Alerts

The Deep Detection Pipeline (DDP) produces the Deep Object Catalog

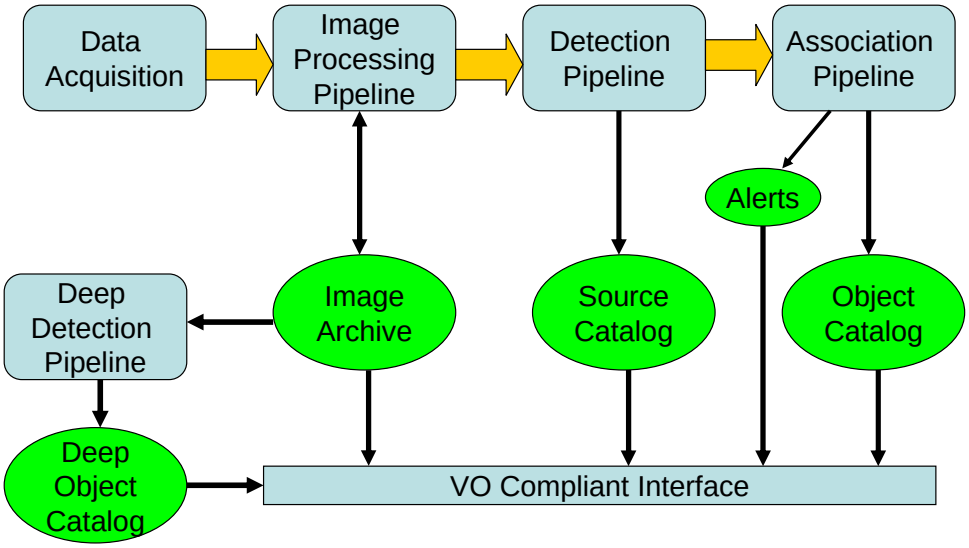


Figure 4.4.3-1 Application layer view of the DMS pipelines

4.4.3.1 Requirements

Application components are required to meet a set of implementation standards, which are intended to ensure that they fit properly into the DMS architecture, and are supportable through the life of LSST.

4.4.3.1.1 Use of LSST Middleware classes

The LSST Middleware Layer is discussed in Section 4.4.4. It provides an API on which all application components are built. This API provides:

Distributed Processing Services

Data Access Services

User Interface Services

Use of the API ensures that only thoroughly tested functions are utilized. Additionally, it insulates application components from changes in the underlying Middleware and Infrastructure layers, and therefore is an important contributor to maintainability of the overall DMS.

4.4.3.2 Image Processing Pipeline

The Image Processing Pipeline (IPP) is used in three different modes:

Level One Science Reduction Mode

Level Two Science Reduction Mode

Calibration Mode

These modes produce the associated Image Data Products. The modes are implemented by sequencing the execution of a mode-specific subset of a common set of image processing tasks, as described below. Our baseline design assumes a focal plane mosaic composed of CCD's. In the event that LSST uses CMOS arrays instead, the details of some of the tasks will change significantly.

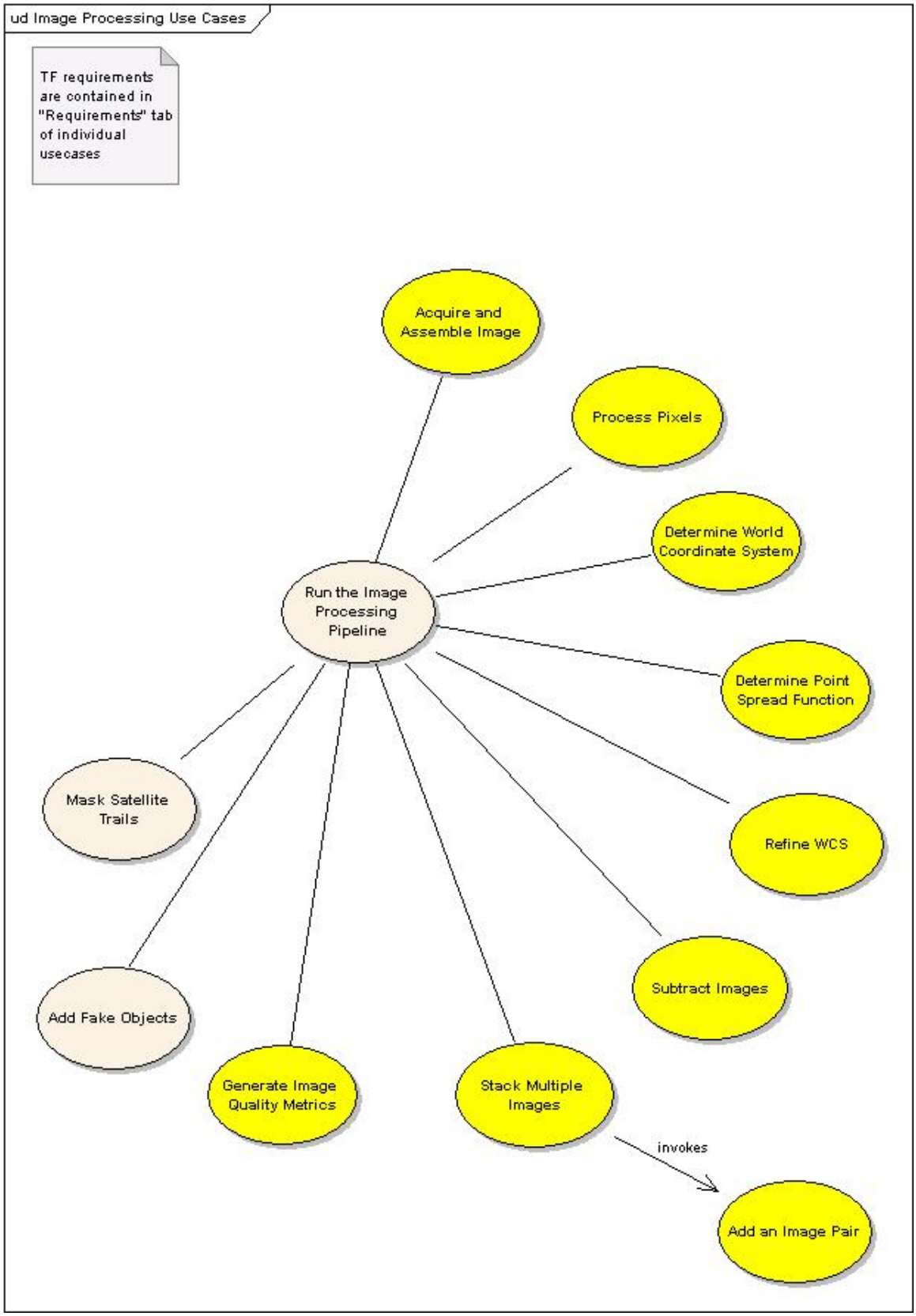


Figure 4.4.3-2 Use case diagram for IPP

4.4.3.2.1 Image Assembly

As discussed in 4.4.2.1.1.2, the first step in the transformation of a Raw Science Image into a Calibrated Science Image is the combination of data from all the individual readout channels from a single CCD to form a single image for that CCD, resulting in 201 individual CCD images. The design of this task is dependent on details of the Camera / DMS interface, not yet finalized. In any case, it will be a simple algorithm that is not data dependent, and is trivially parallelized.

4.4.3.2.2 Crosstalk, Bias, and Flatfield Correction

This group of tasks removes a number of instrumental artifacts from the Raw Science Images. Crosstalk occurs due to the unwanted electrical coupling of the readout signals of different CCD readout channels. The effect is seen most clearly when one readout region contains a bright object. Crosstalk will cause the signal from readout of the bright object to appear as spurious objects (either bright or dark) in other channels. The coupling is linear, and may be removed by a matrix multiplication:

Icorr = Axtalk * I

The structure of the crosstalk matrix will not be known with certainty until camera components are available for test. Our working assumption is that only channels on the same raft (see Section 4.3.4) are coupled. This gives the crosstalk matrix a block diagonal form with 288 x 288 blocks. We note that the matrix multiplication can be performed in parallel by row.

4.4.3.2.3 Masking of Artifacts

The Image class contains a pixel mask used to identify pixels that for some reason should be specially treated in later pipeline stages, in particular source detection. We note that the mask values are not binary, but rather can take on a variety of possible values. This allows subsequent pipeline stages to be highly adaptable. For example, a hot pixel may simply be ignored, while a pixel affected by stray light may have a value interpolated in some way. Reasons for a pixel to be masked include:

"Hot", "cold", near charge trap, or otherwise responds nonlinearly

Satellite trail

Diffraction spike, or other artifact from telescope pupil

Affected by ghost image of bright source, or other stray light

The first category is handled by initializing the mask with a set of "bad pixels" which are a static (or at least slowly varying) property of the detectors. The remaining categories are addressed by a variety of special-purpose image processing modules. Satellite trails can be detected by the Hough transform, which maps linear features in the image to points in the transformed image. Artifacts from telescope optics will likely need to be identified with the assistance of a numerical model of the telescope optics. The best method for identifying these artifacts, and handling their processing in later pipeline stages, will need to be determined during D&D.

4.4.3.2.4 Astrometric Calibration

The astrometric calibration of an image is encoded in its World Coordinate System (WCS). In addition to describing complex geometrical distortion terms that vary with time, the WCS depends on the color of the source. Accounting for this color dependence will be important for achieving LSST's required astrometric accuracy, and may require us to define a data structure for the WCS that goes beyond the existing FITS standard. This aspect will be addressed in the R&D phase.

The baseline design for the IPP associates a WCS with each individual CCD, rather than with the focal plane mosaic as a whole. This maximizes astrometric accuracy, and allows a high degree of parallelism in the WCS determination. An initial WCS is created using the telescope pointing information combined with static information about the mosaic geometry (pixel size; location and orientation of each individual CCD). This initial WCS, relatively crude, is refined in two stages as pipeline processing proceeds.

In the first stage, isolated stars are detected in the image using a fast algorithm such as Sextractor [ref Bertin]. At this point in the IPP there is only crude PSF information, so the positions of the isolated stars are determined with only moderate accuracy. The star positions are then matched to an astrometric catalog, and the initial WCS is refined.

The second stage occurs after an accurate PSF has been determined. Isolated stars are again detected, this time using the known PSF, thereby achieving the maximum possible accuracy. The stars are rematched to the astrometric catalog, and the WCS refined for a second time.

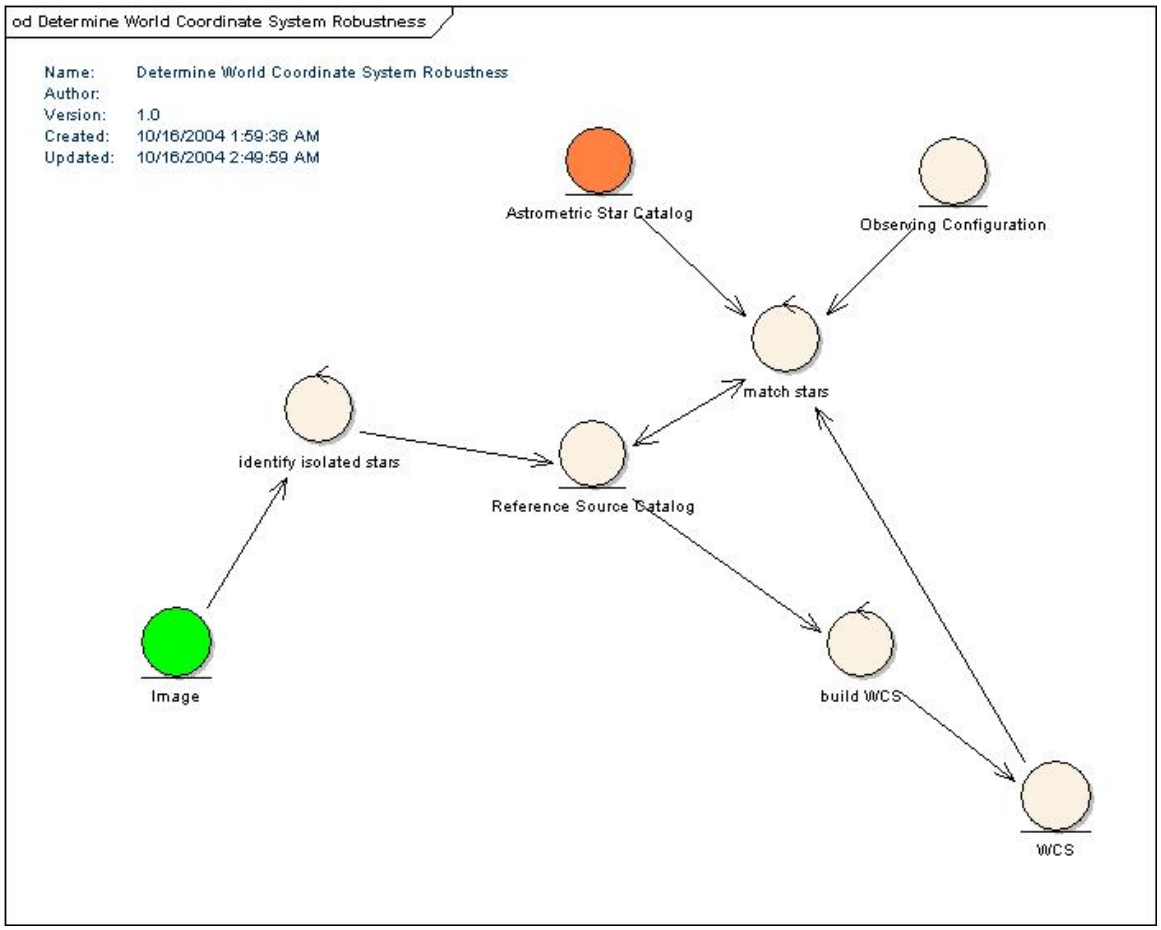


Figure 4.4.3-3 Robustness for "determine WCS"

4.4.3.2.5 PSF Determination

The LSST will pose significant challenges in accurate determination of the PSF across the focal plane. These challenges arise from the combination of fast optics, wide field of view, and demanding photometric and astrometric accuracy requirements derived from the SRD. The PSF shape will vary significantly with field position, source color, and time, due to the changing

atmosphere and imperfect active optics compensation for changing telescope geometry. PSF determination is therefore identified in Section 5.4.2.1.2.4 as a major R&D task. We outline here a baseline approach, certain to be modified or replaced as an outcome of R&D.

As with the astrometric calibration, for the baseline IPP we choose to determine the PSF on a per-CCD basis rather than for the focal plane as a whole. PSF determination occurs after a medium quality WCS is available. This allows a pre-determined set of isolated stars suitable for the purpose to be reliably identified in the image. The space of PSF functional forms is constrained by knowledge of wavefront information supplied by the camera system and included in the Raw Image data structure. Making use of this constraint, an optimal fit of the constrained PSF function to the isolated star data is performed, and a PSF representation built from the fit parameters.

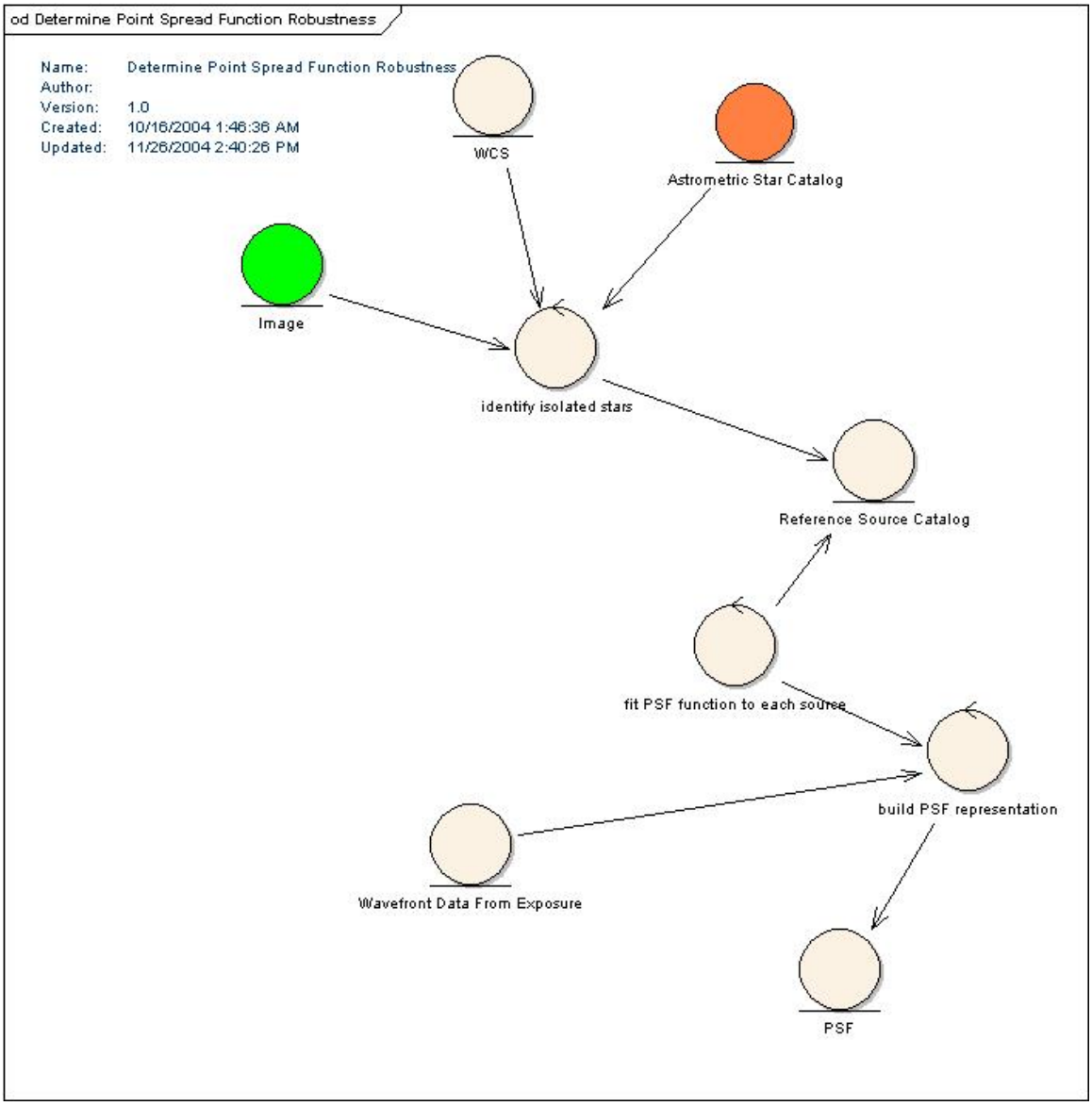


Figure 4.4.3-4 Robustness for “determine PSF”

4.4.3.2.6 Photometric Calibration

The LSST SRD dictates photometric accuracy of 1%. The SDSS experience indicates that this will be possible, but not easy. Among the issues to be faced are:

Flat field accuracy limited by scattered light, sky color mismatch to object color

Mismatch of LSST photometric system and transformed reference catalog system

Nonlinearity of CCD response

Variability of extinction in both field angle and time

Variability of telescope transmission with field angle

While the SDSS experience is valuable, it is not directly transferable to LSST due to the very different observational modes. We expect to perform photometric calibration in two stages. The first, discussed here, is performed on single images by the IPP, and sets the photometric zero point of the image. The second, is performed during a Level Two pass over the object catalog, and makes use optimal use of multiple observations of individual objects to refine the photometric calibration in the object catalog.

As with WCS and PSF determination, the baseline concept is to perform photometric calibration independently for each chip. Making use of the PSF to perform optimal photometry of stars from the photometric reference catalog, the zero point is determined by a least squares fit of the instrumental magnitudes to the reference magnitudes. Due to the wide field of the LSST, the fit must account for the variation in telescope transmission with field angle.

4.4.3.2.7 Data Quality Assessment

The OCS relies on the DMS to provide prompt assessment of a set of image quality metrics:

Site seeing

Telescope aberrations

Detector performance

Transparency

Photometric calibration quality – determined by degree of correspondence of zero points between chips as well as fit quality for each chip.

Astrometric calibration quality – determined by degree of correspondence of WCS between chips as well as fit quality for each chip

All of these "prompt" metrics can be performed at the level of the IPP. More sophisticated metrics, such as evaluating the width of the stellar locus in a color – magnitude diagram, are performed during Level Two passes over the Object Catalog.

4.4.3.2.8 Image Subtraction

Generation of prompt alerts will rely largely on high quality image subtraction performed by the IPP. Image subtraction has been successfully employed by numerous surveys. Nearly all of them have relied on variants of a single subtraction algorithm. In spite of its success, this algorithm often results in significant systematic errors around bright (but unsaturated) objects, leading to elevated false alarm rates during source detection. The algorithm is optimal, but only in a restricted sense, since it is limited to a specific form of convolution kernel. As an R&D task, we expect to generate an improved algorithm that capitalizes on the high quality knowledge we will have of our PSF.

4.4.3.2.9 Image Addition

Image addition is employed in Level Two passes over the image archive to generate images with characteristics that are improved over individual Calibrated Science Images in several respects:

Higher signal-to-noise, ideally increasing as sqrt(N)

Elimination of gaps in sky coverage resulting from inter-CCD gaps

Elimination of artifacts due to cosmic rays, and to some extent, diffraction spikes

There is extensive experience with image addition. During the R&D phase, we will evaluate the existing algorithms using simulated LSST images, and identify possible approaches to improving their performance.

4.4.3.3 Detection Pipeline

The task of the Detection Pipeline is to identify and measure sources in an input image. The input image must have been calibrated by the IPP, but can be one of several types:

Calibrated Science Image

Subtracted Science Image

Co-added Science Image

The definition of "source" in this context requires care. At the depth of LSST images, most astronomical objects are galaxies rather than stars. The identification and measurement of these extended objects poses a number of challenges.

Their surface brightness will be far below that of the sky. Errors in sky subtraction therefore have a major impact on the size and brightness measured for an extended object.

Extended objects generally do not have well-defined edges, and their measured brightness therefore depends on how much of their spatial extent is actually included in the measurement, and how flux beyond that extent is modeled. It is also problematic which photometric quantities are the most useful to store in the catalog.

Their shapes are in general complex. A method for parameterizing them must be chosen which is practical to store in the Source Catalog, but which is also sufficiently general and complete that scientific analysis is not compromised. We need to ensure that extended objects of scientific interest that have unusual shapes, such as arcs from strong lensing, are suitably measured.

We are not yet ready to provide definitive answers to all of these issues. We will target them with a variety of investigations during R&D, guided by experience from preceding large galaxy surveys such as 2dF and SDSS.

As with other pipelines in the DMS the Detection Pipeline can operate in a number of different modes, which are selected by a pipeline policy module (PPM). The PPM for the Detection Pipeline will determine characteristics such as:

Whether an input object catalog is utilized, and if so, how

Choice of sky background estimation algorithm

Values of detection thresholds

The sequence of steps in a baseline Detection Pipeline follows. It assumes that only an input catalog of bright objects is used. Using a full input catalog would affect both sky determination and de-blending:

Interpolate over all image defects, as identified in the input pixel mask. The goal of such interpolation is to simplify succeeding processing steps so that there is no need to handle bad pixels specially. The information that they are bad is, naturally, preserved. An appropriate

interpolation scheme is a linear predictive code, using the known auto-correlation function of the image (i.e. the PSF).

Estimate the sky background. We will do this by using an input catalogue of bright objects (both stars and galaxies). For the galaxies, we subtract a simple best-fit Sersic model; for the stars we subtract some model based on the current scattered light model; this may be a function of the color of the star. Note that we need to be able to handle the wings of objects off the frame. In general, these may be complex due to reflections within the optical system.

Once the large, bright sources are subtracted, we can determine the sky in large, overlapping regions (e.g. 256x256pixels on 128x128 centers. TBD). We then fit a smooth function through these estimates, and subtract the resulting sky image. This sky image is saved as part of the image metadata.

Smooth the image with the PSF. It is probably sufficient to use a simple Gaussian for this step, and this can be achieved by an efficient x- then y- convolution in real space.

Detect sources with at least one pixel above threshold in the smoothed image (which may be thought of as an image of the probability of there being a PSF at a point).

Grow the detected set out to the size of the PSF (as a single pixel over threshold corresponds to the significant detection of a PSF). Detect all distinct peaks within the detected sources.

Merge the models subtracted above into the list of detected pixels. Remove all pixels associated with sources from the sky-subtracted images, and replace them by noise.

Optionally re-estimate the sky level now that all (not only bright) sources are removed.

For every source in list of detections:

Reinsert pixels associated with the source into the image (note that the low S/N per pixel outer parts of e.g. galaxies were not detected, and so are still present in the image). If a model was subtracted, add it back in now.

Run a de-blender, similar to the SDSS de-blender. The output of the de-blender is a set of children associated with each source. Each child is itself an image of a (probable) astrophysical object.

Remove the parent source from the frame.

For each child, measure photometric, astrometric, and shape parameters and enter into Source Catalog.

4.4.3.4 Association Pipeline

The Source Catalog, produced by the Detection Pipeline, contains all measurements of sources detected in input images. Each input image is, of course, taken of a particular sky region at a specific time (or pair of times, in the case of difference images), and in a single filter band. The task of the Association Pipeline (AP) is to associate sources from different times, filters, and perhaps sky positions, into data structures that describe astrophysical objects. Having made these associations, the AP can make further measurements on the full object data to generate astronomically useful quantities such as:

Proper motions and parallaxes of stars

Recognition that an alertable transient event has occurred

Classification of variable stars

Phased light curves for periodic variable stars

Orbital elements for solar system objects

The architecture of the AP splits naturally into three components:

An associator for rapidly moving objects, based on sources detected in difference images (Rapid Mover Association Pipeline – RMAP).

An associator for all other objects, driven principally by sources detected in single images, with some assistance from difference image sources as well (Slow Mover Association Pipeline – SMAP).

An object measurer/classifier, which runs asynchronously with the other two components (CLAP).

The AP is certainly the most challenging component of the DMS from a conceptual point of view. In large part, this is due to the probabilistic nature of the problem. Detections in the Source Catalog may have resulted from a real astrophysical source, or from noise. Depending on the parameters employed in the DP, particularly the choice of detection thresholds, false sources may greatly outnumber real ones. Changing levels of atmospheric seeing will make closely spaced objects separately detectable at some times, and inextricably merged at others. In addition to the noisy nature of the input, the association process itself is inherently probabilistic when moving objects are involved. Without detailed a priori knowledge of a moving object's trajectory, deciding whether or not a moving object detected at some position tonight is the same object detected at another position last night is inherently probabilistic.

Unlike the IPP and the DP, which are not concerned with behavior over time, the AP must make use of multiple detections to reduce to negligible levels the propagation of noise into the Object Catalog. One of its functions is to continually "groom" the object catalog, pruning associations that subsequent information has shown to be incorrect. To do this, it must frequently look backwards in time at the whole history of an object, as well as processing new source detections that are arriving from the DP.

To date we have focused our design efforts for the AP on the rapidly moving object problem, since we consider it the most computationally stressing. We present a baseline plan for the RMAP in Section 3.4.3.4.1.

4.4.3.4.1 RMAP

Associating moving sources is a significant computational and algorithmic challenge. At the depth of a single LSST observation there will be on the order of 107 Main Belt asteroids and 104 NEOs visible across the survey area (together with 105 TNOs and several thousand comets). The rate of motions detectable with LSST range from fractions of an arc-second over a period of a year for proper motion studies through to degrees per day for NEOs. Algorithms and data structures for defining associations between observations must be able to account for this dynamic range in velocities.

Figure 4.4.3-5 provides a basic flow diagram for the analysis of moving sources within the RMAP. Given a pair of observations, separated by a short amount of time, rapidly moving sources (e.g. Main Belt Asteroids and NEOs) will be identified within the Source Catalog entries derived from subtracted images. This initial pairing will result in a position, time and velocity for each candidate track (with an associated error). Utilizing a database of known asteroid orbits these tracklets will be associated with existing orbits and, possibly, pruned from further consideration as new tracks. At this point the analysis becomes one of linking tracklets to form candidate tracks. This must be achieved for a series of visits extending over a lunation or across lunations. Naturally the major challenge in achieving this linkage comes from the combinatorics associated with identifying all possible tracks (naively it will scale as Nk where N is the number of sources identified as moving and k the number of observations).

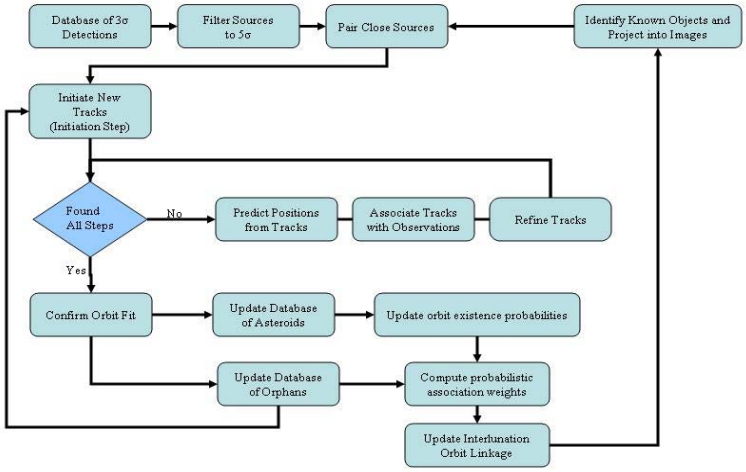


Figure 4.4.3-5 Rapid Mover Associate Pipeline

Once tracks have been initiated, pruned and merged (to exclude overlapping tracks) the resultant candidate orbits are confirmed using orbit fitting algorithms (three pairs are required to constrain the orbit) and the tracks accepted or returned for further track initiation. Given the observational constraints described below, singleton tracklets or orphans, where a moving source could not be associated within a track, must be stored within the association database for linkage across lunations and merging with existing orbit fits.

While the process flow is straightforward, observational conditions and constraints provide substantial challenges in achieving a high degree of accuracy. There is a fundamental trade off between completeness and accuracy. For example, assuming tracks are linear we can tune the algorithms to achieve a 98% completeness in our detections (i.e. we identify most asteroids), but at the cost of only 2% of the detected tracks being real (i.e. there is a large false positive rate in our linkage). As the testing of the candidate tracks using orbit fitting software is a rate limiting step (e.g. gOrbit) this is clearly a suboptimal solution. In contrast quadratic approximations to the tracks result in the same level of accuracy within only 16% of the tracks being false. Observational conditions impose the additional problem of dealing with censored data. Variations in atmospheric seeing, tumbling of asteroids, and the 5_ detection limit for the data will result in sources moving in and out of the detection threshold. Linkage algorithms must be able to account for missing observations, singleton observations (where we have no velocity information) in a way that does not bias the efficiency of the orbit finding. For example, the efficiency of a simple forward prediction or multi-hypothesis testing algorithm can be sensitive to how the data are taken (i.e. missing early observations makes it more difficult to determine accurate tracks than missing observations later in a sequence). Joint forward and backward search techniques might alleviate this problem.

Associations need to be undertaken in a probabilistic sense where we account for the priors on acceleration and velocity when linking tracklets. Current approaches used in extant surveys (ref mlinear) will not scale to the data rates of the LSST. Data structures such as kd-trees and ball-trees do, however, provide a means for achieving efficient nearest neighbor searches in O(NlogN) time. Combined with gating of the possible tracks (to limit the number of tracklets searched based

on the likelihood of a track being physical) we expect that it will be possible to initiate and merge tracks at a rate consistent with the LSST's processing timeline. Efficient mechanisms for associating known asteroid orbits with new candidate tracks, providing fast ephemeris calculations for predicting where an asteroid will be at a given observation time and for identifying sources lying close to a quadratic orbit are required to provide the input to the association pipeline for matching with known asteroids. Data structures capable of dealing with observations without velocity information and with time as a variable that can be applied in an iterative fashion for matching singleton or orphan observations are required in order that the pool of unmatched observations does not become prohibitive and so that LSST is sensitive to very rapidly moving sources. Finally, efficient, robust and fast orbit fitting remains a challenge for confirming a series of tracks as coming from an orbit. Current algorithms are slow and are designed to find the best fitting orbit rather than to cull those tracks that are not physical.

A second part of the RMAP will verify and improve the moving object catalog entries over time, and enable a wide range of queries described further below. As new observations are made and new links among them are proposed, our understanding of orbits, their probability of existence, and the associated data structures need to be updated. The tracking process described above will yield many good candidate orbits, as well as some unmatched singleton observations. The steady stream of candidate linkages and orbits from the tracking process above needs to be integrated into the moving object catalog. Our goal is to have the system evolve over time to behave as though it has closely approximated the probability distribution over orbits, given the data seen so far (accounting for the distribution of moving sources as a function of position on the sky). To accomplish this, candidate orbits that are likely the same object will be merged across lunations and these data representations updated within the moving object database. Likewise, as existing orbital candidates become less likely due to missed observations, their probabilities of assignments are transferred elsewhere. In the extreme case, orbits will be deleted. Notice that such updates lead to additional opportunities to link singletons, as they may gain partners from observations freed from fictitious orbits recognized as such.

Key algorithms and data structures for this part of the association pipeline include: fast selection of candidate objects that appear in an image, accurate projection of candidate orbits into observation space, probabilistic matching of projections to observations, linkage models based on orbital parameters. It is clear that the machinery developed to solve the above challenges will provide the capability to efficiently satisfy a large number of queries related to moving objects (e.g., "which objects will be seen in a given region of the sky, in a specified time frame, with confidence higher than some threshold?") To achieve this we need to develop a framework for very general time domain and moving source queries that will facilitate a better understanding of the existing orbits. We propose that the capabilities developed here are also exported as services, both to other data consumers in the LSST project, other astronomers, and the public at large.

4.4.4 Middleware Layer

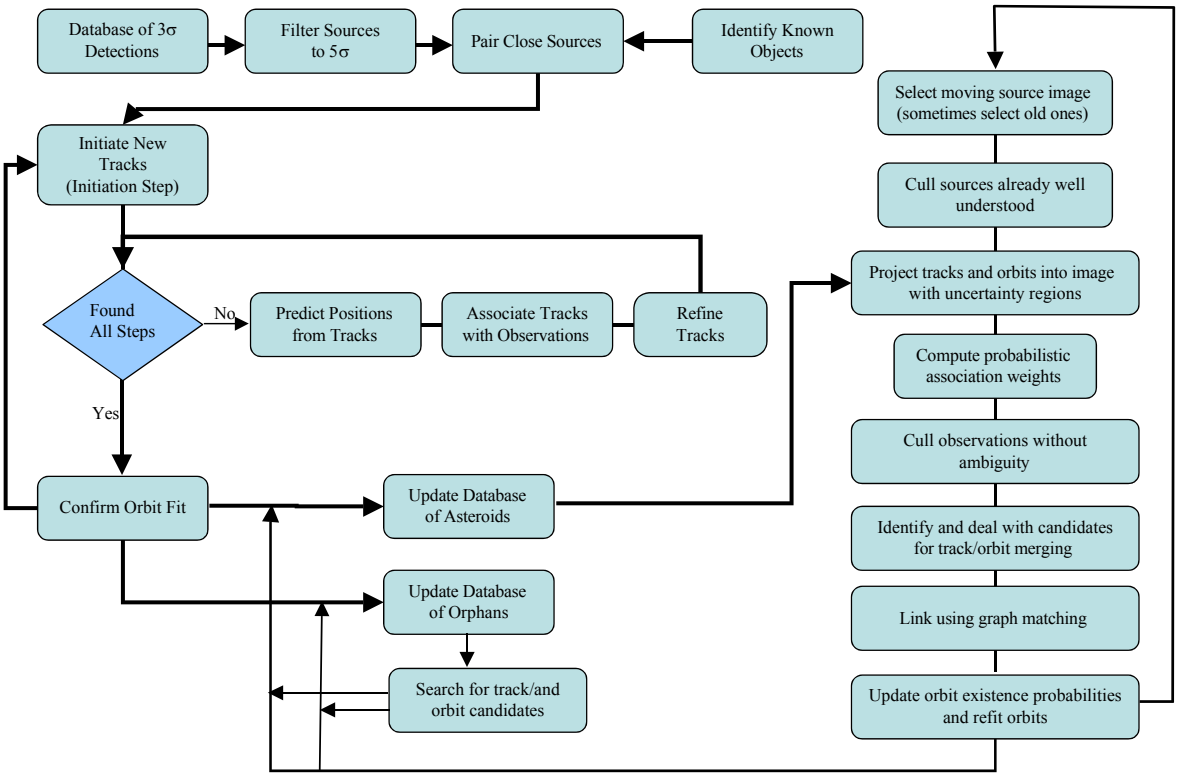


Figure 4.4.4-1 Overview of Middleware Layer

The middleware layer is between the application and infrastructure layer. Its purpose is to ease the applications' interactions with the infrastructure first by hiding many of the hardware details. This can be important for providing a common view on heterogeneous components of the infrastructure (whether that heterogeneity arises by initial design comes about due to the evolution of the hardware over time. More important thought, a middleware interface can provide applications with a logical view of the infrastructure based on the logical operations an application needs to perform—like, "get me this data" or "execute this pipeline"—which engage the infrastructure. Not only does this ease application development, it allows infrastructure interactions to be optimized in a variety of ways for the particular logical operation.

In addition, a middleware layer can provide a standard set of common or shared application functions, encouraging consistency, improving reliability, and reducing implementation effort.

A relatively low level example of middleware is the C++ Standard Library, which provides reusable data structures and algorithms that are needed by most applications, such as container and queue structures and memory management. A large percentage of new C++ production-quality applications utilize this library, so many in fact that it has become an ANSI standard.

A higher-level example of middleware is the now ubiquitous Message Passing Interface (MPI) middleware for parallel processing applications. This middleware is now standard on virtually every available commercial cluster environment and has been used for thousands of parallel applications across many scientific domains. It provides standard services for hosting, data definition, inter-process and intra-cluster communication, and various forms of execution control and synchronization.

An even higher-level example of this layering is the ESO's Common Pipeline Language (CPL) that is specifically oriented at astronomical data processing. CPL enables algorithms to be constructed as plug-in dynamic libraries and executed in sequence using a pipeline recipe, which is a text configuration file. Various data types and data access methods (for example, one is for FITS files) are supported by CPL so that the plug-ins can pass data through the library.

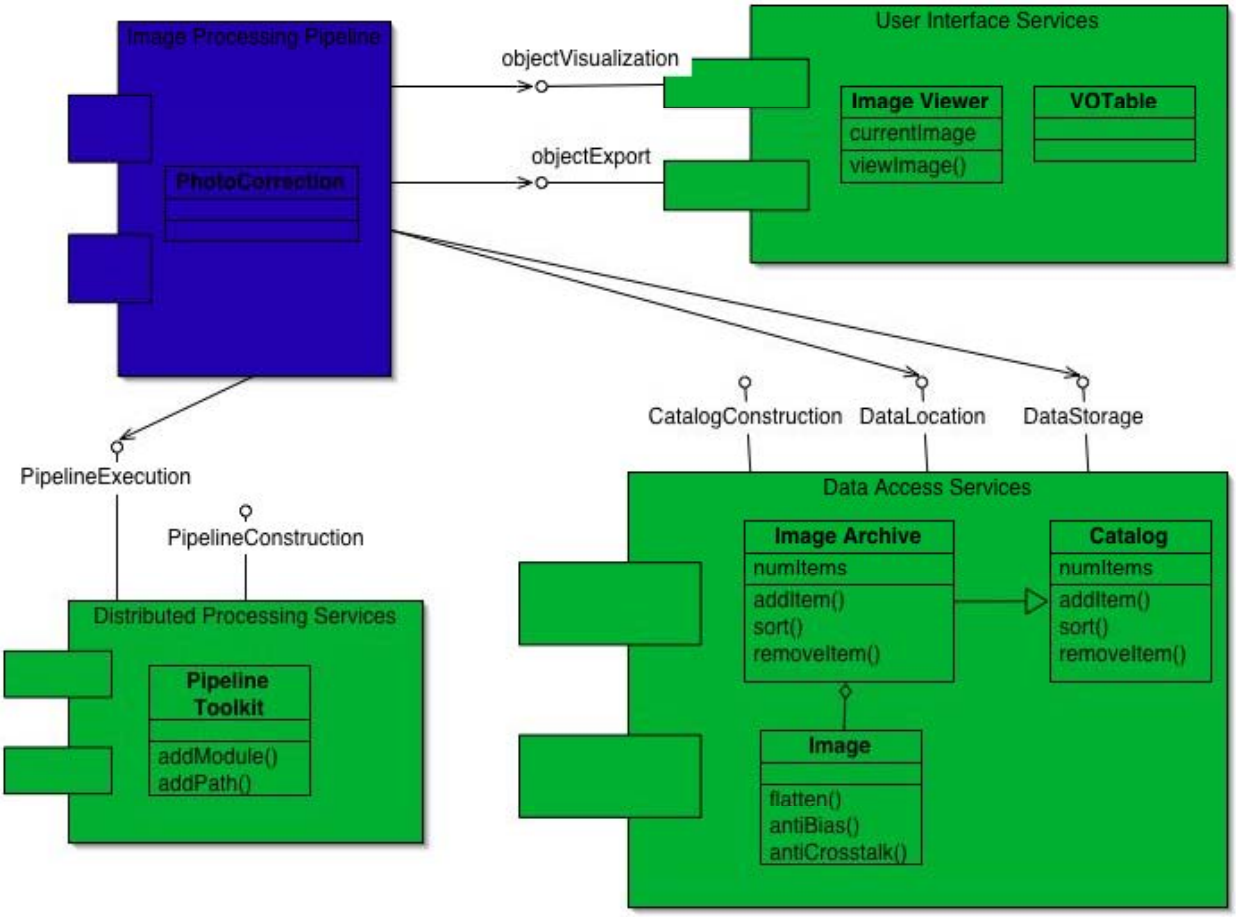


Figure 4.4.4-2 The Middleware Component Diagram

LSST Middleware is broadly divided into 3 components as depicted in the UML Component Diagram in the Figure above. Each component contains objects within it (the 3-section boxes) that implement the component. Each component publishes interfaces (the "lollipop" symbols) that are accessed by other components. The Application Layer components (e.g. the Image Processing Pipeline in blue) access the middleware layer via these interfaces. Note, the above diagram is representative only, the actual components will contain many objects and other sub-components.

The overall requirements for middleware are presented next, followed by a section on each major component.

4.4.4.1 Requirements

The requirements of the Middleware Layer derive from our overall baseline requirements for data management, many of which address the need for timely processing of large data volumes.

In particular, the baseline requirements for performance—especially the low-latency required for supporting alerts—push the middleware layer to require a high degree of automation and fault tolerance.

Coordinating processing and data flow across the three types of LSST facilities (mountain/base, archive centers, and mirror data centers)—necessitates a high degree of flexibility in where and how the processing pipelines are deployed. For example, we will need a more precise understanding of the processing needed for generating alerts, the amount processed data produced, and the effective bandwidth achievable between the base camp and the archive center in order to decide whether all subsequent products are transferred to the archive or whether some products are recalculated after transfer.

Furthermore, our baseline requirements call for the need to periodically reprocess all data as our software algorithms improve. Because of the extensive computational power required to do this, the data will not necessarily be reprocessed on the same platforms as the original real-time pass. Thus, it is important that a pipeline be easily run at any of the LSST sites.

In order to support the level of performance, automation, fault tolerance, and flexibility needed, the middleware layer needs to:

- support parallel processing on high-performance platforms as needed to meet performance requirements. Parallel platforms could include distributed-memory and/or shared-memory architectures.
- provide common interfaces to storage, communications, and system facilities that are independent of platform location.
- provide transparent access to pipeline input data regardless of location.
- enable transparent access to all output data products across LSST sites.
- provide real-time monitoring of execution of pipelines including automated detection of and recovery from faults.
- provide mechanisms for re-configuring and redeploying pipelines—in particular, change where the particular processing takes place—with minimal effort.
- allow a pipeline to be paused in order to change module parameters as well as allow the pipeline to be restarted without extensive repetition of previously executed processing steps.
- allow portions of pipelines to be executed on selected sets of data without major reconfiguration.
- provide component interfaces that allow both local and secure, remote control of the data management system.

In order to establish a flexible data management system as well as promote reusable components, we envision creating Distributed Processing Services environment that provides a common set of facilities, including:

- adaptive and prioritized process management/control/scheduling
- workflow management
- monitoring, error handling, and failure recovery
- automated ingestion of data products
- data and process packaging and deployment
- secure access control
- data access for distributed processing including files, structured objects, and database query results (via Data Access Services),

Flexibility also refers to how the pipeline processing modules that encapsulate the scientific algorithms are connected together: not only do we need the ability to connect modules in different ways, we need to be able to swap out modules with more modern implementations as they are

developed. For this reason, we envision an integrated framework for developing pipelines and pipeline modules. In particular, this framework must:

- support the ability to statically (i.e. not during execution) insert a new module in an existing process flow without recompilation of modules whose interface has not changed.
- provide support for both composition-time and run-time type checking of modules and data structures for compatibility.
- provides mechanisms for estimating the processing and storage workload represented by a given pipeline configuration and dataset to be processed.
- provide mechanisms for integrating a data-driven approach to pipeline execution management (e.g. GridDb).

Several of the general middleware requirements listed above bear specifically on the role of the Data Access Services in processing the data and delivering it to the community at large. Another important aspect we need to consider (as with the Distributed Processing Services) is how we allow the Data Access Services to evolve. While especially important during the initial construction and commissioning phases of the telescope, we expect the evolution of the system to continue at some level into the operation phase as well; thus the Data Access component must allow for changes. In particular, it should be possible to:

- define new data types, including new science-related entities such as object classifications, new meta-data entities, and new data relationships.
- extend catalogs to include new data types and concepts
- integrate new types of storage hardware
- support new data access patterns
- define new access control policies

Finally, we need a set of user interfaces for driving the data management system. We envision two functional classes of interfaces: administrative interfaces for driving the creation of standard LSST data products and end-user interfaces for data access and remote analysis. On the user-side, we will need interfaces that allow users to:

- submit queries against all LSST catalogs and receive results in a research-ready form
- browse LSST data products through standard astronomical views or visualizations
- create “best” images of selectable regions of the sky in a research-ready form (e.g., FITS format, with access to ancillary images such as varience, etc.)
- download any of the released LSST data products, from raw to stacked, as desired, subject to the users’ computing capabilities (e.g., network bandwidth, storage)
- track the provenance of a data object, such as a transient detection in a catalog, back through the processing used and to the original raw data from which it was generated.
- perform selected analysis operations remotely
- integrate user-generated query or analysis modules into the LSST processing infrastructure at LSST facilities, subject to security restrictions and available LSST resources.

In general, administrative interfaces will need all the capabilities of the user interface; in addition, administrators will to be able to:

- start, stop, and configure standard processing pipelines
- monitor pipeline status and analyze errors and failures
- flexibly visualize data via techniques appropriate to the data type
- manage the replication of data across LSST facilities as needed.

In general, these interfaces need to be flexible to modification to accommodate not only adding new capabilities but to present different views for different audiences. This need for flexibility calls for the use of toolkits that support modular, “pluggable” views. We expect the Web to be

the dominant platform for user interactions; given the distributed nature of our project—that is, distributed not only in terms of facilities but also people—the Web may be important for administrative interfaces as well.

4.4.4.2 Distributed Processing Middleware

The LSST application layer is composed of 3 major elements:

- data products (images, catalogs, and alerts)
- pipelines to process raw data into data products
- interfaces to allow access to the data products and pipelines.

This section focuses on the second element, pipelines and the underlying middleware needed to support the pipelines. The LSST distributed processing middleware handles the creation, scheduling, and execution of the pipelines described in section 4.4.3 on the infrastructure described in section 4.4.5. It is a framework that enables pipelines to run efficiently and consistently on the LSST distributed systems, and that maintains this level of service while algorithms and systems undergo continuous improvement.

4.4.4.2.1 Critical Issues

There are several critical issues in distributed processing middleware that must be addressed in order to meet LSST requirements for pipeline functionality, performance, reliability, and extensibility.

4.4.4.2.1.1 Pipeline Extensibility and Algorithm Modularity

Pipelines are collections of loosely coupled algorithms that are applied to a stream of data in successive stages. Algorithms have interfaces (inputs and outputs) that must be rigorously defined in order for the middleware to determine which interfaces can be "plugged together". The middleware provides a programming environment that allows algorithms to be represented to the middleware as objects containing interfaces and methods. The objects, which can be developed and tested independently of the pipelines, can thus be thought of as replaceable software modules.

4.4.4.2.1.2 Portability

The distributed processing middleware must be adaptable to hardware infrastructure and operating system software that may differ between sites and evolve over the project life cycle. To achieve this adaptability, the middleware interfaces with native parallel resource managers (also sometimes called middleware) through an abstraction layer that allows this software to change without major impact on the rest of the system.

4.4.4.2.1.3 Fault Tolerance

With large distributed systems comes an increased component failure rate. The middleware provides a level of fault tolerance for pipelines. If a cluster node that is executing an algorithm on a particular data item in the stream fails, the middleware reruns this computation on another cluster node without stalling pipeline data flow. Other more difficult to handle failure modes such as software failures, network outages, and storage system failures will also be considered for incorporation into a fault tolerance strategy.

4.4.4.2.1.4 Provenance

Recording and understanding the history of a dataset—the conditions under which the original data was taken and the subsequent processing that was used to bring that data to its current state—can be a complex proposition. Further, LSST provenance tracking is made drastically

more complicated by the expectation that we will periodically recreate data products when improved algorithms become available.

Good provenance records that allow us to determine which algorithm versions were applied to a given set of data could in principle help us to avoid re-doing processing steps unnecessarily; however, the full complexity of the provenance tracking problem could generate large volumes of data and a high computational load. Even if the information could be easily tracked, we must be sure that we can provide a clear, straightforward, and self-consistent picture of the state of the data to the user community. This may require strategies such as regular scheduled "releases" of the entire LSST data collection and "retirement" of deprecated versions in order to keep the community's understanding of provenance simple and manageable.

4.4.4.2.1.5 Security and Access Control

With the extensive, valuable data products that will be created by the LSST, it is critical to manage access to the data, or potentially suffer a catastrophic loss of data or data integrity. Without this control, the LSST could be subject to inadvertent data corruptions or malicious hacks. We must also manage access to computing and communications resources, or potentially be overwhelmed by the processing workload.

The distributed processing middleware is integrated with the data access middleware to provide uniform security/access control to LSST data products and compute resources. As we consider the ways we might support remote processing and analysis by users, the issue of security begins to extend beyond protection from undesired uses of the system to providing access control for different classes of users. This access control will be implemented in various levels using the concept of service tiers described in Section 4.4.1.3.

4.4.4.2.2 Distributed Processing Middleware Approach

To accomplish the above, the distributed processing middleware provides several elements:

- an application programming interface (API) to algorithms that enables them to function as replaceable modules in pipelines
- a language for expressing structures that describe the input and output streams of data between the pipeline modules and data and control flow across the modules
- tools that describe the available execution infrastructure (computing resources, storage, and networks), map a pipeline run on a specific dataset (i.e. a "job") onto the infrastructure, schedule the job for execution, monitor the job, and clean up after job completion.

4.4.4.2.2.1 Pipeline Module Application Programming Interface

The pipeline modules must all provide a standard interface that allows the entire pipeline to operate in an integrated way, including:

- moving data and control information between modules
- synchronizing parallel parts of the pipeline
- querying/logging execution status
- error handling and recovery.

The level of detail needed in the interface definitions of modules (e.g. the tightness of the binding) and the amount of module framework infrastructure that can be reused in LSST middleware without impacting performance is an area for further research. Several current technologies that are listed in Appendix X Candidate Technologies will be considered.

4.4.4.2.2.2 Pipeline Definition Language

With algorithms wrapped as objects with well-defined interfaces, pipelines can be constructed using a pipeline definition language that expresses the data and control information flow between algorithms. The pipeline definition language could be a functional data definition language with algorithms represented as composite functions, such as that proposed by GridDB, or it could be an imperative scripting language. It must enable pipelines to be easily extensible and reconfigurable, and implement type safety to ensure interfaces between algorithms are compatible.

4.4.4.2.2.3 Parallel Execution and Scheduling Capabilities

To achieve the required processing throughput, we expect to employ a variety of execution strategies that takes into account the different modes of parallel processing we will need. The simplest of these modes capitalizes on the independence of data in the pipeline data stream and executes multiple copies of the algorithms on multiple data when possible, to maximize throughput. This is called "data parallel" execution.

In “algorithm parallel” execution, the algorithm is decomposed and parallelized as needed to reduce the time needed to process a single data item. This may require that each instance of an algorithm be executed on multiple CPU's using a message passing or shared memory parallel programming paradigm.

These two different modes tend to be launched in different ways and may benefit from being launched on different types of platforms. In general, we expect to have to manage chains of execution at a high level. We plan to leverage existing and emerging resource management tools from the grid community to assist in this.

Some LSST pipelines have fixed latency requirements. As a pipeline executes, statistics can be gathered on the time needed to run each data item through each algorithm, and this information can in turn be used to predict future execution times. The middleware can use this information to perform smart scheduling of multiple pipelines, guaranteeing that latency requirements are met, while efficiently utilizing computing resources. Parallel resource scheduling for pipelines, some of which may have fixed latency requirements, is an area of research.

Grid computing environments orchestrate pipeline workflow within a distributed execution environment. Pipelines that leverage these technologies are more portable, robust, and better suited to heterogeneous environments than their predecessors. These environments also provide advanced scheduling, security and access control, execution monitoring and logging, and fault tolerance in the distributed environment.

LSST will examine a wide range of these technologies, such as those described in Appendix Y Candidate Technologies, against the middleware requirements.

4.4.4.3 Data Access Middleware

The purpose of the data access middleware is to provide consistent, efficient mechanisms for accessing large amounts of data that are distributed across multiple sites. In fact, LSST requirements to support the camera’s high data rate with low-latency processing for time-critical alerts are made more challenging by the fact that storage and processing of the data must occur at some level at all LSST facilities, from mountaintop to data centers.

Furthermore, processing at these sites cannot operate in complete isolation. For example, image differencing that occurs at the Base Facility site will often require previously processed data from the Archive Center. One lesson from scientific data grid applications today is that even when high-bandwidth networking is available, it is often difficult to take full advantage of that bandwidth, and applications quickly become I/O-bound. The data access middleware, then, must

be more than a set of tools and APIs for moving data through the facilities; it must be an environment for coordinating strategies that enable high data throughput.

Though implementations may share common APIs and storage mechanisms, data access middleware generally falls into two classes: access to files and access to databases. Typically, the former are used to store raw and processed images, while the latter are used to store structured catalogs of astronomical objects. Some middleware attempts to hide the storage type behind an abstract interface, but does so at the risk of implying equivalent performance in both instances, which is generally not true.

On the other side, there are generally two classes of customers for the middleware services: pipelines that must process the data and archive services that must deliver data to users. In both cases the capabilities of both the receivers and the senders can vary across LSST facilities and user community. Different users will have different access to network bandwidth and client-side software to handle the data. The same can be said of analysis pipelines running at different end user locations.

The commonalities across the classes of data, as well as how the data are delivered to clients, calls for a unified approach to a data access framework. We are currently prototyping such a framework, drawing on existing technologies from the grid and high-energy physics communities. The prototypes will allow us to explore three key strategies for meeting our requirements:

- *Parallel data streams*: we can apply a variety of techniques for parallel I/O that utilize multiple nodes for streaming data. These techniques include the use of parallel file systems, multi-threaded applications, and volume management. Refer to Appendix Y, Candidate Technologies, for a list of these technologies that we are evaluating.
- *Data caching*: as a survey instrument, we can predict to good measure what processing—and therefore, what data—will be needed on which platforms beforehand. To avoid long delays waiting for data to arrive from the archive to the processing platform, we can ensure the appropriate archive data is pre-cached whenever possible. For raw data streaming from the telescope, our caching strategies must take a comprehensive approach to managing all storage, including storage on processing platforms along with archival storage, to minimize the amount of copying that can slow the net data storage rate.
- *Adaptation to available capabilities*: through simple techniques employed in existing middleware, we can provide different views of the data and methods for getting it based on the capabilities (bandwidth, available storage, computational environment, and access to software) of the site receiving data. The operational model for LSST presumes that sites will be categorized into a standard set of "tiers" that define these parameters for typical usage patterns and capabilities.

The LSST will require Data Access Middleware that supports the following features:

- Collection-based access where files and collections of files can be referenced via location-independent identifiers. The middleware can resolve one of these identifiers to a list of physical locations or to the closest location. Collection identifiers can be used to initiate bulk operations on collections.
- Multiple protocols (e.g. http, ftp, gridFTP, srb) for data transfers. Clients can negotiate with a remote server to choose a supported protocol.
- Third-party transfers; automated and/or scheduled transfers.
- Ingestion of new data into the archive.
- Remote queries of database catalogs.
- Automated, fault-tolerant management of hierarchical storage.
- Support for VO-compliant interfaces that are critical to the LSST community.

- A security framework that supports a group-based authorization model.

4.4.4.3.1 Critical Issues

There are several critical issues in data access middleware that must be addressed in order to meet LSST requirements for data access functionality, performance, reliability, and extensibility.

4.4.4.3.1.1 Data Access Middleware Technology Assessment

In order to avoid prematurely locking into technology selections that will not be the best available, it is important to project the state of middleware technology in the 2007 – 2010 time frame when key technology choices must be made. Certain choices will need to be made in 2007 during Preliminary Design, in preparation for the Construction phase (e.g. the overall architecture for storing images and objects in file systems or databases); other choices will be deferred until Detailed Design, which occurs during the Construction Phase (e.g. the specific version of a database system).

LSST will be aided greatly in this assessment by remaining connected into the communities developing and using the relevant tools and thereby benefiting from the broader experience, particularly those that are grid-related (e.g. data transport, replica management, and authorization). Our job during research and development is to understand how these emerging technologies should be applied to our application. We will do this via a combination of analysis and prototyping.

Database technology must also be tracked and evaluated closely, because many of the performance issues are related to application-specific factors. Indeed, the problem of selecting a database technology encompasses a broad set of critical issues on its own. The choice of relational vs. object-oriented vs. “roll your own” must be balanced against the requirements for performance, scalability, fault tolerance, and long-term maintainability.

4.4.4.3.1.2 Data Architecture

Data structure and partitioning issues are critically important and drive the selection of data access technology. The way data is organized across storage devices can significantly impact system performance: inefficient data partitioning leads to 'hot spots' that slow search and retrieval.

Related to the issue of partitioning is the issue of indexing: the LSST data access middleware must satisfy both spatial and temporal queries across multiple catalogs, as well as other pre-defined and as yet undefined notions of proximity or association (e.g. classification/type, color and other physical characteristics). Duplicating large quantities of data is expensive; therefore, other approaches must be investigated. To achieve good performance, we will like have to look at a variety of strategies for using the different tiers of storage, from the storage on the processing nodes accessed directly by pipelines to nearby archive storage to the remote or slow archive storage

To address the partitioning question, it is critical to have a good sense of the most important and most common user queries and access patterns to be supported. We will employ the Iconix process and UML use cases to capture these queries and access patterns, extrapolating from current large surveys to those likely to be required for LSST. Representative queries include:

- Extract object time series (general variable objects)
- Extract objects by cone search (anything)
- Extract variable objects by type
- Extract variable objects by variability parameters (amplitude, period, ...)
- Cone-magnitude-color search (white dwarfs, quasars, brown dwarfs, ...)

- Extract galaxies + cone search (weak lensing)
- Extract proper motion objects (gal. structure)
- Extract parallax objects (solar neighborhood)
- Extract new objects (transients, SS objects)
- Extract transient object close to a galaxy object (SN)
- Extract color gradient objects (blue core galaxies, AGN)
- Cone + photometric redshift + shape (N-point correlation function)
- Find stars of nearly constant brightness (photometric standards)
- Identify objects indicating motion (to derive parallax or proper motion)
- Select multi-band photometry (for photometric redshifts)

The database schema must not only be optimized for the above usage, it must satisfy additional operational requirements specific to the LSST. The schema design can impact performance. What would be a minor deficiency in a smaller system can have a dramatic impact on system performance in a multi-petabyte regime. Even though the LSST will have many TFLOPS of computing power, typical relational approaches involving extensive inner or outer joins will not be possible in many cases, given the size of the catalogs. Techniques that have emerged out of data mining/data warehousing disciplines (n-dimensional indices, star schema, federation of multiple schemata into one virtual schema, etc.) are likely to offer some solutions.

The schema has to support complex science. The schema has to express dynamic nature of the LSST science, where the understanding of orbits/objects and the associated data structures need to be updated with time. The association of detections into objects (in a single image, in multiple images of the same field, and in images from multiple fields over time) will naturally mutate as more survey data is collected and algorithms improve. Altering existing associations in a destructive way will create problems, especially if done on a fine grained basis (individual objects) because users and application layer software will have difficulty adjusting to the large number of individual changes in the catalogs. The alternative is to "publish" snapshots of the catalog at periodic intervals, but this means users may not have all the latest processing results. Therefore, a blend of these two approaches will be required. Consequently techniques for data immutability and versioning as well as data provenance become requirements.

The schema/persistency has to be shielded. While it is undesirable to evolve schema, there will almost certainly be cases impossible to dismiss. To allow schema changes without impacting user applications, persistency details including the schema need to be shielded from users and application layer software.

4.4.4.3.1.3 Scalability

Database scalability is a particularly difficult but critical issue. Simple, well-known approaches to providing efficient search do not scale to the levels demanded by LSST. For example: the size of an index for a petabyte of data is likely to reach 2 terabytes, which means a single sequential scan at 100 megabytes/sec would take over five hours. The latest techniques to address scalability problems of today's large databases include data partitioning, index partitioning and query parallelization. While database vendors have started to incorporate these features into their products, to the best of our knowledge no system was designed with petabytes in mind.

Whether we are considering databases or file collections, we need to understand what APIs we need for data access. As described earlier, the data access API goes beyond the simple *get* and *put* operations. The distributed nature of the LSST system complicates access patterns, for example when Base Facility processing requires old observations from the archive in order to calculate a difference image.

Furthermore, the obvious choices for accessing catalog data will depend greatly on the type of database (relational, object-oriented, or other) we build upon. Finally, there is the question of

how our data model will be represented in the API. Should images and catalogs be treated as fundamentally different products with different APIs, or should we, for example put "everything in the database" and access it with a single API?

4.4.4.3.1.4 Security and Access Control

The appropriate security model is also a critical issue. As a public archive with access to high-performance resources, we could easily imagine becoming a target of malicious users, so at a minimum we must prevent unintended usage. On top of that, we need to support different levels of authorization for accessing data and services not only within the archive but also across the various tiers of LSST user sites.

4.4.4.3.1.5 Implications for Infrastructure and Low-level software

The choices we make at the middleware level can have important consequences for the hardware infrastructure (e.g., storage systems, clusters, network configurations) we need to support along with the low-level software needed. In particular, as we noted above, a successful parallel I/O system requires multiple storage nodes to stream the data in parallel.

For example, NCSA is prototyping different configurations for storage clusters that specialize in delivering data to pipelines and users. The choice of file systems is also affected as well. Multiple approaches to managing parallel access range from true parallel files systems like Lustre and IBRIX to file management systems like xrootd and the Storage Resource Broker (SRB). Refer to Appendix Y Candidate Technologies for a list of file systems that we are evaluating against LSST requirements.

4.4.4.3.2 Data Access Middleware Approach

In this section we describe the components of a prototype data access framework and highlight some existing and emerging technologies that can support it.

4.4.4.3.2.1 Data Collections and Identifiers

This framework calls for all non-catalog data, both raw and processed, to be organized into logical, hierarchical collections. Every collection and every dataset has a unique identifier associated with it that is made up of the identifier of its parent collection appended by a local name. This is analogous to paths to directories in a file system; the difference, of course is that the identifier need not imply an actual physical location on disk. Collection identifiers, then, can be used to access all datasets in a collection as simply and efficiently as accessing a single file.

This can also reduce the overhead of accessing large numbers of files by eliminating the need by an application to make database calls or directory listings to determine what data is needed. Datasets, then, should be organized in such a way that files that will commonly be processed together (in time, but not necessarily by processor or platform) should share a common ancestor collection.

Multiple collection levels provide greater granularity—and thus, flexibility—for accessing the data. Such flexibility is critical for supporting effective parallel I/O. Consider the case of single-frame calibration in which each detector image can be processed independently on a different processor; it should be easy from an application or pipeline management level to send an entire frame (via its identifier) to a cluster and yet transparently distribute the detectors one-to-a-processor. Making the detector images accessible as separate FITS images (or some equivalent content object) provides the greatest flexibility to how the data is distributed.

4.4.4.3.2.2 Replica Location Service

This component provides a service that is used to resolve a logical, location-independent identifier (URI) into physical locations accessible through some protocols (URLs). With a list of such URLs, clients have the flexibility not only in choosing where to retrieve the data from, but also what protocol to use. We expect typically that this kind of decision making can be built into the middleware layer itself so that applications can, for a collection ID, get back a unique list of URLs to retrieve the individual datasets from, perhaps characterized by the anticipated retrieval performance.

In general, the replica location service would be used to track datasets stored across LSST facilities, from mountaintop to the tiered access sites. Two leading grid technologies support replica management: the Globus Replica Location Service (RLS) and the Storage Resource Broker (SRB). Replica management with such tools can provide a natural way to connect clients with parallel streams as the URLs can point, for example, to different nodes in a storage cluster; however, other volume management tools, like xrootd, can be quite effective for enabling parallel streaming from a single site.

4.4.4.3.2.3 Authentication and Authorization Framework

At any time, the data management system will be handling data that are in some part publicly accessible and some part restricted. While in principle we expect most if not all LSST-generated data products to be made public "immediately", in practice, such data will likely require a period of time to allow for data verification. We do not want to release incorrect data to the community.

In addition, different user communities represent different access levels or tiers, in terms of the bandwidth and computational resources they have at their disposal. Finally, we cannot expect all user profiles to be administered centrally by a single entity, as this would be onerously cumbersome.

The grid community has honed solutions for authentication based on X.509 certificates that work well across administrative domains. Such solutions work well when users are not expected to have their own account on the individual grid machines to begin with. Less well-developed but emerging are solutions for authorization to enforce access control policies.

Both SRB and Globus address the problem of access control by assigning policies to groups of users. With the Globus Community Authorization Service (CAS), policy management is handled within a centralized service, reducing the amount of information about users a particular data service needs to manage. The Shibboleth framework, used more widely in the digital library community, has similar components as CAS and may be useful to LSST. The challenge then becomes determining the best way deploy these solutions across the LSST facilities.

4.4.4.3.3 Data Replication Tools

We envision four modes of data replication across LSST facilities and to end users:

- *Downloading data to the user*

Typically this mode would happen through an LSST portal or via VO interfaces. To support this mode, we need to address efficient ways to retrieve whole collections and make the best use of available network. The BIMA Data Archive's Data Retrieval Tool (DaRT) is an example of a helper application that helps the user manage collections of data from an archive and even schedule the download for a later time.

- Cache-triggered replication

This mode would be used to move data from the telescope to the archive sites and beyond automatically; in this mode, the arrival of data at one LSST site (say, in some watched data cache) triggers its replication at another site.

- On-demand send and retrieve

This mode would be used by pipelines to get the latest data for processing. It would also provide the underlying replication mechanism for the other modes.

- Scheduled, third-party transfers

This mode would be used to cache data on pipeline platforms prior to when it is actually needed. Supporting all these modes is fairly straight-forwarded when handling files; however, the real challenge is enabling replication of catalog data stored in databases.

4.4.4.3.3.1 Ingest and Metadata Services

Data products are either created by the telescope or by processing pipelines; in both cases, we need a mechanism to ingest these products into the archive as they become available. Ingest usually involves copying data to safe storage and extracting metadata for loading into the archive's holdings database. Once the products are ingested, applications need a way to not only discover them but also retrieve their metadata (e.g. to plan and configure processing pipelines).

4.4.4.3.3.2 Catalog Access

This component provides users and pipelines the means to access catalog data. It is likely to be backed by a database management system (DBMS) capable of supporting a petabyte-scale database. For users, a VO-compliant interface (e.g., OpenSkyNode described later in this section) should be sufficient; however, pipeline systems will need a more tuned, high-performance interface.

No DBMS is capable of meeting the LSST needs today, and it remains uncertain whether there will be any in the timescale of the next few years that will deliver the scalability, performance and fault tolerance needed. Given the uncertain future of object oriented database market, a relational database system is currently favored. Another reason to focus initially on a relational technology is existence of promising open source database engines, like MySQL.

4.4.4.3.3.3 Relevant Virtual Observatory standards and interfaces

The Virtual Observatory (VO) will play an important role in distributing data to the community. Because the VO is largely about interoperability, supporting standard VO services will be key to enabling science requiring synthesis of LSST data with data from other archives. The key standard services that exist today and enable cross-cutting research are the OpenSkyNode and the Simple Image Access Protocol (SIAP).

The former provides interoperable access to astronomical catalog data. In particular, it enables efficient joins—particularly object cross-correlation—across distributed catalogs. This standard allows us, for example, to cross-correlate objects detected by LSST with 2MASS over the network.

SIAP provides a uniform interface for retrieving images from an archive. These can be static images, as well as images created on-the-fly to the users specifications of such things as region, projection, resolution, filter, and even observatory-specific parameters that might govern the manner of co-addition.

Two additional standards are emerging that will be important for LSST. VOEvent is a standard for publishing time-critical, astronomical discoveries, such as supernova detections. In fact, LSST is helping to drive this standard by providing requirements and assisting in the design of the VOEvent interface.

VOStore will govern how users can interact with remote write-able storage in an interoperable way. Given the data volumes involved with LSST-based research, greater emphasis will be placed on remote analysis services through a portal. Remote storage, where users can store such

things as results from database queries to specially processed images, is critical all but the most trivial portals. An important aim of the VOStore standard is to make it possible for users to gain a global view of all storage from multiple observatory portals and enable cross dataset comparisons and synthesis.

Refer to http://www.ivoa.net and http://us-vo.org for more information about the Virtual Observatory.

4.4.4.4 User Interface Middleware

The User Interface Middleware facilitates our view into LSST data, analysis, and operations. As described in the requirements listed in 4.4.4.1, we recognize two classes of interfaces to the LSST Data Management system: the administrative interface for operating the system and the end-user interface for making use of the LSST data products. In general, the former includes the latter. The interfaces must address the fact that LSST community—be they scientists using the data or administrators running the observatory—is distributed, so remote access to services is necessary. Web access is certainly paramount to users, but it may be important for administrative interfaces as well. The functionality we need to expose is multi-faceted and complex; thus, web access will need to be organized into a portal.

It is often useful to classify some functionality that lies just below the user interface as part of the User Interface Middleware. This can include various the aspects of managing a portal, including user databases, session management, user-accessible persistent storage, and—most importantly—documentation management for learning how to use the interfaces and the products they access. Fortunately, numerous tools from the IT community for managing portals continue to emerge.

4.4.4.4.1 Critical Issues

4.4.4.4.1.1 Mitigating the Cost of Developing User Interfaces

User Interfaces are perhaps the most costly class of software component to develop well. Often it is difficult to construct a user interface right the first time before the underlying functionality is complete (and perhaps exercised for common use patterns); thus, it can easily get left to the end of a project when costs and schedules are tight. One strategy we plan to pursue is to attempt to leverage VO standards and the large number of existing and emerging VO-ready end-user tools to minimize what we have to develop ourselves. This strategy, though, begs its own critical issue, described next.

4.4.4.4.1.2 Setting the Scope of LSST User Interface Development

It is hoped that we will not have to develop basic visualization tools. There are many that exist today that cover the variety of visualizing we expect to need for LSST data, and many such tools understand standard VO Data Access interfaces. Thus, we trade off the expense of developing the end-user tool for supporting a standard interface. The unanswered issue then is to what extent can this trade-off be applied to all areas of user interactions. Is it sufficient to provide a public programming interface (e.g. Web interface) and leave it up to the community to bring the tools? How can we engage the community to build tools for LSST products and services that may not yet exist?

4.4.4.4.1.3 Types of Interfaces for Administrative Interactions

The people that will operate and monitor the LSST data management system will be widely distributed; remote monitoring, assessment, and control in one form or another will be necessary. Web browser-based interfaces will likely be the best choice for some of these interfaces—

particularly monitoring. On the other hand, a Web browser-based interface for controlling data management operations raises issues of security and development costs. We will need to understand what will be the best mix of command-line tools, GUI-driven applications, and browser-based interfaces.

4.4.4.4.2 User Interface Middleware Approach

As described above, the LSST intends to make maximum leverage of standard interfaces and data interchange formats provided by the International Virtual Observatory Alliance (IVOA), including those described in section 4.4.4.3. In particular, we plan to collaborate directly with the US-based VO effort, the National Virtual Observatory, to help drive the development standards that will aid the community's access to LSST products. (A current example of this is the VOEvent standard for publishing transient event detections.) By supporting these standards, our community will be able to leverage the growing body of VO-compliant end-user tools to do such things as access images, query catalogs, and visualize results.

While VO tools may well cover some of the more complex yet necessary end-user interface issues, it is likely that we will have to implement some more basic interactive services that are specific in some way to LSST. This would likely include basic archive browsing capabilities, including those for drilling down through the provenance of a data product.

The end-user access points will be gathered together through an LSST portal. The level of sophistication of this portal will depend on the extent of interactive interfaces we will need to provide directly (as opposed to what we rely on VO tools to provide). As appropriate, we will study the use of portal management and content management tools. Today, there are two emerging standards for portal management [names]; leveraging a standard may be important for hooking in community generated tools.

On the administrative side, we will need to build a variety of tools for monitoring and controlling data management system. An important subset of these will be tools for testing and validating the pipeline software and data products. Web browser-based tools are especially well suited for monitoring and quality assessment, and the web often provides a convenient way to integrate the use of visualization tools. However, control interfaces will likely be manifested as a combination of command-line tools that must be run at the site where the system being controlled is running and GUI-applications. The latter can potentially communicate remotely with a system via web services. For this type, we expect to need to employ a GUI toolkit that provides common look-and-feel.

4.4.5 Infrastructure Layer

The infrastructure layer on which the application and middleware layers are deployed for execution contains the computing, storage, and networking hardware together with the systems software needed to utilize it. In this section, we first describe overall requirements and the method by which we can trade off between these technologies to achieve an LSST DM architecture that optimizes cost, reliability, and availability, i.e. our ordered project priorities. Next, we describe the trends in those technologies and how we are employing them in the tradeoffs. Finally, we describe a baseline architecture that, if the technology trends hold, represents such an optimization.

4.4.5.1 Requirements

This section describes the requirements that must be satisfied by the infrastructure, assuming the operational model and facilities described in Section 4.4.1.2.

- System level requirements

- The allocation of processing and storage to facilities will be done to optimize (in priority order) cost, reliability, and availability of the total data management subsystem
- The infrastructure will be sized such that the net throughput of the data management pipelines will permit complete processing of a night's observing data prior to the start of the next observing night, assuming no system outages. Anticipated throughput by pipeline is:
  - Data Acquisition pipeline generates raw images as 16-bit integers at 4Gbps
  - Image Processing pipeline generates 32-bit floating points
    - Calibrated images at 8 Gbps (36 TB/night)
    - Noise maps at 8 Gbps (36 TB/night)
    - Subtracted images at 8 Gbps (36 TB/night)
  - Detection pipeline generates raw object information at 0.1 – 1.0 Gbps, depending on object density in field, and chosen detection threshold. (0.45 - 4.5 TB/night)
  - Association pipeline generates linked objects into catalog at 0.05 – 0.5 Gbps, depending on object density and observing cadence (0.225 - 2.25 TB/night)
  - Deep Object Detection pipeline generates objects into catalog at (TBD)
- The infrastructure will be sized such that complete re-processing of the entire raw image dataset may occur once per year without interrupting observatory operations
- The infrastructure will be provide for temporary storage for a minimum of 150% of the mean time to repair of any communications network link at the source end of that link
- The infrastructure will be sized such that after outages, "catch up" processing of the temporarily stored raw image data may occur at the rate of one night's observing data processed per day, without interrupting observatory operations

- Mountaintop site requirements
  - Provide sufficient infrastructure to support observatory operations that operate in real-time and engineering and calibration activities, including the Data Acquisition Interface and the Camera, Telescope, and Observatory Control Systems
  - Computing and storage equipment will be preferentially located at base versus mountaintop due to lower support costs and fewer reliability issues (lower altitude), therefore any processing that can be done in either location will be allocated to the base site
  - Provide two nights of raw data storage in the event of Mountain to Base network outages
  - Provide infrastructure to perform data quality assessment on a percentage (<10%) of the data large enough to ensure optimal observatory operations.
  - Must be co-located with observatory

- Mountain to Base network
  - Must be highly available, Mean Time Between Failures (MTBF) > 90 days
  - Must be highly reliable, Mean Time to Repair (MTTR) < 24 hours
  - Provide secondary link or transport mechanism for minimal operations support in the event of extended outage

    - Prefer link to be owned and operated by LSST and/or the operations entity due to responsiveness of support

- Base site
    - Provide sufficient infrastructure to support rapid-turnaround processing for image processing, detection, and association pipelines, and generation of all time-critical data products, i.e. alerts
    - Provide at least three nights of raw data storage in the event of Base to Archive Center network outage
    - Prefer location at existing facility to leverage existing support and facility resources

- Base to Archive international network connection
    - Must be highly available, MTBF > 180 days
    - Must be highly reliable, MTTR < 48 hours
    - Provide secondary link or transport mechanism for minimal operations support in the event of extended outage
    - Prefer link to be leased from public or commercial entity due to high cost of support for long-haul link

- Archive Center
    - Provide infrastructure to support pipeline processing and reprocessing, permanent storage for all data products (with provenance), and source of data for replication to data centers and end user sites
    - Provide disaster recovery support preventing loss of LSST data in the case of infrastructure or facility-threatening events
    - Prefer hosting at existing NSF/DOE-funded supercomputing center due to expense of creating new facility
- Data Center
    - Provide infrastructure to support storage for a subset of LSST data products (with provenance) for Tier 1 access by end users
    - Provide access to LSST data products via VO services in accordance with IVOA/NVO standards
    - Prefer hosting at broad range of facilities to permit widest possible access to LSST data

4.4.5.2 Tradeoffs in the LSST DM Infrastructure Architecture

The baseline infrastructure architecture provides a structure within which one can develop the tradeoff model for cost, reliability, and availability. The accuracy of this model is dependant on analysis that will occur repeatedly during the entire R&D period to validate the cost/performance trends of the technology. It will be updated to reflect the current "most likely" scenario and serves to provide a departure point for further investigation. At the key points of Construction phase planning/proposal development and start of development during the Construction phase, the model will be re-baselined.

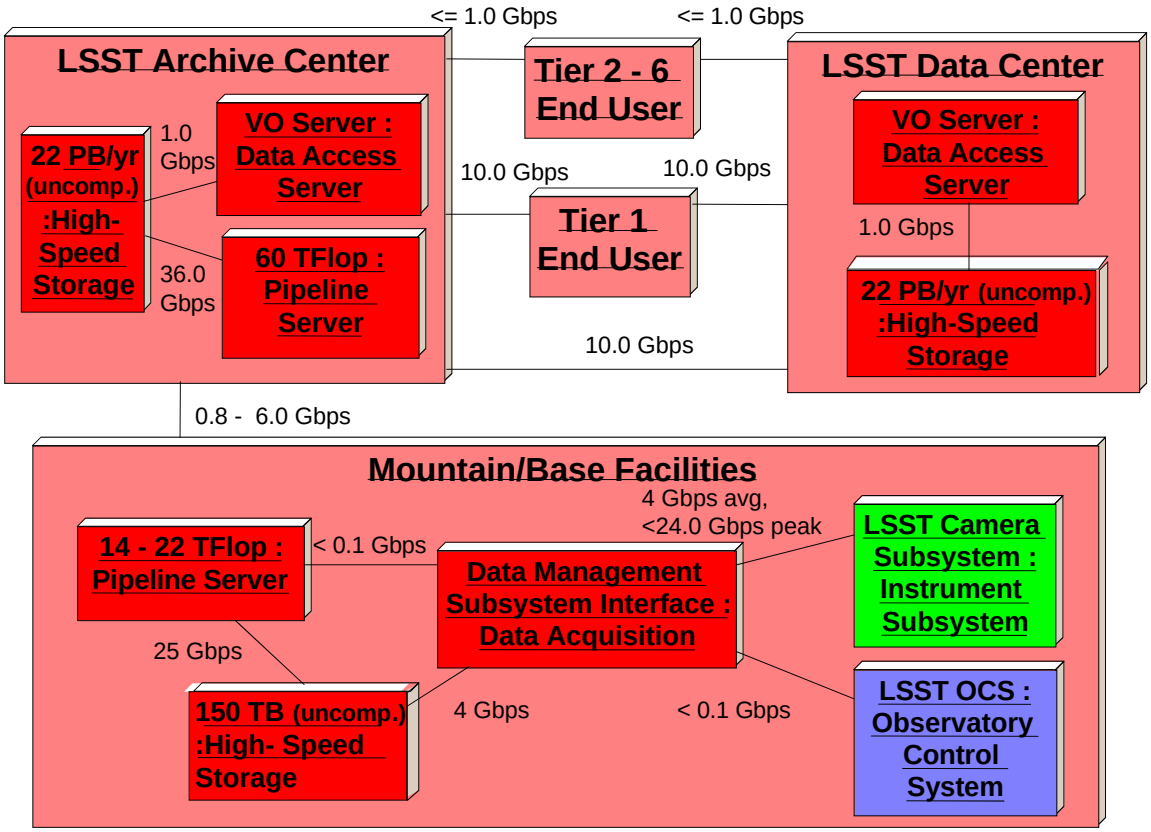


Figure 4.4.5-1 LSST DM infrastructure with tradeoff points

Figure 4.4.5-1 shows the principle components of the infrastructure and the potential range of the computing, storage, and network elements, depending on the results of the multiple tradeoffs involved in finalizing the design of the data management system. These tradeoffs are further elaborated below.

4.4.5.2.1 Real time data quality analysis and alerts: locating computing and storage at the mountaintop versus base facility versus archive center

The principle real-time requirements on the LSST data management system come from the need for extremely rapid data quality analysis (DQA) and transient alerting. These analyses must be performed in as close to real time as possible, the former to ensure that the control system can react within the time before the next exposure is finished, that latter to ensure that follow-up can be done in a timely fashion.

The current operational model of two 10 second exposures approximately every 30 seconds, with 6.4 gigabytes per image, provides a real-time data rate from the Data Acquisition Interface of 12.8 gigabytes approximately every 30 seconds. Refer to the system level requirements above for the pipeline throughput required to process this data and generate real-time alerts. Full analysis of this large data stream requires many teraflops of computing power, as well as associated storage capacity.

The optimal location of this computing power and storage depends on tradeoffs between cost and latency time. The closer the processing is to the telescope and camera, the lower the latency for DQA and alerts will be. However, complex computer systems have reduced reliability at high altitude. Also, the cost of operating such systems at a remote site, such as a mountaintop, is several times that of operating them at a well developed site, such as either a base facility or

archive center, given not only the power and infrastructure requirements but also (possibly more importantly to ongoing operations) the manpower required to support this complex infrastructure.

Given the limitations on the mountaintop computing environment, as well as the availability of highly reliable technologies for implementing a multi-gigabit/s link between the mountain and the base, it is envisioned that limited processing of the images and engineering data will take place at the mountaintop. The bulk of the computational processing of the received data will be split between the base site and archive center, with the real-time alerting portion being done at the base site.

4.4.5.2.2 Data reduction scenarios: compute and store versus on-the-fly re-compute

There are multiple drivers in determining the split in computation work between the base and archive center, including:

- Latency requirements for real-time alert generation
- Sufficient, reliable, and cost-effective network access to the base facility to support the end-user science requirements.
- Cost of power, cooling, and floor space for a supercomputer and multi-petabyte data storage system at the base facility compared to the archive center.
- Cost and availability of a large, dedicated, supercomputer system support and administration staff at the base facility compared to the archive center.
- Cost of redundancy, to provide for continued operations in the event of computing and storage failures, at the archive center versus the base site.

The computing and storage requirements of the archive center would optimally be co-located at an existing supercomputer center. This leverages the existing infrastructure of computer room floor, power, cooling, high-speed networking, and expert staff of system administrators and maintenance people. The cost-sharing benefits are enormous compared to procuring an LSST-specific building and hiring, training, and keeping a full-time, dedicated staff.

There are multiple scenarios for the handling of derived data and data products, depending on whether they are computed and stored, or computed on the fly each time they are needed, or some combination of the two. The tradeoff is still being evaluated and is dependent upon a final LSST DM approach. The two extremes are explained below; the final architecture will almost certainly be in between.

4.4.5.2.2.1 Compute-intensive Extreme

In this approach, LSST DM would archive the minimum data sets (ie, raw images, sums, and object catalogs) on disk and re-compute most everything else that is required as needed. Only raw data and metadata are transferred from the base facility to the archive center. The data are then reprocessed at the archive center to produce all of the high-quality data products to be delivered for the LSST scientific analyses and archive end users.

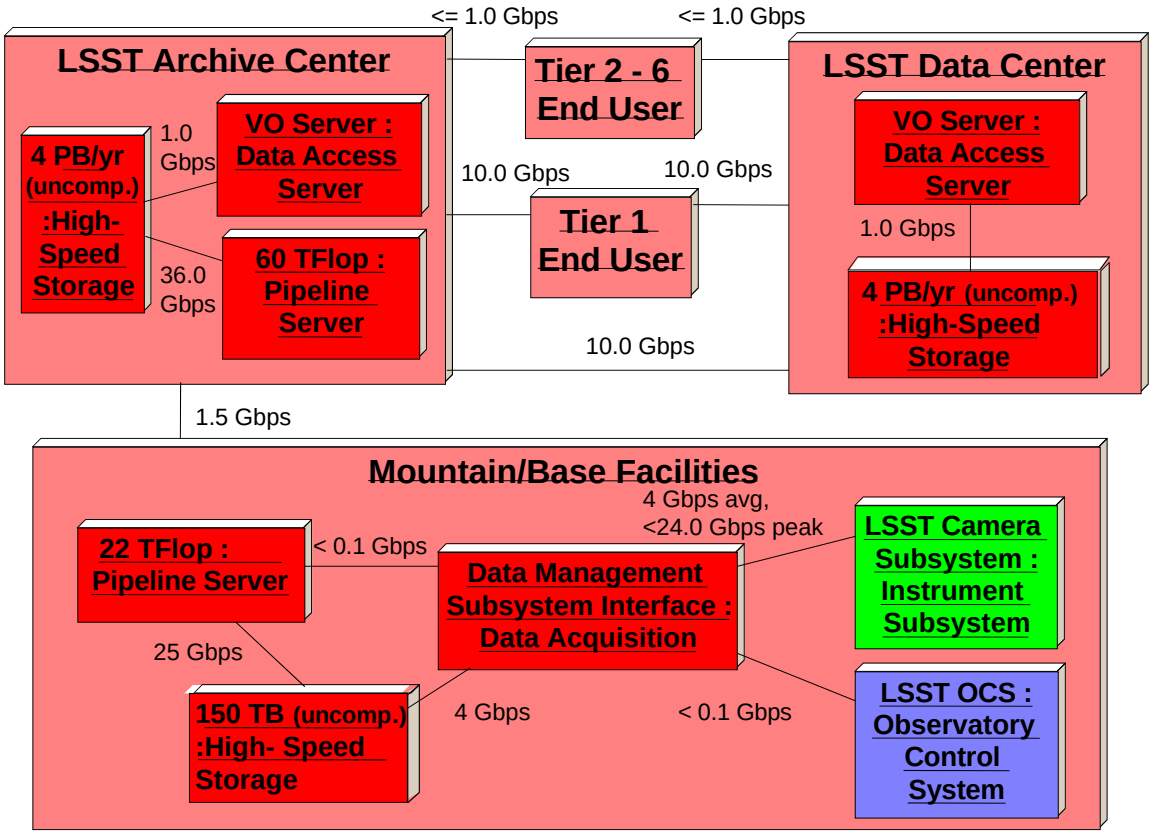


Figure 4.4.5-2 Compute-Intensive Architecture

This scenario represents the lower network bandwidth limit; the connectivity from the base to the archive center is sized to allow for one night's raw data to be transmitted to the archive during the following day (within 24 hours). In order to provide for rapid DQA (beyond that which might be done on the mountain) and transient alert generation, this processing must be occur at the base facility and the alerts must distributed from there.

The second processing pass at the archive center is not subject to the real-time latency requirements since it does repeat the distribution of the alerts that were already generated at the base. Also, since the real time processing must use library calibrations, while to obtain the highest quality data products of one night of data, one relies upon calibrations derived from that night's data that can only be calculated after the full night of data is available. Some of the data from the base facility reductions may also be sent to the archive and vice versa to provide a form of error checking, both for the transmission of the data and also to identify possible divergence between the base facility and archive pipelines.

Thus, significant computational and storage resources are necessary at the base facility, amounting to roughly 22 TFLOPS of processing capability and at least 150 terabytes of storage. Very little contingency in processing is needed, since once the window for alerts and real time DQA has passed, there is no need to catch up in these areas. The storage requirements are driven in this case by the needs of the real time processing combined with some contingency in case of problems sending the raw data to the archive center.

The base facility system will need to store all of the relevant catalogs and reference images needed for alert generation. Updates to these catalogs will have to be sent to the base facility from the archive center, along with reference images produced in further processing of the data at the archive center (e.g., the cumulative sum and/or optimized template images). As the network

is full-duplex, the required network bandwidth will be defined by the larger of either the transmission of the raw data from the base facility to the archive center or transmission of this information back from the archive to the base facility, not the sum of this traffic.

The archive center will require roughly 3 times the base center computing capacity, to permit repeat processing each night, plus re-processing for improvements in data quality and algorithms. The minimal archive center data storage needs are on the order of 4 to 5 petabytes for the data products each year. However, care must be taken to insure that sufficient bandwidth to disk exists to satisfy the processing needs imposed by all the processing requirements, ranging from real time processing to user inquiries, all simultaneously. These use-scenarios are still under active research, but the results of these analyses will be used to provide specifications for the compute needs, storage, and bandwidth needs.

Finally, in this scenario, there is no bandwidth allocated between the base and archive to support science inquiries from users; this will all be done using the archive center as the host on other network channels. All VO-related inquiries will be run from the archive center and data center sites.

4.4.5.2.2.2 Network/Storage-intensive Extreme

The network/storage-intensive alternative is to store nearly everything that is computed. This changes the required system to support 6 gigabits/s in the base to archive link and a much more massive database at the archive center. Under this scenario, it is estimated that the DM pipeline will store about 22 petabytes of useful data and data products per year, including reduced images, difference images, summed images, and a wealth of catalogs and other data products.

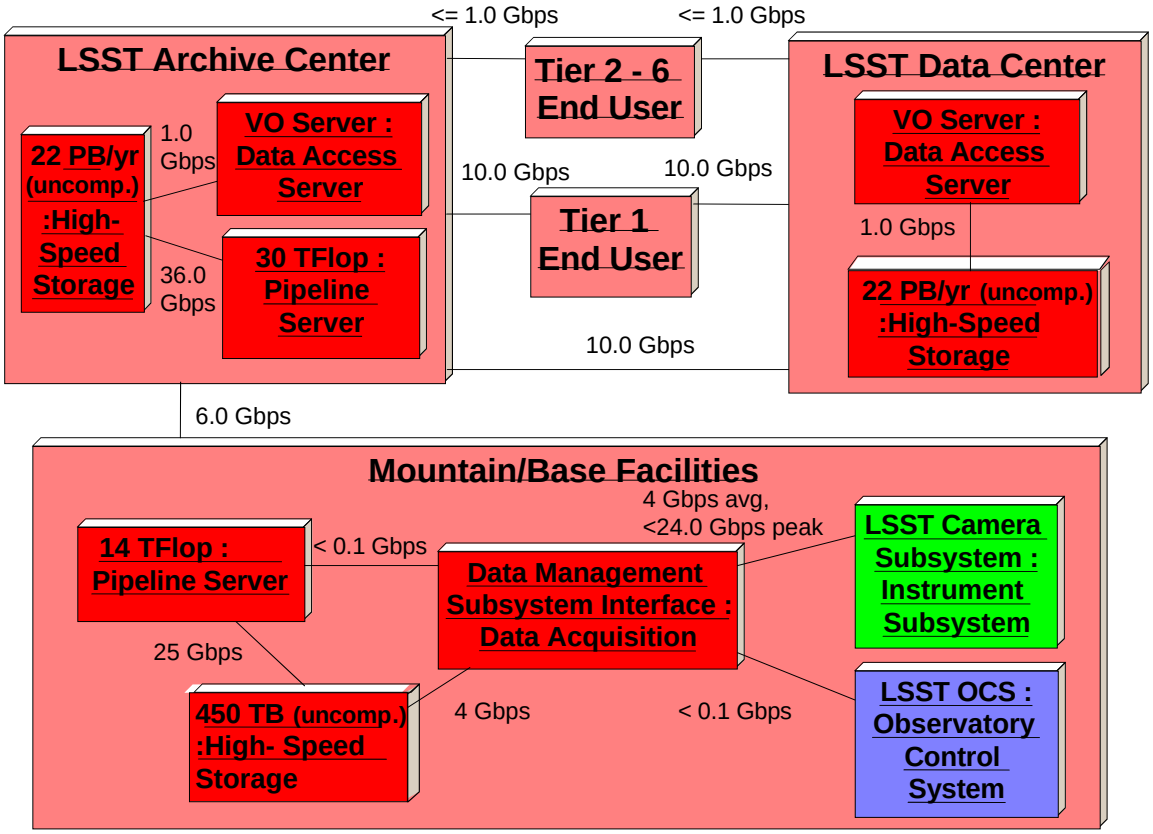


Figure 4.4.5-3 Network/Storage-intensive Architecture

The higher data rate between the mountain and base camp allows the bulk of the computation already performed at the base to be avoided at the archive center, since the data products are now transferred to the archive center. The archive center must only be sized for the increased storage and re-processing, approximately 30 TFLOPS.

The advantage of this approach is more immediate access to all data products by end users and application pipelines. There are two perceived disadvantages to this approach:

- The cost/performance trend of computing resources is increasing more rapidly than either long-haul network bandwidth or memory-to-disk i/o rates
- The second processing is not performed, therefore additional processing must be performed to detect errors due to transmission from the base or problems in the base pipeline processing.

Given that both processing capabilities and storage technologies are advancing rapidly and to some extent at similar rates, and that both are moving faster than high-speed wide area networking and memory-to-disk i/o rates, it is unlikely that either of these two extremes will provide the optimal solution to the science requirements of both the LSST survey science and the archive end user needs. Detailed analysis of both the end user usage patterns and balance between technological advances in CPU and storage options will be required to determine the solution which provides the optimal performance for the LSST data management system.

4.4.5.2.3 Communications paths: implementing the networks

An analysis of networking paths from each of the proposed LSST observatory locations has been conducted, to identify the key initiatives that might provide the bandwidth required in the two scenarios above. This analysis is summarized below.

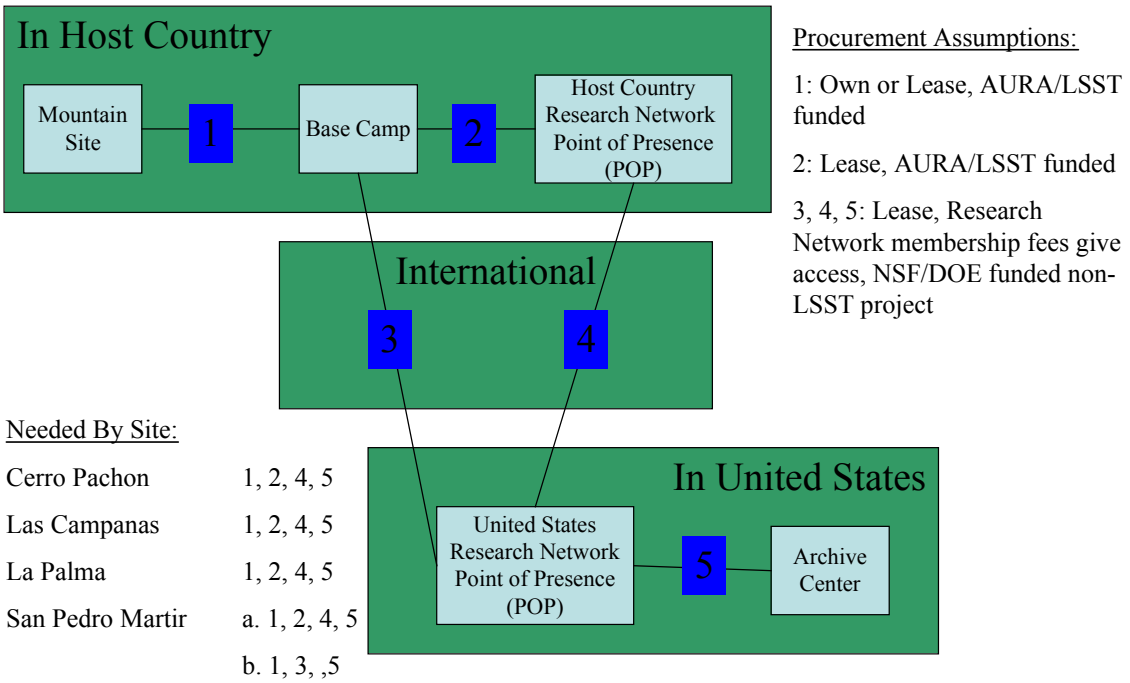


Figure 4.4.5-4 LSST Communications Links by Site

Figure 4.4.5-4 shows the communications links and procurement assumptions for the end-to-end communications from the observatory back to the archive center. Table 4.4.5-1 lists options for each link.

Table 4.4.5-1 LSST Communications Links Options

Site	**Link 1 Primary/ Backup**	**Link 2 Backup is integrated in network**	**Link 3 Backup is integrated in network**	**Link 4 Backup is integrated in network**	**Link 5 Backup is integrated in network**
Cerro Pachon	4 Gbps Fiber Optic to La Serena 80 km Own/ 0.5 Gbps Microwave Own $1.5M Total	4 Gbps Fiber Optic La Serena to Santiago Entel, telefonica, or REUNA Lease	Not Applicable	4 Gbps Fiber Optic Santiago via Miami to Chicago AMPATH/ LambdaRail Non-Project Lease	10 Gbps Fiber Optic Chicago to Archive Centers ABILENE /internet 2 and TeraGrid Non-Project Lease
Las Campanas	4 Gbps Fiber Optic to La Serena telefonica 25/200km Lease 0.5 Gbps Microwave telefonica Lease $1.5M Total	Same as above	Not Applicable	Same as above	Same as above
La Palma	4 Gbps Fiber Optic to Santa Cruz 20km Lease 0.5 Gbps Microwave Lease $3M - 4M Total	4 Gbps Fiber Optic Santa Cruz to Madrid Lease	Not Applicable	4 Gbps Fiber Optic from Madrid to Chicago CLARA or DOE/Atlas/CERN Non-project Lease	Same as above
San Pedro Martir option a	4 Gbps Fiber Optic to Ensenada 100km/200km TelNor Lease/ 0.5 Gbps Microwave TelNor Lease $3M - 4M Total	4Gbps Fiber Optic Ensenada to Tijuana TelNor Lease	Not Applicable	4 Gbps Fiber Optic Tijuana to San Diego AMPATH/ LambdaRail Non-Project Lease	10 Gbps Fiber Optic San Diego to Archive Centers ABILENE /internet 2 and TeraGrid Non-Project Lease

San Pedro Martir option b	Same as above	Not Applicable	4Gbps Fiber Optic Ensenada to San Diego ABILENE/internet 2 Non-Project Lease	Not Applicable	Same as above

As the table indicates, there are numerous leased and research network options that have or will have the capacity to support the LSST data transmission requirements. It is essential that LSST make these requirements clear to the appropriate funding agencies, to ensure that they are taken into account in long-range capacity plans.

4.4.5.2.4 Continuous operations and data safety: reliability, availability, and failover modes

The driving requirements in this area are the need for almost continuous operations in order to not get behind in data reduction and the need to ensure that no data are lost. To meet these needs requires not only high reliability and availability systems, but also backup emergency operations modes to provide service in the event of failures. Such backup modes may have limited functionality with respect to the baseline system. The tradeoffs here involve risk mitigation and provision for at least basic functionality of the system during possible failures. This area is a significant cost driver, as the more stringent these requirements, the more robust and redundant the architecture must be.

Although the baseline operational model is scoped to provide minimal downtime, any given segment can and almost certainly will develop problems during the lifetime of the project. As such, backup measures must be identified to provide for all of the functionality described above. The solutions for storage and processing are well known and common practice in the large scale computing communities, ranging from fault tolerant storage such as RAID systems to redundancy in storage and processing units with automatic failover. With standard practices common today, the risk of losing data can be minimized at both the base facility and archive facility.

The network links, however, are more difficult to duplicate for redundancy and therefore must have failover strategies to provide for effective data transport in the case of unexpected downtime. As the data flows from the data acquisition component, it would be cached in storage sufficient to provide for several nights of observations in case the network link from the mountaintop to the base facility was unavailable. The bandwidth of this network segment would be sized with a contingency to allow the data management system to catch up relatively quickly in the case of a brief outage (minutes to several hours).

In the event of a major interruption in mountain to base facility communications that couldn't be repaired in the MTTR of <24 hours, manual transport of the cache storage medium (such as shipping disks down the mountain) would provide for data transport to the base facility on less than 24 hour timescales. Clearly this places additional constraints on the mountain "cache" system (that it be transportable), but existing technologies (e.g., hot swap disks and USB style connections) already meet these requirements.

In any case, only those transient events that have a shorter window than the recovery period would be affected, all other science would still be achieved. Also, in both the short and long outage cases, a slower backup link (microwave or other lower speed connection) could provide bandwidth for an interim period. This link would most importantly provide continuous connectivity for command and control of the LSST system, such as observing database updates from the base facility and/or archive center. In addition, this backup link could be used to transmit some fraction of the data in real time, providing for some limited alert capabilities and minimizing the backlog of data to be transferred.

While this risk mitigation is necessary, if LSST owns and operates the mountain to base facility network link, the additional measure of having trained staff able to effect repairs provides for minimal downtime in the case of major faults such as cut fibers. Similar measures of in house repair are more difficult to put in place for the longer base facility to archive center network segment, given the cost and complexity of running what will almost certainly be a long (>100 km), international network connection. However provision for manual transport of large storage units is also practical for this link, via FedEx or similar courier service. While the turnaround time might be >24 hours, leading to some backlog of data processing, the archive center is sized to provide for the necessary additional processing to rapidly recover from the backlog. Alternative lower bandwidth communication channels are also necessary, in addition to the manual transport, to provide for flow of status and related command information.

The possibility of such network outages also impacts the tradeoffs related to the location of computational resources among the sites. In order to continue observing during such an outage, at least some computational capabilities must be sited on the mountain to provide the necessary real time DQA. Thus while both full DQA and alert generation may be impacted, observations could continue during either a short or long interruption of network access between the mountain and the base facility.

Similarly, the risk of network outages between the base facility and the archive center produces tradeoffs in the site of the more advanced data processing required for full DQA and alert generation. If the base facility has the computing resources necessary to perform this data processing, alerts can be generated during a network outage, and a much smaller backup bandwidth is necessary to distribute only the alerts to end users throughout the world.

In the case of a network outage with the minimal base facility described above, the archive center would not receive the data and could not generate alerts until the connectivity was restored. Once restored, however, the archive center would utilize the re-processing capacity described previously to catch back up relatively quickly.

4.4.5.2.5 Data access: extracting the dataset needed from multiple petabytes

The computational resources devoted to servicing user queries and related computational loads must be added to those for data reduction, and are likely to take a different form, that is they must be optimized for rapid search and retrieval of data, even as the source data is updated.

Current production examples of petabyte scale, file oriented datasets that are optimized for query and search (e.g. the Google File System and computing infrastructure) utilize large quantities of commodity processors and a great deal of replication to achieve high performance.

Research is also underway to implement parallel queries for petabyte scale database systems, using multi-tiered architectures with clusters as the primary hosts. LSST is likely to need both types of data access systems for query and external interfaces, and therefore both types of data storage infrastructure may be required. R&D tasks to define the access patterns and map those to storage technologies will inform the architecture in this area.

4.4.5.3 Technology Trends

With the rapid pace of cost/performance improvements in computing, storage, and network technology, it is most prudent for LSST to defer selections in these areas until the latest possible time.

This section presents those trends as they are manifested today. This analysis will occur repeatedly during the entire R&D period to validate the cost/performance trends of the technology. It will be updated to reflect the current "most likely" scenario and serves to provide a departure point for further investigation. At the key points of Construction phase planning/proposal development and start of development during the Construction phase, the model will be re-baselined.

4.4.5.3.1 Processor and Architecture Trends

This analysis is based upon the 2002 CERN PASTA report and the Semiconductor Industry Association *International Technology Roadmap for Semiconductors* (ITRS), dated 2004, plus conversations from key vendors.

Typically, advances in processor speed and power are driven by market demand that has remained steady for the last decade. We should expect to see a slowing of demand for the highest possible performing CPUs in the desktop and home markets. Gaming systems and professional high-end workstations will continue to drive CPU development, but obviously they represent a much smaller market. The big market push over the next few years is not in the desktop arena, but will be in the mobile and wireless markets.

The ITRS report covers all of the relevant technologies for the semiconductor industry; several of them are especially relevant to projecting trends for future LSST computing systems. While the density of transistors will continue to increase, the heat that a CPU chip dissipates will quickly reach a limit that has far reaching impacts on logic design and chip architecture.

The highest performing CPUs and, therefore, the hottest CPUs this year dissipate 167 Watts. The 2003 ITRS projected 200-Watt parts appearing in 2008 and growing to 300 Watts in 2018. One year later the ITRS update significantly changed the power dissipation to cap it at 198 Watts from 2008 through 2018. This 10 year long cap on power dissipation reflects the industry view that air-cooling is about to reach a physical limit and other cooling technologies, such as liquid impingement or spray cooling will be neither cheap nor attractive to the bulk of the market. Several mainline chip vendors have said that clock speeds will peak out in the 5-6 GHz range because of the limit on heat dissipation. The ITRS report also noted that cooling the chips is becoming a serious issue for system designers.

The densest CPU parts will grow from about 300 million transistors today to 1.2 billion transistors in 2009. Since the power dissipation of a transistor is proportional to the square of its switching frequency, and power is capped at 198 Watts, one cannot build chips containing more transistors switching at ever increasing clock frequencies. To stay under the maximum heat dissipation of the chip package, more widespread use of clock gating and other power-minimizing architectural techniques will take place.

One technique for staying within the maximum heat dissipation of the chip package is multi-core chips. Instead of faster and faster clocks, with their associated squaring of power consumption, major CPU vendors are moving to multi-core chips where multiple copies of a processor occur on each chip. This approach permits an effective use of the higher transistor accounts available through 2015 and beyond, with only linear scaling of power demands. IBM is already shipping dual core chips in BlueGene/L and their high -end server line. Intel will have dual core processor chips in 2005 and AMD in 2006.

An alternative technology being aggressive explored by industry to evade the heat dissipation limit is simultaneous multi-threading (SMT). Intel refers to this approach as Hyper Threading Technology (HTT). CPUs with SMT have sufficient redundant hardware to allow the simultaneous execution of 2 or more processor threads. The real purpose behind this technique is to avoid excessively stalling the processor while it waits for data from memory. This approach provides a 30% performance increase without generating significantly more heat. Multi-core and SMT/HTT are not orthogonal technologies; the current Power 5 chip from IBM is a dual core chip with each core implementing simultaneous multi-threading.

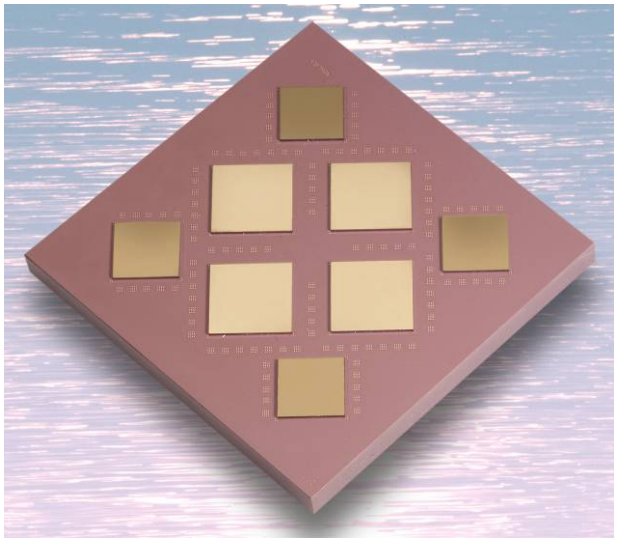


Figure 4.4.5-5 IBM Multiprocessor module with 4 dual-core Power 5 CPUs and 4 L3 cache chips.

The LSST data pipelines must run in parallel to meet critical data quality analysis requirements. The programming models to effectively use multi-core or simultaneous multi-threading are not significantly different so either technology or even a mix of both technologies will run the data pipeline efficiently.

There are several recognizable disruptive technologies emerging that can have a significant impact on the computing architecture chosen for both the base and archive center sites. System on a chip (SOC) systems are available today. The BlueGene/L system is a perfect example of this technology. A fully integrated, dual core chip with cache and all interconnects on a single die, the addition of 9 DRAMS makes a complete computer.

The IBM design goal was highest possible compute density in a cabinet, not the fastest possible CPU chip. The compute density, in terms of GFLOPS/foot^2 is 2 orders of magnitude better than any other supercomputer. The energy consumption of BlueGene/L, measured in GFLOPS/kilowatt is also 2 orders of magnitude better than other supercomputers.

The current one-half BlueGene/L at Lawrence Livermore is 3 times faster than the second fastest computer in the world, and will shortly be 6 times faster when fully installed. A single BlueGene/L rack today is rated at 5.7 Teraflop/s; it is expected that a follow-on system in 2008 will exceed 15 Teraflop/s in a cabinet. The following table compares the BlueGene/L system on a chip approach to more traditional clusters of SMPs and vector supercomputers.

Table 4.4.5-2 Current supercomputer performance

Effectiveness Measure	ASCI White	ASCI Q	Earth Simulator	BlueGene/L	Units
Memory-Space	8.6	17	3.1	140	Gigabytes/m^2
Speed-Space	13	16	13	1600	GFLOPS/m^2
Speed-Power	12	7.9	4.0	300	GFLOPS/kW
Speed-Cost	~100	~100	~100	>>2,000	GFLOPS/$M

Another disruptive technology that is on the horizon is systems composed of large numbers of small but highly integrated floating-point processors. Examples are the IBM Cell processor and FPGA-based boards made by companies like ClearSpeed. The first implementation of the Cell chip has a theoretical peak of 0.25 TFLOPS; FPGA-based boards are capable of 0.1 TFLOPS today.

Arrays of floating point processors that are designed to support image-processing functions can deliver huge amounts of computational power. While these floating processor boards offer neither the compute density nor the power consumption advantages of the BlueGene architecture, they could possibly be more cost-effective in the 2009-2010 time frame. During the course of the R&D phase of the project, the LSST DM team will closely monitor progress on these fronts as well.

The 2004 ITRS reports the industry consensus that volume production of 4-gigabit memory chips will occur in 2008 and the 8-gigabit parts will be available in volume in 2011. This has implications for the maximum memory that one can install in a system built from commonly available motherboards. Motherboards today have a very limited number of slots for DRAM modules. Slots for DRAMs are highly unlikely to increase over the next few years and could possibly decrease as a cost-savings measure. This will not be a serious issue for the data processing pipeline but may impact the choice of platforms for the database servers. This question will be explored in more detail during the R&D phase as the requirements for the database servers are further defined.

Projecting costs out 5-7 years is very risky. Relying on historical trends is a basis for a cost projection model but it neglects any severe impacts caused by fluctuating business climate, failure to solve any one of 134 technical challenges listed in the ITRS report, or the financial inability of any company to continue IC manufacturing capital investments in excess of $10 billion every few years. As a result, we will revise the cost projections every 6 months during the R&D phase and evaluate whether the basis of estimation has remained consistent or undergone fundamental changes.

Based upon some actual system costs, not merely components or individual nodes, but integrated systems (except for mass storage), the LSST program will use this cost model:

Table 4.4.5-3 Projected cost performance trends in computing

Year of purchase	Computing Price-performance
2004	3 Teraflop/s/Million $
2008	14 Teraflop/s/Million $
2010	22 Teraflop/s/Million $
2012	35 Teraflop/s/Million $

The 2004 cost is based upon the actual BlueGene pricing quoted by the vendor. The 2008 and 2010 numbers are based upon information from vendor internal programs. The 2012 number is a projection based upon 2008/2010 numbers and historical trends.

4.4.5.3.2 Storage Trends

Disk drives with a 1 Terabyte capacity will be readily available in the 2006-2007 timeframe. The capacity of disk drives continues its upward trend with the high volume availability in mid-2005 of disk drives that use perpendicular recording techniques to hit an areal recording density of 133 gigabit/ in^2. It is interesting to note that longitudinal recording techniques which had been the norm since 1960s hit a physical limit at 100 gigabit/in^2.

While 3.5" drives are the standard today, there will be increasing use of 2.5" drives because systems will need more spindles to saturate faster host connection pipes such as FC-10, Infiniband 4x and Infiniband 12x. While the areal density of the disks will continue to increase, disk vendors have a tremendous cost incentive to reduce the platter count in a drive. While full size 3.5" drives are expected to be available for the next few years, this form factor will probably be discontinued long before 2010. By 2010, LSST expects 2.5" drives spinning at 23,000 RPM to be the norm at the high end, with 1" drives becoming cost-competitive.

One advantage to LSST of the smaller format disk drives is that the cost of a failed component is lower and the time to reconstruct the data from the parity drive in a RAID configuration is lower. It lowers the MTTR of the storage subsystem. The RAID feature more than offsets the slight increase in MTBF caused by having so many more disk drives.

Projecting out to 2010, it should be very easy to store a few petabytes of data in a single rack at the base and archive sites. For a given capacity disk drive, the purchase price drops about 5% per quarter or about 20% per year. However as new drive technology come out the cost per MB and also cost per GB bandwidth has more drastic reduction that is closer to 30-40% for equivalent capacity. The following table contains cost and bandwidth projects for a 1 Petabyte integrated storage system from a major disk storage subsystem provider. The storage capacity here is the raw storage; the usable space is reduced from this depending upon LSST's choice of file system and RAID technology. These costs do not include the host system bus adapters nor the interconnect fabric.

Table 4.4.5-4 Projected cost of 1 Petabyte of disk storage

Year	Bandwidth	Cost
2007	10 gigabyte/s	$ 1.45 Million
2008	10 gigabyte/s	$ 1.15 Million
2009	20 gigabyte/s	$ 950,000
2010	20 gigabyte/s	$ 760,000
2011	30 gigabyte/s	$ 620,000
2012	30 gigabyte/s	$ 500,000
2013	30 gigabyte/s	$ 425,000

The LSST program will carefully evaluate the projected IO bandwidth required for the archive site. The desired bandwidth is the sum of the incoming data stream rate, the IO rate needed to sustain data queries from the astronomical community, plus the IO rate required in the event of reprocessing all of the archive images. In the event of a reprocessing of the entire archive late in the survey period, the total number of bytes read and written could easily exceed one hundred petabytes. Increasing the bandwidth to storage does add costs. A significant increase in bandwidth is expected to add 10-30% to the cost of the disk storage.

4.4.5.3.3 Networking Trends

There are thee networks required by the LSST program. The first network connects the mountain top site to the base camp; the second network provides the connectivity between the base camp and the archive site(s); and the third connects the archive center to the end user and data access sites. The mountain-base camp network will most likely be a 10 gigabit/s fiber network. While the telescope site decision has not yet been made, the distance from the mountain to the base is likely to be less than 100 km. Given the remoteness of the mountain site, it will be cheaper over the life of the project for the LSST project to install, own, and maintain this network link. The infrastructure at both ends of the fiber will be standard commercial networking products which will be reliable (high MTBF), easy to repair (low MTTF), and inexpensive to maintain

The distance from the base camp to the main archive site is going to be several hundred to several thousand kilometers. The most cost-effective implementation of this network is for the project to purchase the necessary bandwidth from a telecommunications provider. There are no technological problems standing in the way of building a 10 gigabit/s network between North America and South America. At the Supercomputing 2004 conference, AMPATH demonstrated a 2 gigabit/s connection between Pittsburgh Convention Center and Brazil. The LSST project will treat the long-distance network as a utility to purchase, just like electricity. LSST will examine the data network choices made by Gemini South team to leverage off of their experiences.

During the design and engineering phase of this project, more detailed cost projections from several long distance carriers will be developed. This will be refined once the final site for the telescope is chosen.

The costs for the long haul network are composed of the hardware cost of equipment to drive the 10 gigabit/s fiber, the maintenance and support costs for the hardware, and the cost to lease the data capacity from a long haul provider. The hardware cost is insignificant; the personnel costs are also quite low since a small team at the base camp will provide service to the entire complex of computers and networking. The NOAO experience in Chile has shown consistently over the years that this is a very cost-effective and successful approach. The most costly component is the lease of bandwidth. These costs are driven by demand and not technology. There are 2 data points to examine. NOAO leased a 10 megabit/s line from Chile to the USA via AMPATH for the 3 year period of 2002-2005. This was at a fixed price cost of $250K/yr. The newest 3 year contract is for a 45 megabit/s connection at just under $200K/yr. This shows a 5-fold improvement in price-performance over 3 years. If this trend holds, then the cost for a 10 gigabit/s connection in 2011 would be about $500K in the first year of operation and would decline in each succeeding year.

There are two commercial providers of long haul networking between North and South America – Global Crossing and Telefonica (Emergia). The network topologies for these companies are remarkably similar. The total capacity of these networks between North and South America is in excess of 4 terabit/s. Competition between the suppliers will continue to keep intense pressure on bandwidth pricing.

LSST will explore partnerships with several Research and Education Network (REN) organizations who have current relationships with the NSF. The Cooperacion Latino-Americana de Redes Avanzadas (CLARA) is an organization of South American national RENs who are cooperating on building a high-speed connection into US networks. This is a possible opportunity for LSST to participate in this effort. These various approaches will all combine the LSST purchasing power with other Research Education Network (REN) partners to achieve our required bandwidth at the lowest possible cost.

As an alternative to a high speed network, a model for transport of physical media between the base camp and archive site was explored. At current pricing, the cheapest bulk delivery service between South America and North America for a 100 pound box is $568. Given the weight of a DVD, about 1,000 DVD-style disks could be safely packed into a 100 pound box. In the LSST timeframe, this collection of disks would hold 2 nights of data. The yearly cost to ship the boxes would be $98,000. While the capacity of disks is high enough, the disk data transfer rate is not measured in gigabytes/s, but megabytes/s. In order to keep up with the nightly data rate, 25 to 50 DVD writers would be needed at the base camp and another 25-50 DVD readers would be required at the archive site. It is estimated that at least 1 person at both locations would be needed just to keep the disk writers/readers busy. The combined costs of 2 people, 100 DVD units, and shipping is considerably higher than acquiring the necessary long haul network bandwidth. The following table summarizes these costs over a 9 year period, using the current 4.5x bandwidth improvement every 3 years under constant dollar cost. The physical shipping model assumes a $1.00 cost for a quality DVD-style disk is constant, that shipping costs are constant, and that fully burdened personnel costs for 2 people of $250K/yr increase at a 5% annual rate.

Table 4.4.5-5 Cost Comparison of Media Shipping versus Network

YEAR	Media and Shipping Costs	Personnel Costs	Total Media Costs	10 gigabit/s network
2012	$263K	$250K	$513K	$500K
2013	$263K	$262K	$525K	$500K
2014	$263K	$276K	$539K	$500K
2015	$263K	$289K	$552K	$125K
2016	$263K	$304K	$567K	$125K
2017	$263K	$319K	$582K	$125K
2018	$263K	$335K	$598K	$30K
2019	$263K	$352K	$615K	$30K
2020	$263K	$369K	$632K	$30K
9 Year Total			$5,123K	$1,965K

The table does not include the cost of 50-100 heavy-use DVD reader/writers nor the costs to maintain them. During the course of the telescope lifetime, higher capacity disks with sufficiently high data transfer rates will exist to reduce the shipping costs, but the network connectivity cost will be so low that the network solution will always be the superior choice. Note that shipping physical media is a one-way data flow and imposes a several day latency period, so a lower speed network would still be required between the base camp and archive site for problem tracking and resolution and data base updates to meet the science alert requirements. This would be an additional cost. Better science can, therefore, be done more cheaply by the use of the 10 gigabit/s network link.

The archive center to data access center network bandwidth/cost is a relatively low risk in that there are several research oriented networks within the United States that already provide bandwidths up to 40 gigabits/s and they are constantly increasing the available capacity. Internet2, TeraGrid, National Lambdarail, OSG/Grid3, ESNet and many others are examples of networks in this class. Given the NSF and DOE funding profile of LSST, the existing access to these networks by most of our partners, and the anticipated growth in bandwidth, it is virtually unthinkable that bandwidth/operating costs will represent a significant risk issue to LSST. The

main issue here, already discussed in the Middleware section, is to implement robust and tunable access control via service tiers.

4.4.5.4 Baseline Infrastructure Conceptual Architecture

In this section we define concepts for an architecture driven by the above tradeoffs and trends. It should be noted that the architecture will not be final until the end of the R&D phase, to permit maximum leverage of the positive trends in each technology. Rather, this is a preliminary architecture that, should the trends continue unaltered, represents an optimization of cost, reliability, and availability consistent with the goals of the project.

4.4.5.4.1 Mountaintop Site and Data Acquisition Interface

The infrastructure requirements for the Observatory, Camera, and Telescope control systems are discussed in the System Engineering section. This section discusses the Data Acquisition Interface infrastructure.

The camera will produce data at a rate of up to 24 gigabits/s during readout. The data management system must accept this data flow into a rapid cache interface and then stream the data out during the following exposure (nominally 15 seconds with overheads for sensor readout and telescope slew/settle time).

The nominal plan for data acquisition has 25 fibers from the camera to 25 DAQ compute nodes. This is one per “raft” of 9 CCDs where the camera is populated with 21 full rafts and 4 partial rafts with just 3 CCDs. Additional fibers and DAQ compute nodes would also be needed to receive the wavefront sensor data and optionally the guidance system data.

Each DAQ node will have a sufficient memory cache (each raft generates 288 Mbytes/image) to buffer several images. These data will be received from the camera in a short time (1-2s) and then injected into the DM system over the course of the next camera exposure. In order to keep the operation as automated and fault tolerant as feasible, the data would flow from this rapid cache directly into both the network connection to the base facility and a cache (first in, first out) adequate to store 2 nights worth of data.

4.4.5.4.2 Mountain to Base network segment

As described above, the mountain to base facility network segment will be used to transfer data in near real time, so that the computing facilities at the base facility can be used to do additional data quality assessment and to produce alerts on a timescale of a minute or less.

This bandwidth would nominally be supplied by two sets of fibers providing 4Gbps per pair to load balance the data to the base facility, additionally giving a redundant fiber pair in case of failure on one circuit. There would be another fiber circuit to accommodate command and control communication from the base facility as well as video and maintenance services. A slower backup microwave link will be employed to minimize the catch-up time needed in the event of outage in the primary link.

These links would be owned and operated by LSST and/or the operations entity, providing a high degree of control over the maintenance and mean time to repair (MTTR << 24 hours) in order to meet the minimal downtime specifications of LSST operations.

4.4.5.4.3 Base Facilities

At the base facility, the data would flow into a computing center including a pipeline server and associated high-speed storage needed for image processing, detection, association and alert generation. The minimal operational requirements of the base facility facilities also involve serving as a staging area for the data to be transferred from the base facility to the archive facility.

This computing center will be optimized for floating point calculations, and therefore is anticipated to be based on a Bluegene/L or Cell computer architecture, with fiber-interconnected disk storage. The details of this architecture will be elaborated during the R&D phase.

4.4.5.4.4 Base to Archive International Network Connection

The requirements for this link are dependent on the operational model, but the absolute minimal requirement is that all raw data and associated metadata be transferred on the timescale of less than 24 hours.

It is anticipated that this link will be a leased fiber optic link from a public or commercial carrier. Refer to section 3.4.5.1 for a discussion of options.

4.4.5.4.5 Archive Center Facilities

The archive center represents the most sophisticated and extensive computing center, given it's multiple roles of pipeline processing and re-processing, permanent data archiving, and providing data access for distribution to data centers and end users.

The pipeline processing architecture will mirror that of the base pipeline server, but will be expanded in capacity to support re-processing. The data access server architecture is anticipated to be a combination of distributed file system commodity servers and a multi-tiered database architecture on clusters. Both are likely to be interfaced to physical storage devices via a Storage Area Network (SAN). Again, the details of this architecture will be elaborated during the R&D phase.

4.4.5.4.6 Data Center Facilities

This facility will support data access servers that mirror the architecture at the archive center, typically without the pipeline servers.

5 R&D Investigations

5.1 Simulations

The data products of the simulations group need validation. The key validation is ensuring the modeling of the atmosphere is sensible. We have taken data on an existing large-aperture instrument, with 15 seconds of integration time. The observing included seeing conditions from good through bad. These data were taken in tandem with MASS and DIMM atmospheric measurements in order to provide input to our atmospheric models. Analysis of these datasets is in progress, and they will used to update the fidelity of our simulator accordingly.

Additional improvements include the proper modeling of the telescope control system based on inputs from the wavefront sensors, and more realistic treatment of the telescope and camera responses to environmental disturbances. These will be incorporated as the design matures.

The simulator will be used to establish the photometric and astrometric performance of the system. To facilitate such studies, we will put some effort into repackaging it into a more user friendly format with a graphical user interface.

The atmosphere seeing varies over time, and different parts of the sky have different densities of stars (used to correct PSF distortions). Due to the stochastic nature of seeing and appearance of transient and high proper motion objects, the LSST run strategy evolves in real time.

Therefore, one doesn't know a priori the precise nature of a data set in terms of its quality of sky coverage. The LSST group therefore models the real time run strategy with a Cadence Simulator, as described in Section 4.1.3. This simulator takes as input realistic atmospheric conditions, and makes decisions about pointing direction and filter. Its output is a stack for each accessible direction in the sky with pointers to atmospheric conditions, filter, etc. Ultimately, the optical simulations group will take these stacks, fill in images (real or synthetic), then apply cosmology, atmosphere and instrument distortions. At present, the simulations group is processing a single operations simulator stack (one particular sky direction) over a 10'x10' (CCD-sized) piece of the image. Extending this to a full year over the full sky will require considerable parallelization and automation; this in the charge of the proposed programmer/librarian.

5.2 Telescope R&D; Site Investigations

The telescope and site research and development efforts will focus on the critical technical issues, the reduction of technical and schedule risks, and the development of designs with fully supported costs and schedules. Efforts to address these objectives will be planned to support incremental design development for full project coordination, to support development of construction proposals and to support and coordinate the ongoing development of long lead items funded privately. A full plan of efforts aligned with the work breakdown structure has been developed. This plan has already been initiated using both in-kind resources and funding from the NSF Design and Development Cooperative Agreement. The most significant of the efforts is further described in the following sections. They are the efforts to characterize and select a site for the telescope, Design and develop the telescope mount and drive systems, the efforts to address the optics, and the wavefront sensor and alignment system.

5.2.1 Site Evaluation and Selection

The LSST project has completed two phases of site investigation, narrowing the potential sites to three observatories in two regions. One potential site is located at San Pedro Martir, in Mexico. The other sites are located in Chile. A plan for further evaluation leading to a final site selection has been defined to gather the information necessary for the final site selection. This includes additional studies of existing data, gathering more historical data, collaborating on new data generation, and tasks for new equipment for LSST and site specific testing. The objective is to assemble the most complete and uniform data set possible for these three sites considering limitations on funding and time. The plan includes both internal and external tasking as well as tasks being performed by others that could benefit LSST. The plan is to make the final site selection in April 2006.

5.2.1.1 Cloud Cover

One of the important criteria that potentially differentiates the three sites and has direct impact on the LSST survey is Cloud Cover. To further investigate the statistics at each site a campaign will be conducted to simultaneously analyze date developed from both satellite and ground detection methods. LSST will continue to work with Dr. Andre Erasmus to evaluate GOES satellite data with the highest fidelity tools developed for this downward looking analysis. Figure 5.2.1-1 shows one set of GOES images that are available every 3 hours.

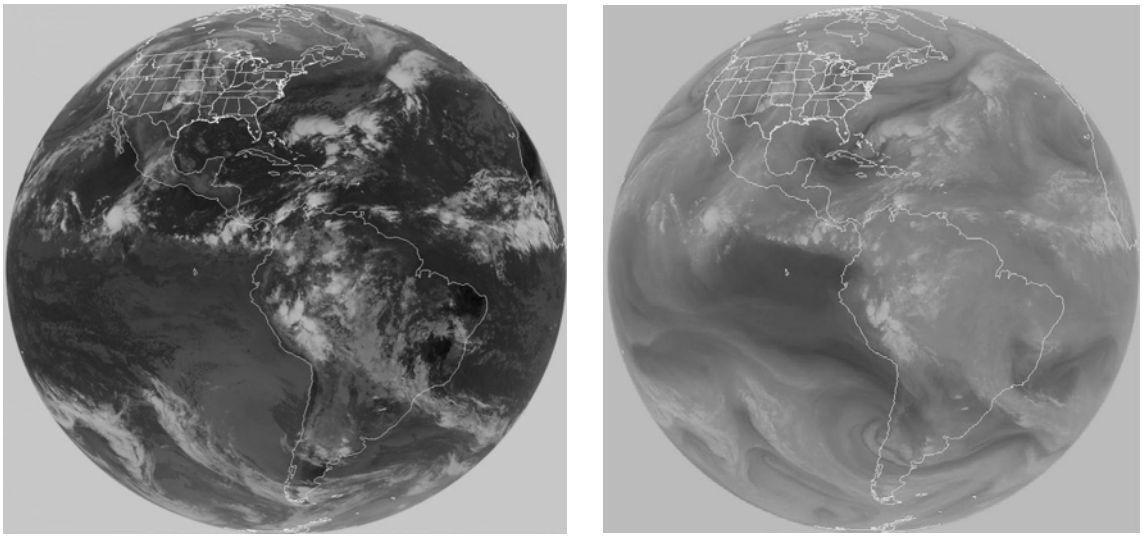


Figure 5.2.1-1 Goes-8 Satelite image at 11:45UT on October 25, 2000. Left, infra-red window channel (10.7micron) and right, water vapor channel (6.7 micron). LSST study evaluates 25 pixels above each site to establish level of cloud cover.

For a contemporaneous year we will also deploy an all-sky camera at each of the three potential sites to measure the cloud cover from the ground looking up..The All-Sky cameras have been deployed in the past near observatories for cloud detection. The camera's that LSST is placing (1) on each site has been proven out at Cerro Tololo Inter-American Observatory in Chile. Figure 5.2.1-2 shows the output of one such early model deployed on Cerro Tololo and Cerro Pachon. Upgrades for the LSST version include observations in additional filters to evaluate the sky extinction and to estimate the possible impact of light pollution on sky brightness from neighboring towns. It will also include auto-photometry evaluation software to quantify the photometry to 2% as a function of time and area. This quantification of the data will directly feed into the LSST observing simulator to provide information on the number of observed fields and the observing statistics, field by field for the entire simulated survey.

The correlation of the ground and satellite data over the year will increase the confidence in the use of this data for site evaluation purposes. The satellite analysis is already well established but this direct correlation has never been done systematically. It is important to understand the significant of the satellite data as a longer historical data set exists from these satellites. These results are expected to be strong evaluation criteria in the site selection process.

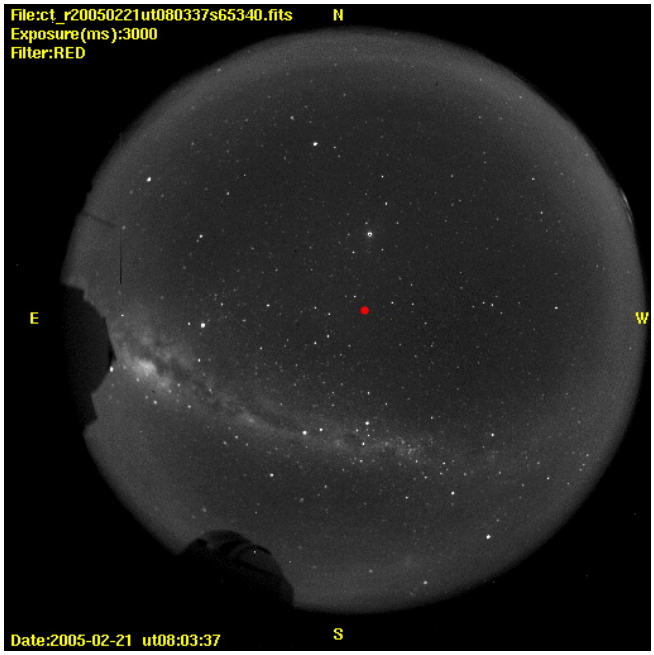


Figure 5.2.1-2 Image from CTIO All-Sky camera system called TASCA (Chile). The buildings in the background are the observatories located near the camera. The red point represents the pointing of the main telescope.

5.2.1.2 Astro-Climate Data

Atmospheric seeing data, a measure of the atmospheric impact on light propagating through the atmosphere, is a traditional method for astronomical site evaluation. This is characteristic which is very site dependant and is customarily monitored continuously at sites as a reference to the quality of the atmosphere hpe and hence the level the observatory should be able to achieve. For the three LSST candidate sites, the historical values for seeing are very similar. The final year of LSST site evaluation will still include new measures of seeing to add to, and validate early measurement statistics. This data set is also used by the LSST observing simulator to plan and assess the survey. Differential image motion monitors (DIMMs) will be used at each site to evaluate seeing. Figure 5.2.1-3 shows such an instrument. These systems have been in use for many years at major observatories and recent efforts in Chile have calibrated the various DIMMS in use at sites so that comparisons have credibility. Las Campanas (Chile) and San Pedro Martir (Mexico) have their own systems already in place. Cerro Pachon also has a DIMM but it is 2 kms away. A dedicated LSST DIMM will be installed at there, directly on Cerro Penon to evaluate the specific peak and to offer insight into the difference with the neighbor DIMM giving LSST a better time baseline of measurements to compare in the selection process.

Figure 5.2.1-3 Example of a Differential Image Motion Monitor (DIMM)

Beyond the use of weather station data already installed at the sites, we are planning to measure physical parameters to model local turbulence and ground layer height in an effort to compute the final height of the telescope. For Pachon and Campanas it is suspected that the layer height is much lower then the likely height defined for handling issues. At San Pedro Martir this is less certain due to the unknown affect of the trees. The objective of this task is to get a tower with either micro thermal gauges, ultra sonic, or mechanical 3 axis wind speed monitors at various altitudes to find the height of "laminar" flow. A second objective for this task is to develop a data set of the wind power spectrum for the site. This will be folded into the design and analysis of the dome and telescope structure in efforts to mitigate the impact of wind on the structural and optical performance.

5.2.2 Mount Development

Since the LSST is a survey telescope, a significant portion of its duty cycle occurs during slewing and settling. It has been estimated that the best slew and settling time that can be obtained using a purely conventional telescope structure is ~8 seconds. Meeting the present 5 second goal will require significant improvements.

Although the minimal dynamic requirements of the telescope can be marginally attained with a conventional telescope structure, a significant reduction in the telescopes slew and settling time should be attainable by incorporating advanced technologies. These technologies include advanced motor controls, advanced structural materials and smart structures (active damping).

The application of more conventional technologies which are not normally applied to telescopes and will also be investigated. Most telescopes are not very sensitive to settling time, consequently, they do not incorporate significant damping and are not maximized for stiffness. The incorporation of passive damping into the telescope assembly will be investigated. The stiffness of the telescope assembly, which limits the natural frequency, is significantly affected by

the hydrostatic bearing stiffness. Maximizing the bearing stiffness is another specific area that will be investigated.

5.2.2.1 Advanced Motor Control

Very conventional methods were used to estimate the 8 second slew and settling time. The maximum power was used to accelerate and decelerate the telescope. The settling time was then determined based on the maximum acceleration. Larger accelerations reduce the slew time but increase the settling time. The power was varied to minimize the combined slew and settling time.

Advanced motor control systems do not operate under the above simplistic motor control scheme. The accelerations and decelerations are controlled to minimize the excitation of the principle vibrations modes, which reduces the combined slew and settling time. These control systems are already in use for large radar structures. These radar structures are similar in configuration to large telescopes. Their tracking rates, however, are much higher. Incorporation of these advanced motor control systems will require significant input from, and interaction with, the motor control industry.

5.2.2.2 Advanced Structural Materials

There is little difference in the specific modulus of the common structural materials, steel, aluminum etc. The specific modulus is the elastic modulus divided by the density. Since the natural frequency of a structure is proportional to the square root of the specific modulus of its material, common material substitutions produce negligible effects on a structures natural frequency. Geometrically identical steel and aluminum structures will have nearly identical dynamic characteristics.

There are advanced materials with significantly higher specific modulus. Among these are graphite epoxy composites, beryllium and silicon carbide. Of these only the graphite epoxy composites have been used for large structures (aircraft).

Although it would not be practical to build the majority of the telescope mount from advanced materials, the design does incorporate large cylindrical tubings which would be ideal candidates for graphite epoxy composite construction.

Non graphite composite tubing is commonly produced in the appropriate size for the piping industry. Unfortunately, this tubing does not have the appropriate characteristics. Smaller tubings with the appropriate characteristics are commonly produced by the aerospace industry. This application would require that aerospace quality tubing be produced at piping industry size.

Determination of the feasibility of incorporating advanced materials into the LSST telescope mount will require significant industrial input. The performance of these materials is significantly more fabrication sensitive than conventional structural materials.

5.2.2.3 Smart Structures

Smart structures use actuators and sensors at milli- and micro-scales to achieve a certain goal. These are very short stroke, high speed actuators, piezoelectric ceramics for example. In the case of LSST the principle goal will be vibration attenuation (active damping). The vibration induced accelerations are sensed and the actuators are used to counteract these accelerations. Through this application, appreciable reductions in settling times are possible.

Smart structures can also be used for thermal and gravitational compensation. However, thermal and gravitational compensation is better provided by conventional positioning actuators (hexapods). These types of compensation are already common for large telescopes. The larger displacements and slower actuation requirements are better suited for conventional actuators.

5.2.2.4 Structure R&D Plan

The research and development plan to address the issues of the telescope structure will include concept design studies pursued by industrial vendors and targeted studies with vendors of specific advanced technologies. During the first year of the plan a contract will be let to a large precision structures vendor to elaborate on a concept design with the objective of exposing the critical issues, like those identified above, and to elaborate on the performance benefits in incorporating these additional technologies. Following this study, LSST will let specific study contracts with specialty manufacturers of those products and technologies that are determined to provide the highest payoff in the performance and/or cost of the LSST telescope.

5.2.3 Telescope Optics

The Research and Development plan for the LSST reflective optics addresses the glass and support designs of each mirror. The design trade and concept analysis of the mirrors is a very short period because orders for these mirrors are being made early using private funding. The primary / tertiary monolithic mirror glass is in fact already ordered and LSST plans to order the material to build the secondary mirror in early 2007. All aspects of the telescope optical systems are schedule and cost risk items due to the long lead and very custom nature of their development.

5.2.3.1 Primary/Tertiary Monolithic Mirror

The R&D effort for the Primary/Tertiary mirror is focused on the optical fabrication and testing of this unique mirror as well as the support system for its adaptive control. Achieving the tight alignment tolerances of the two surfaces with respect to each other is a challenging task that requires incorporating additional state of the art measuring systems into the metrology system. LSST and the University of Arizona are working closely to develop a suitable approach.

The active support and thermal control system will be developed from existing systems already deployed. LSST engineers will take the lead in evaluating the previous approaches to supporting and controlling these large borosilicate mirrors at the Large Binocular Telescope, the Magellan telescope, the MMT and others. The support actuators have evolved with each mirror deployment and several different approaches to thermal control have been taken at the different telescopes. Working closely with the University of Arizona engineers and the operators of these other working examples, LSST will generate the final requirements and concept definition and will establish a design contract for the Primary Mirror support system with an outside vendor.

The R&D effort will also need to support the development of the mirror at the University of Arizona. The contract to provide the glass substrate and the subsequent figuring and polishing is defined in 4 phases. The first phase is the casting engineering and bulk material purchase. The second phase is the actual casting process and the third phase is the optical finishing. The last phase integrates the glass with the LSST provided support system and tests the system under the metrology tower. The first two phases and the start of the third phase will happen during the R&D effort. LSST engineers must complete sufficient systems engineering to continue the primary mirror development passing on all final critical design details to the U of A.

5.2.3.2 Secondary Mirror

The baseline design has been established for the secondary mirror system by alternate materials will also be investigated and evaluated for superior performance. Preliminary design for secondary mirrors of various materials and configurations will be developed, analyses, priced and evaluated. The secondary mirror might take advantage of lightweight high stiffness materials like silicon carbide or beryllium or it might be best to stay with conventional glass construction. The mirror may perform best as an adaptive mirror, a thin sheet of glass supported by many actuators,

or it may be best as a stiff mirror with few support actuators. All these design choice will be evaluated against predicted performance, cost, and fabrication risk. Contracts will be established with several vendors to develop suitable designs for various material choices to insure an optimal design is evaluated for each material and fabrication technique considered. Following the final design selection the long lead material will be purchased in early 2007.

5.2.4 Wavefront Sensing and Alignment

The wavefront sensing and alignment systems of LSST are the central nervous system for the telescope and the active optical system. All three mirrors will have rigid body position degrees of freedom and all three will have some degree of figure correction. In addition to he importance to the functioning telescope the wavefront sensing for LSST must happen at the focal plane so it a critical part of the already critical focal plane development in the camera. For both of these reasons the research and development of the Wavefront sensor system and Alignment system are critical aspects of the overall development plan. The alignment system, also referred to as the auxiliary alignment system, is a collection of metering devices to keep the rigid body alignment of the optical system within a certain level of capture and in some cases hold those alignments to specification. This auxiliary alignment system is to be developed to reduce the degrees of freedom that must be addressed by the wavefront measuring system. To this end both systems are intimately related to one another and further, both have several possible supporting technologies and additional restraints. For example, the wavefront sensor system must be compatible with the overall focal plane layout.

5.2.4.1 System Engineering Tools

Several tools will be developed to address the many design variations and assess the requirements and performance of the system architectures and characteristics. The first of these tools is a model of the optical system that focuses on the transfer function of optical errors measured in the focal plane to the source error in the system. This is the reconstruction algorithm that utilizes field position measurements and provides the necessary correction matrix for all degrees of freedom. The basic reconstructor algorithm will be assembled into a system model that accepts noise and measurement parameters associated with different sensing technologies and also allow modifications to the degrees of freedom to be controlled within the system. The later allows the incorporation and impact of an auxiliary alignment system operating in the telescope. For example, the primary and tertiary mirror will be outfitted with edge sensors that will keep the two surfaces in rigid body alignment well within tolerance for extensive periods (of order 30 days). This reduces the demand on the wavefront sensing system and needs to be included in the system simulator (model). In addition, the tool will have the fidelity to apply additional filters that account for natural modes of the final system design and time scales that relate to operating conditions. The first bending modes of each mirror surface need the different bandwidth sensing then other system modes. Additionally, the reconstruction math may simplify and be more accurate if system behavior modes are considered. This reconstructor simulation tool is the heart of the engineering effort and the algorithms will be the heart of the operational system. An early precursor reconstructor has been developed at NOAO to validate the concept for LSST. A new effort has been imitated to develop the algorithm more thoroughly and parallel efforts at other organizations will be conducted to cross correlate results of this critical effort.

5.2.4.2 Wavefront Sensor Technologies

Several technologies and approaches are under consideration for the wavefront sensing method for LSST. (see Section 4.2.5) In the research and development phase the focus will be on

development of viable methods sufficient to compare the performance and compatibility of each for the LSST System. Specifically three steps will be followed.

1. The curvature sensing method in use today relies on analytic code developed several years ago. This code will be updated and optimized for use in the LSST application and tested to determine additional performance capability.
2. Two competing concept studies will be let with industrial sources of custom wavefront sensing solutions. These studies will address the critical performance and interface issues and provide sufficient programmatic information to assess the viability of the selected approach for LSST.
3. An additional in house effort will be started to further investigate methods of low order aberration measurements directly form out of focus images.

All of these efforts are planned to be completed by December 2006. At this time the complete baseline system will be defined and detailed development of the relevant subsystems will be initiated. The final definition will be done in concert with the camera focal plane developers and the critical design phase will include final focal plane constraints worked out and documented in a wavefront sensor - focal plane interface control document.

5.2.4.3 Auxiliary Alignment Systems

The development of the Auxiliary Alignment System will be carried out in a stepwise fashion from requirement definition through prototyping and final definition. Aggressive requirements will be defined to establish the goals for a system, available technologies and tools will be researched that address the system goals, and a conceptual design will be defined. The intent is to stay with proven and commercially applied technologies. A prototype system, or parts of the system, that represent the critical and high risk aspects will be procured and tested to the fullest extent possible. The key to this development work is to establish and verify the accuracies of the alignment system to insure a proper allocation is made within the alignment budgets. This will impact the development of the wavefront sensor system and will potentially impact the design of the optics, camera and structure so that each critical element in the aligned system is compatible with the measurement techniques.

5.3 R&D Investigations for the Camera

5.3.1 Focal Plane Development Plan

5.3.1.1 Overview and General Discussion

This section details the R&D activities associated with the focal plane array (FPA) of the LSST Camera. These activities include:

Sensor development

Sensor test plan and test stand development

Raft design, development and testing

The current state of conceptual design of the elements of the focal plane array and these R&D activities are described above in Section 4.3.2.

The development plan for the sensors themselves involves an iterative process with potential vendors of both CCD and Si PIN-CMOS arrays. This iterative process is intended to start in early FY 2005. Based on evolution of the current LSST strawman design for the CCD sensor, R&D contracts could be awarded to several potential vendors in late FY 2005 and would lead to tested final prototypes about 2 years later. For CMOS arrays, testing would continue on devices as

available at no cost from interested vendors; a strawman design for an LSST PIN-CMOS sensor would emerge at a later date.

In the process of FPA development, LSST will be testing early prototypes with equipment presently available at the collaborating institutions. This equipment is suitable for both CCD and PIN-CMOS sensors. This testing will guide the FPA development and will also define the test and certification procedures and equipment (the test stand) that LSST will employ in the production fabrication phase of the FPA. The handoff from the vendor or vendors to LSST will also be established – that is, what testing and certification will be performed by vendors prior to delivery and what will be performed by LSST after delivery of production sensors.

As indicated in Section 4.3.2, the sensors will be mounted in groups of 9 on intermediate mechanical structures called "rafts." The rafts are in turn mounted on a mechanical backplane called the "FPA Integrating Structure." The rafts of sensors must meet certain metrological and thermal specifications, especially the flatness of the section of the FPA represented by one raft. Flatness of the multi-sensor rafts, both on the bench and in the camera at operating temperature must be measurable and possibly adjustable. Techniques and equipment for this will be developed during the FPA R&D phase as well.

5.3.1.2 Sensor Development Plan

The LSST FPA will comprise an order of magnitude more pixels than any array realized so far. The principle underlying the development plan is that for an FPA involving 250 to 300 large format sensors, one cannot rely on the handcrafting of individual sensors. Rather, an industrial approach has to be developed and adopted. Therefore the plan that LSST proposes to follow in sensor development is to employ targeted R&D by qualified vendors guided by the LSST camera team, with the goal of producing a prototype suitable for industrial production.

Toward this end the LSST sensor group has completed a Strawman CCD design that will keep the development by the vendors focused on LSST requirements.

The CCD sensor strawman design (Figure 4.3.2-12) is a 4000 × 4000 array of 10 μm pixels, with specifications enumerated in Table 4.3.2-2. A Vendor Information Package for prospective vendors was completed and distributed December 2004. On this schedule, a formal RFP for R&D could then go forward to those vendors who express interest and capability. This would occur happen in March 2005. A possible schedule for the subsequent steps in this plan is provided in Table 5.3.1-1:

Table 5.3.1-1 CCD sensor development schedule

Date	Milestone
March 2005	RFP to potential sensor vendors
June 2005	RFP response and vendor(s) selection
September 2005	Vendor contract(s) award
September 2006	1st testable sensor from fabrication efforts
November 2006	Evaluation of testable sensor
March 2007	Iterate design, production of 2nd version sensors
May 2007	Evaluation of 2nd version sensors
September 2007	Delivery of tested final prototype sensors

Note that it is essential to have two or three such R&D efforts running concurrently with different vendors if a second time-consuming round of development is to be avoided. The

resources requested in this proposal for FY 2006 are dominated by the funding of 3 such contracts to have been awarded in September 2005.

5.3.1.2.1 Sensor Development Resources

R&D funding for sensor development is needed to address:

- Contract costs of the development work by the vendors
- Salary and MS&T for sensor development staff not covered by institutional resources

Over the 2+-year period covered by this proposal, the vendor development contracts are estimated to have a value of $1000k per vendor. This includes $300k in FY 2005 to cover vendor-related activities prior to contract award.

The scientific and professional manpower involved in LSST sensor development activities to date has been supported by internal funds at the collaborating institutions. In the period covered by this proposal, additional professional and technical staff (plus associated MS&T) will be required. As tabulated in Section 7 of this proposal, we require R&D funds to support 2.5 FTE in FY 2006 and 2007 on sensor development activities.

5.3.1.3 Sensor Testing and Test Stand Development

The testing of prototype sensors during the R&D phase will be accomplished with existing equipment suitable for measuring the performance of individual sensors. Tests include (see Sections 4.3.2.2 and 4.3.4).

- Flatness
- Quantum efficiency
- Point spread function
- Dead/noisy channels
- Effects of thermal cycling between operating temperature and room temperature

Because of the very fast optics of the telescope and the consequent angular width of the light cone at the focal plane, flatness of the FPA is of paramount importance. The flatness requirement on the whole FPA assembly is 10μm peak-to-valley (p-v). The sensors themselves must be built flat to better than 5μm p-v, with the rest of the flatness budget devoted to the raft and integrating structures. As shown in Figure 5.3.1-1 below, the current state-of-the-art, as represented by the HIRES imager mounted on the Keck II Telescope, is not far from meeting this requirement.

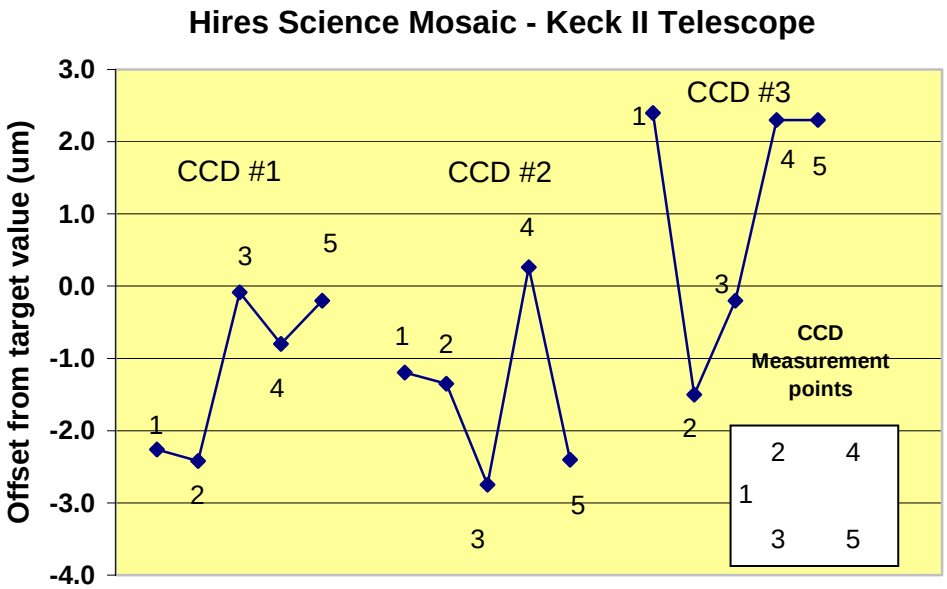


Figure 5.3.1-1 Measurement of as-built CCD flatness in focal plane of Hires imager mounted on the Keck II telescope.

In the course of determining vendor performance in meeting LSST sensor specifications, we will gain the testing experience necessary to design the production sensor test stand and procedures.

Since the plan for production quantities of FPA sensors will be based on industrial-style sensor fabrication, we expect that a significant portion of sensor testing will also be part of the production fabrication contract. Hence an important part of the testing program during the R&D phase will be to determine the split between production testing and certification by the vendor and acceptance testing and characterization by LSST.

The schedule for development of a test stand must be consistent with the sensor development schedule shown in Table 5.3.1-1. On this schedule, a prototype test stand should be available for use with 2nd or final prototype sensors in the summer of 2007. Design of this prototype would start when 1st sensor prototypes are available in late FY 2006.

5.3.1.3.1 Testing and Test Stand Development Resources

BNL, Harvard, CFA and STScI will carry out testing on sensor prototypes. In addition, the R&D-supported BNL staff discussed in Section 5.3.1 will set up a lab at BNL to do CCD prototype testing and the development of the test stand. Additional MS&T funds in the amount of $300k in FY 2005 are required to set up the test stand development laboratory.

5.3.1.4 Raft development and prototype raft testing

The sensor raft is an important element of the FPA design concept (see Sections 4.3.2, 4.3.3 and 4.3.4). It has to accommodate multiple sensors and their connections to the readout electronics and provide for mechanical adjustments and thermal connections to the FPA integrating structure. A key activity of the R&D phase is to design, build and test prototype rafts. The sensor, the readout electronics and the thermal/mechanical design of the FPA all interact with the raft. Figure 4.3.2-12 shows these interfaces schematically.

The interfaces at the raft level are illustrated in the following discussion of the testing associated with the FPA flatness specification:

- The sensor vendor will be responsible for controlling the height of the sensitive surface with respect to three mounting pads. This will require proprietary R&D by the sensor vendor.
- Early prototyping by the LSST camera team will validate the sensor-to-raft interface design.
- The LSST camera team will validate the design of the rafts using a 3x3 array of mock sensors. Passive flat rafts will be tested at operating temperature and all use angles.
- A raft or rafts with active sensors will be tested on an existing telescope.
- Flatness and stability of the integrating structure will be tested in stages: Unless data exists, candidate materials will be thermally cycled and measured for instability. The kinematically adjustable raft-to-backplane interface will be validated for stability.

5.3.1.4.1 Raft development and testing resources

On the schedule shown in Table 5.3.1-1, passive sensors should be available in FY 2006 and active, testable sensors should be available at the end of FY 2006. Thus raft design and early prototyping should be under way in FY 2005 and early FY 2006.

Because of its key role in the integration of the FPA, the raft design and development involves the activities of several institutions in the Camera Team – SLAC, LLNL, BNL and Harvard.

5.3.2 ASIC Development Program

The LSST focal plane will require low-noise analog processing circuits at each sensor output port. At this time it is not known whether the sensors will be n-channel CCDs, p-channel CCDs, or p-i-n/CMOS hybrids. If CCDs are chosen, there will be of order 7,000 such ports and the anticipated pixel rates are 250 - 500 kHz per port to achieve a fast readout. In addition, there will be local clock drivers to supply the parallel and serial clock voltages to each CCD. Both of these circuit functions will have to be miniaturized and dissipate low power, so that they can be located within a few cm of the rafts carrying the sensors. To this end, we will investigate ASIC approaches to the analog front end functions.

If p-i-n/CMOS hybrids are chosen, the front-end functions will be minimized. Hybrid sensors have built-in multiplexing and preamplification, do not require high voltage clock and bias signals, and may incorporate a vendor-supplied readout ASIC. If hybrid sensors are chosen it will probably make custom ASIC development unnecessary. This development plan assumes that CCDs will be the used.

5.3.2.1 Signal Processing ASIC

For processing the CCD output signals, we will develop a multichannel signal processor which will perform dual correlated sampling and integration. This method of signal processing has been found offer a good combination of low noise and flexibility for pixel rates below 1 MHz. The goals of this readout ASIC are:

- negligible contribution to the noise from the CCD output source follower (< 4 electrons rms equivalent input noise)
- ability to handle signals up to 100,000 e- (implies 10^5 dynamic range)
- power dissipation below 25 mW/channel
- channel-channel crosstalk below 0.1%
- linearity 1%
- ability to operate at the focal plane temperature (expected range -100+/-20C)

A tentative set of specifications for this ASIC has been given in Section 3.3.3.2.4.

Development of the signal processing ASIC will be done by the BNL Instrumentation Division, which has substantial experience with low-noise, highly-integrated front end electronics for radiation detectors. Design and prototyping will be done in a commercial submicron CMOS technology and will follow a standard sequence of steps that include:

- Definition of system requirements
- Signal chain simulation using behavioral models
- ASIC architecture definition
- Technology selection
- Transistor-level design and circuit simulation
- Circuit layout and verification
- Prototype fabrication using multiproject service (e.g. MOSIS, www.mosis.org)
- Construction of test fixture
- Device test and evaluation
- Second and further prototype cycles as necessary
- Fabrication and test of production devices

Experience with other ASICs indicates that the time required for first prototype of a circuit of this complexity is between six and 9 months including fabrication. In the case of the LSST, it will be necessary to allow additional time to develop simulation models that accurately describe the device behavior at the chip operating temperature (about 170K), which is below the temperature range typically modeled by the vendors. Successive iterations can take from 4 to 6 months including fabrication.

A set of milestones for the signal processing ASIC development is shown in the table below (FY dates beginning October).

Table 5.3.2-1 Set of milestones for the signal processing ASIC development

System requirements, signal chain simulation, architecture definition	3Q 2005
Technology selection	3Q 2005
Transistor level design and simulation	4Q 2005
Layout and verification	4Q 2005
First prototype fabrication and test	2Q 2006
Second iteration fabrication complete	3Q 2006
Third iteration fab and test (if needed)	1Q 2007
Fabrication and test of production devices	3Q 2007

Since the lead time for ASIC development is long, it will likely have to start before the final sensor has been defined. We will therefore have to design to a set of flexible specifications that may need to be modified during the course of sensor evaluation and selection.

5.3.2.2 Clock Driver and Bias ASIC

This chip will follow a similar development plan as the signal processing ASIC. However, it will be necessary to choose a different process technology since the CCD clock and bias voltages typically require a signal swing of over 20V, which rules out the use of standard submicron CMOS. We will survey the possible vendors of high voltage CMOS processes and characterize test devices for low temperature behavior. The remainder of the development will follow the same cycle as the signal processing ASIC. At this time the collaboration does not have the resources to support the simultaneous development of two ASICs, therefore the development of

the second chip will be delayed by about one year while the signal processing ASIC enters its first prototype fabrication cycle. At this point we will be able to interleave the design and fabrication intervals of the two chips.

5.3.2.3 Supporting activities

In parallel with the chip design and prototype fabrication, we will develop test fixtures necessary to evaluate the performance of the ASICs. As described in the Electronics Development plan, an electronics test bed will be developed that will include discrete circuit implementations of the front-end functions, cryogenic cooling, and the digital back end electronics (digitization, buffer memory, serialization and output to a PC). We intend to make the test bed electronics modular to allow a continuous evolution towards the final design. When first ASICs become available, we will be able to substitute them for the early discrete implementations, which will aid in testing.

The final ASICs will be installed on flexible printed circuit boards that will mount just behind the sensor focal plane. These circuits will have to comply with the stringent mechanical and thermal constraints of the FPA. Design of the flex circuits will proceed in parallel with the development of the ASICs.

A CCD simulator will also be needed to test ASICs. The CCD simulator will have to provide signals and impedance levels similar to the real sensor to allow realistic testing of the ASICs. The CCD simulator will be developed on a parallel path with the signal processing and clock driver ASICs.

5.3.3 Assembly and Alignment of the Focal Plane

5.3.3.1 Mechanical/Thermal

The two leading drivers of the mechanical design are the requirement of 10 micron peak-valley flatness across the entire focal plane array when the camera is operated at -100^{0}C in different orientations and simultaneously, the need for temperature uniformity (and stability) of ~0.1^{0}C across the sensors.

These two requirements compete with each other, and are likely to lead to the need for state-of-the-art mechanical solutions. They imply the need for careful mechanical and thermal engineering of the sensors, their packaging, their fixation onto the rafts, the mounting of the rafts to the integrating structure, and control of external loads (cables, cooling etc.).

While much of the structure can be simulated, we anticipate that there will be a need to verify these calculations by laboratory measurement on mechanical and thermal prototypes. R&D towards development of the final design, will be associated with verifying these calculations.

5.3.3.2 In-Situ Metrology

During the assembly and testing of the focal plane array (FPA), we will need to measure the flatness of the subcomponents (sensors, rafts and integrating structure) as well as the fully loaded integrating structure under a number of conditions simulating operation on the telescope. In particular, we anticipate that the fully loaded structure will need to be measured under five distinct configurations as components arrive at SLAC:

1. The FPA is warm and L3 is removed. Measurements are done in a clean room environment.
2. The FPA is measured warm under vacuum and then cooled down and measured cold under vacuum with a thin glass plate substituted for L3, (there being a vacuum on both sides of the plate).

3. The FPA under vacuum is measured warm and then cooled down and measured with a thick planar glass plate substituted for L3.
4. The FPA is warm and the measurement is taken through L3 both with and without a vacuum.
5. The FPA is then cooled down, under vacuum and the measurement is taken again through L3. This is the real operating condition.

These measurements will be made with the camera in different orientations and will all be performed in a laboratory clean room environment.

We anticipate that non-contact commercial laser-triangulation heads and/or confocal measuring heads can be used to do the initial measurements of the FPA flatness, at the sub micron level shooting through thick lenses (L3) or through windows. Vendors such as Keyence, Micro-Epsilon, and Acuity produce potentially suitable devices. By mounting these heads on an xy stage and providing an optically flat reference, we anticipate that a metrology system can be built for use in the lab during assembly, and then carried to the telescope site for subsequent use if there is a need to disassemble and/or repair elements of the camera.

The first phase of the R&D is effort is to confirm that these devices are suitable for use with anti-reflection coated silicon, with/without windows and lenses and under the conditions (i.)-(v.). We will then use the prototype system in the study of differential in-situ measurements described below.

5.3.3.3 Differential In-Situ Metrology

Once L2 is installed these optical systems will no longer be useable. Nor would they be useable with the camera mounted on the telescope. While there is the potential to use “data” from the stars to detect changes in the flatness of the FPA, those measurements can only take place long after the camera is transported and assembled, and by virtue of requiring the telescope optics, couple together many sources of error.

We are therefore investigating several possible systems that would allow us to detect changes in the flatness of the focal plane at any time after the FPA is assembled, and in particular with L2 in place, and otherwise in operation.

Such in-situ measurements would have the advantage of detecting shifts in the alignment of the rafts very rapidly, and independently of the other optics in the telescope. They can also potentially be used to verify / backup the non-contact optical measurements made during assembly and cooldown of the camera and other mechanical testing of the camera.

At the present time we are planning to investigate four different techniques which may be used singly or possibly in a hybrid mode – combining two of the methods. The techniques break down into two major categories depending on whether they assume the integrating structure is designed sufficiently rigidly to be characterizable and usable as an absolute and repeatable reference in all orientations and temperatures. Briefly the four techniques we have identified for investigation are described below:

1. **The “optical straightedge”** A grid of laser lines to measure offsets in the raft alignments. The beams are created by a set of diode lasers in x and y mounted on the integrating structure. The beams intercept pellicle splitters on each raft, and the beam centroid position detected in a miniature ccd camera.
2. **Frequency scanning interferometry.** This technique has been used in high energy physics experiments to measure the distortions of precision vertex detectors. Frequency scanned laser light is brought into the camera on a set of optical fibers. A fiber is brought under each corners of each raft, and split. Part of the beam goes to a return fiber adjacently mounted. The rest of the beam travels to a retro-reflector mounted to the raft

corner itself. The return beam is captured by the return fiber and produces interference. By scanning the laser frequency and counting fringes (either an APD or photomultiplier as a detector), the optical path is measured to better than 0.1 microns.

3. **Capacitive edge sensors.** This more conventional approach might be used in conjunction with (1) or (2) to stitch together a map of raft positions and changes over time. In particular, the number of optical elements in (1) and (2) could be reduced.
4. **Direct pattern generation**: Use of one or more lasers and diffraction gratings (or pattern generating grating) mounted on the edges of the integrating structure to project a well defined / repeatable pattern of light (ellipses) onto all elements of the FPA sensors either directly or by reflection off the inner surface of L3. Measurement of the resulting ellipses, their axes and orientation at a large glancing angles is very sensitive to the FPA position. Comparison with the reflected images off L3 provides information on the deformation of L3 as well. This technique requires the camera to be cold and fully functional to read out the images of the spots.

It is of course critical that any system we choose to employ is both more accurate and significantly more stable over time than the changes we are trying to detect. R&D will involve feasibility studies of each of these techniques (in the vacuum and -1000C camera environment), as well as an investigation into the ultimate stability of the measurement. We will use the prototype laser triangulation system that we develop, as a way of studying these four approaches.

These techniques could prove especially valuable if it is found that the FPA flatness stability is not sufficient, and we are required to employ micro-positioners in the array on each raft. These alignment systems would then provide the immediate feedback required for the positioners.

5.3.3.4 Camera Body Evaluation

Vacuum testing of the camera body will proceed in stages with they empty body receiving the preliminary testing. Thermal conduction measurements will be made on the camera body as well as monitoring of the overall pressure.

Considerations for contamination control of all aspects of the assembly process will also be developed.

5.3.4 Filter Development Program

In order to provide an initial specification to the vendors from which they can begin their design studies, we are currently considering the filter parameters as illustrated in Figure 5.3.4-1.

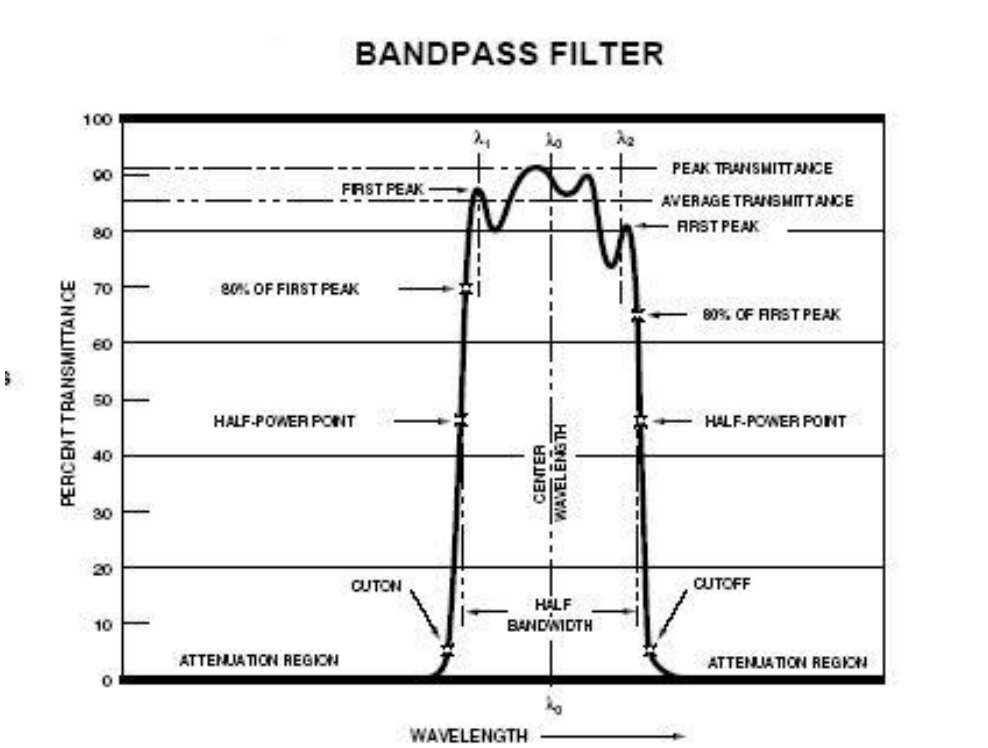


Figure 5.3.4-1 Illustration of generic filter characteristics

In order to assess industrial capabilities for supplying the coatings required for the LSST optics, LSST project representatives have visited and/or initiated discussions with several commercial vendors. The preliminary feedback from vendors indicates that there are no unsolvable technical challenges, and at least two U.S. vendors have previously coated optics of similar size with similarly complex coatings, as illustrated in Figure 5.3.4-2. Other vendors are also interested in extending their current capabilities to enable coating of optics of the required sizes.

Figure 5.3.4-2 Photo of a large (3 m class) coating chamber operated by a commercial U.S. vendor (left) along with a coated NOVA laser optic of similar size to the LSST optics (right).

Based on the initial assessment of the capabilities of commercial coating vendors, the following plan for completing the LSST filter coating procurement has been formulated:

5.3.4.1 R&D Plan for LSST Filter Set

1. Design Study
   - Define performance tradeoffs including shape
   - Coating designs, uniformity, repeatability
   - Define possible parameters to relax without compromising science
2. Risk Reduction Study
   - Engineering proof of concept.
   - Required uniformity and spectral performance developed and tested
   - Fabrication risks identified and addressed
   - Creation and analysis of witness samples
   - Develop final cost/schedule estimates
3. Production of Filters
   - Create handling tools
   - Coat filters

5.3.4.2 LSST Filter Coating Risk Reduction Study

The following research and development tasks have been identified that should be performed by one or more vendors in order to qualify them for a fixed price bid for the LSST filter coatings.

1. Establish procedures to distribute a uniform coating over the entire filter surface. This includes evaluating several coating techniques to determine best method of coating.
2. Set-up test procedures to measure optical performance of filters.
3. Determine optical quality of glass and coatings necessary for rejecting out-of-band transmissions.
4. Develop techniques to ensure wavelengths of pass band edges are met.
5. Establish ability to coat on two sides for spectral performance.
6. Determine exact substrate thickness to achieve desired performance goals. LSST will supply substrate material to vendor.
7. Monitor techniques to reduce variations.
8. Explore coating hardness
   - soft/hard coating vs. performance and cost
9. Spectrum shift with temperature (has some leakage with shift). Needs to be characterized and tested by vendors.
   - Spec at OºC. Measure shift at different temperatures.
10. Generate ROM.
11. Generate fiscal plan for coating filters.

5.3.4.3 Design Specifications for the LSST Filter Set

All specifications in this section refer to the finished filter as delivered by the vendor.

5.3.4.3.1 Design Characteristics:

- Beam that is incident on the filter has a focal ratio of f/1.25 with a 61.5% linear obscuration.
- The filter is concentric about the chief ray so that all portions of the filter see the same angle of incidence range, about 14.2º to 23.6º.
- At the filter, the sub aperture is about 100mm in diameter.

All measurements of filter transmission shall be calculated or measured using the above beam profile.

5.3.4.3.2 Detector Response Curve:

The detector response curve (QE) below is presented to allow calculations of out-of-band vs. in-band transmission. The goal shall be to have less than 1% out-of-band transmission.

Table 5.3.4-1 Detector response curve (QE)

400 nm	600nm	800 nm	900 nm	1000nm
60%	85%	85%	85%	40%

5.3.4.3.3 Substrate Material: Fused Silica, round, parallel surfaces.

Note that if colored glass is available in 770 mm dimensions and is desirable for the definition of the filter band-pass, this is acceptable. But manufacturer must specify the composition of the glass chosen ahead of time since it will require small changes to our optical design. LSST may provide fused silica substrates if the vendor considers all dielectric solutions.

5.3.4.3.4 Transmittance:

Transmitted average should be 95%±3%

5.3.4.3.5 Spectral Cut-on/Cutoff Slopes:

<5% (slope between 90% and 10% points on filter edge-See Table 2)

5.3.4.3.6 Out of band transmission:

< 1% below cut-off points

5.3.4.3.7 Spectral characteristics and coatings:

All of the filters should have good uniformity of transmission across their full clear apertures and should include anti-reflection layers to inhibit ghosting in the telescope. AR coatings on filter [R <0.5%] should span region where T > 10%.

Aperture: 770mm
Clear Aperture: 750mm

All filters meniscus with equal radii of curvature of 5.9m

Angle of incidence: 14 – 23.6 degrees

Structure: Laminate or Single Layer.

Filters can be made of more than 1 substrate if this aids in manufacturing and/or cost reduction while maintaining optical quality.

Transmitted Wave-front Error After Coating: ≤ 0.33 λ PV over any 4 inch diameter at 610 nm.

The finished filter surface needs to be flat to better than 0.33 λ PV over any 4 inch diameter subaperture to preserve the optical quality of the telescope.

Surface Wedge (parallelism): ≤ 45 arc-seconds.
A wedge of 1.0 arcminutes will result in a change of about 5% in the telescope RMS spots before atmospheric distortion is considered. The specification is set to stay below this limit.
Surface Quality: 60/40

Inspection of the filter surfaces is to be consistent with MIL-O-13830A. This specification applies to all substrate surfaces before and after coating and for external surfaces for the case of a

laminated filter. The argument for the relaxed surface quality specification here is similar to that given for the inclusion allowance. The filters are used far from focus.

Coating Durability: MIL-C-48497
Excluding tests for immersion in saline solution.

Operational Temperature Range for Filters: -5 to + 25 C

Total Thickness:

Thickness will be determined on a filter to filter basis and will be optimized for best PSF over entire filter (14-21mm thick filter). Thickness tolerance shall be ±0.25mm.

Environmental Stability:

Coatings shall be stable from -20ºC to 50ºC, humidity range >24hr@98%, slow #600 tape pull (no degradation in performance or λ shift).

Humidity Range: 5% to 95%

Lifetime: 10 years of use

Note that it is anticipated that the filters will reside in a dry air environment for more than 80% of this time.

Pinholes: 1/10 of 1% over total area

5.3.5 Camera Integration and Testing

5.3.5.1 Calibration Activities During the Proposal R&D Phase

The science requirements laid out for LSST become challenging when it comes to working out details of design, defining precise tolerances, physical modeling efforts and physical description of the actual parts. The following is an incomplete list that describes our activities during the proposed R&D phase.

- With the aid of our modeling capabilities, we will continue the pursue the (currently incomplete) task of determining the multi-dimensional tolerancing requirements according to the current strawman optical design, refered to as "3.5°–short". This will include defining the volume in parameter space, for example, that can result in at most a 10% degradation in imaging performance across the FOV given a perfectly flat focal plane. This effort will be largely complementary (yet parallel) to the "modeling" effort, which currently emphasizes atmospheric effects. Should the baseline design change in the future, all of our tolerancing work can be repeated in short order, since the input parameters can easily be changed.
- Items contributing to estimates of the magnitude scale zero point s_0 will be estimated, and where available, real measurement results from similar parts will be used. Particular emphasis will be placed on items with complex dependence on wavelength and incidence angle, such as transmittances through ARCs.
- Terms that contribute to the sky function $f(\lambda)$ (*cf.* Eq. 3.3.9-9) will be studied and collected to include in the candidate function, so that *a priori* estimates of its effect on the magnitude scale zero points may be made.
- Engineering grade CCDs delivered as part of the sensor R&D effort will be validated in terms of their voltage potential profile within the Si bulk. Such verification can be made using either soft X-ray illumination (via observation of photoelectronic charge cloud diffusion) or by illuminating the CCD with optical light on small regions at the

sub-pixel scale. Comparison of data to model predictions will provide means for validation.

- Exercises in performing the signal map conversion from raw frames, according to the reverse of Eq.. 3.3.9-1 will be performed on other stacks of raw CCD frames. If and when insufficient information is available to adequately perform the inversion, calibration measurements that can perform them will be devised and simulated. Delivered engineering grade CCDs could be used for this purpose, but are not necessary, since the emphasis here is on algorithms that linearize the data. Some controlled experiments will also be desirable, such as temperature and clocking dependence to the various charge transfer inefficiencies.

Candidate systems for calibration transfer will be considered. Calibration transfer systems must have simple, optical paths and detectors, and calibrated to better than the required magnitude scale zero point tolerance of 0.005 magnitudes and capable of measuring stars on the bright end of LSST's dynamic range (m_V) to this accuracy, on the time scale of LSST exposures.

5.4 Data Management R&D

LSST data management must achieve unprecedented levels of automated quality assessment and transient alert accuracy. It must process an unprecedented volume of data reliably and quickly. These requirements contain embedded technical risks, which are mitigated by the planned research activities described in this section.

5.4.1 Summary of Research Activities and Primary Technical Risks

These research activities and risks were introduced in section 3.4 above. They are summarized again here for the convenience of the reader.

Table 5.4.1-1 Summary of Data Management Research Activities and Primary Technical Risks

Research Activities	**Primary Technical Risks Addressed**
Optimal PSF estimation/reconstruction and co-addition algorithms Application Layer	Insufficiently Precise PSF Estimation/Reconstruction It will be difficult to determine PSF's sufficiently accurately to meet weak lensing and astrometry science goals. (Refer to the SRD for these goals.)
Optimal subtraction/detection/alerting algorithms with probabilistic classification of detections Application Layer	Unacceptably High False Transient Alert Rate LSST breaks new ground in the ratio of data rate to the number of humans available to examine it. How can we make it practical for humans to make use of real time alerts from LSST without being overwhelmed by false alarm rates?
Optimal object association and moving object orbit determination in densely populated	Inadequate Scalability in Object Association Pipeline

images Application Layer	The LSST data will contain unprecedented numbers of sources in each image and will observe areas such as the galactic plane with high concentrations of these sources. Readily determining which detections represent the same object across these images will be challenging for static (proper motion only) objects, and extremely so for transient objects (e.g. varying luminosity, moving objects. Etc.) Computational requirements are highest in this area of all the LSST pipelines.
Storage-efficient vs. access performance-efficient data structures tradeoff analysis for provenance Application, Middleware Layers	Inadequate Provenance Mechanisms The level of provenance information needed for effective re-processing of LSST data is unknown, but will certainly include not only the raw data, but also extensive status and control information from the telescope and instrument, software and hardware versions, and configuration used to originally process the data, and dependencies on other data.
Petabyte scale database structure and query optimization Application, Middleware Layers	Inadequate Data Serving/Query Performance Given the unprecedented size in terms of individual objects (3.0 * 10^9 or more), associated detections (10^2 * number of objects), and images, data access mechanisms must be very well crafted to avoid searching large, non-contiguous portions of the data unnecessarily. Database servers with sufficiently high performance may not be available when needed at LSST first light, particularly for supporting a high volume of complex science queries.
Extensible pipeline and catalog construction set prototyping Middleware Layer	Inability to Evolve Pipelines and Data Structures with Emerging Science, Instrument, or Telescope System Changes Over the course of Construction and Operations, new algorithms, new scientific requirements, and instrument and telescope system changes will occur. If the pipelines and data structures are not general or extensible enough, it will be very costly or infeasible to adapt them to these changes.
Scalability and fault tolerance in multi-teraflop data management pipelines and multi-Gbps data transmission Middleware, Infrastructure Layers	Unacceptably High Down-Time in Large Scale Computing, Storage, and Network Resources Hardware failures will be routine for LSST, due to the large number of CPUs and disk drives, and reliance on high-speed network

	connectivity. It may be difficult to create a system sufficiently robust to these failures.
Technology trend analysis, projection, and validation strategy Infrastructure Layer	Unacceptably High Cost of Computing, Storage, Networks due to Premature Technology Selection/Purchase The design of the DM architecture is influenced by the technology we expect to be available to implement it, starting with Construction in 2007 – 2008 and continuing through the principal survey period until 2022. We need to predict the characteristics of CPU, network, and storage hardware, and of database software, sufficiently well that our design is appropriate, and insulate the design as much as possible from underlying platform dependencies.

5.4.2 R&D Investigations

The R&D Investigations are directly driven by the risks above. Each R&D investigation is described below, in terms of:

- Overview and rationale
- Activity summary and relationship to other research activities
- Tasks
- Schedule
- Resources and collaborations
- Budget

5.4.2.1 Optimal PSF estimation/reconstruction and co-addition algorithms

5.4.2.1.1 Rationale and overview

As documented in the LSST Science Requirements Document and elsewhere in this proposal, the scientific missions require LSST to achieve a high degree of precision in flux, shape, and shear measurements. Given these requirements, the LSST must explore new approaches to co-addition and PSF estimation.

For example, weak lensing is one of the primary LSST science goals. The requirements for shear systematic errors for LSST are more stringent than for current surveys. We must identify areas where current algorithms would be weak when confronted with the LSST dataset, and develop algorithms that will stand up under the volume and resolution of the LSST dataset. While LSST is being designed for superb image quality control, the large number of exposures enables new approaches to systematics detection and control at the software level.

5.4.2.1.2 Activity summary and relationship to other research activities

5.4.2.1.2.1 Shear Measurements and Co-Addition

The LSST dataset has several properties that depart sharply from current weak lensing datasets. The most outstanding feature scientifically is the large number of exposures of each area of sky,

hundreds rather than one or a few. Each exposure will have a different value of seeing, atmospheric transparency, etc., seeing being the critical parameter for lensing. The problem has generally been framed as one of "image stacking," but we believe this is too narrow. We will drop the assumption of image stacking and investigate ways to extract the shear from the entire dataset. This approach has two great virtues:

5.4.2.1.2.2 Preserving information content.

Exposures with the best seeing have a much higher information content than exposures in typical or bad seeing, because there are many galaxies with intrinsic size roughly the size of the LSST point-spread function (PSF). When the PSF shrinks, many more galaxies become available for shape measurements. In the new paradigm, we measure all these shapes when possible, and combine all the shape measurements in a weighted manner. With image-stacking, one can try various tricks such as discarding the worst-seeing images or de-weighting them, but it is inevitable that galaxies which are resolved only in the best-seeing exposures become unresolved when stacked along with many other average exposures. Of course, image stacks will remain important for the LSST, and stacking algorithms may improve. But stacking is inherently a form of lossy compression of the dataset. Whether the losses are tolerable for shear measurements is one question this proposal seeks to answer, by comparing the results of the new approach with the results from a stack.

5.4.2.1.2.3 Providing external error measurements.

The variance in a series of shape measurements of the same source naturally provides an error estimate that includes not just photon noise, but also errors in PSF estimation for each of the exposures. A monolithic stacked image, on the other hand, has only a single PSF, and errors in that PSF are called "systematics" and promptly swept under the rug. There may be other systematics that also become quantifiable in the new paradigm, such as optical distortions and residuals in registering the exposures. The main drawback to this approach is that it is new. It must be demonstrated to be effective and computationally feasible by the time of the construction proposal. The LSST team is in a great position to do this, by virtue of its pre-cursor datasets, including Deep Lens Survey (DLS) dataset, which has 20–80 exposures per area of sky. This makes it the best LSST testbed among existing lensing surveys.

Once we have the software working, we will address computational efficiency. At first sight, our method looks very I/O intensive compared to image stacking: It revisits each source on each exposure, after making an initial stack for detection purposes. But this can be mitigated by a careful implementation that keeps in memory all the relevant input pixels for a small area of sky, from the creation of the initial stack, through the shear analysis. This is trivially parallelizable, because analysis of galaxies separated by more than a few galaxy widths can proceed independently. In any case, we will provide a realistic estimate of the computational burden of this method, for the construction proposal.

5.4.2.1.2.4 PSF Estimation

We will develop a prototype PSF (Point Spread Function) pipeline based on LSST stellar image precursor data and simulations served. The prototype PSF pipeline will produce discrete Point Spread Functions encoded as FITS images from calibrated stellar image CCD data. Algorithms for PSF reconstruction with wavefront sensor (WFS) data will be investigated with the collaboration of the telescope and camera teams; if the inclusion of WFS data can be demonstrated to significantly enhance the scientific return of the LSST project, PSF reconstruction using WFS data will be implemented within the prototype PSF pipeline.

The telescope and camera response will determine the fundamental PSF parameters, therefore this activity will require close interaction with the LSST telescope and camera team research activities as described elsewhere in this proposal.

5.4.2.1.3 Tasks

5.4.2.1.3.1 Initial shear extraction algorithm development

Demonstrate working algorithms on simulated data without noise or PSF estimation errors.

5.4.2.1.3.2 Extension to noisy data and data with mis-estimated PSFs

Prepare simulated dataset with realistic errors for a "bakeoff" with BJ02. Make some speed optimizations so that simulated dataset can be processed in reasonable time. Begin analysis of simulated dataset. Specify how DLS data should be loaded.

5.4.2.1.3.3 Analyze simulated dataset and comparison with BJ02

Further speed optimizations so that DLS dataset can be processed in reasonable time. Stack DLS dataset so that BJ02 can operate on it. Begin mass reconstruction comparisons. Load DLS data into prototype database.

5.4.2.1.3.4 Compare algorithms on DLS dataset

Make fixes for errors not seen in simulated datasets, such as cosmic rays, asteroids, blended galaxies, etc. Explore further speed enhancements via parallelization, if needed to accurately estimate computational burden.

5.4.2.1.3.5 Complete mass reconstruction activity and estimate LSST computational burden

Prepare preliminary report based on in-house computing facilities. Optimize algorithm code and prototype database integration. Estimate computational burden for each method and prepare final report.

5.4.2.1.3.6 Prototype PSF pipeline

Develop a prototype PSF (Point Spread Function) pipeline based on LSST stellar image precursor data and simulations served.

5.4.2.2 Optimal subtraction/detection/alerting algorithms with probabilistic classification of detections.

5.4.2.2.1 Rationale and Overview

As defined by the LSST Science Requirements Document and elsewhere in this document, the science mission places high demand on the LSST's ability to rapidly and accurately detect and classify varying objects (those with proper motion and optically varying properties, as well as moving objects) and to achieve a low degree of false alarms. Given the very high data volume produced by the LSST, the corresponding large number of detections in each image (10^5??? objects detected per image), as well as the likelihood of entirely new classes of transients, the LSST will not be able to rely on traditional labor-intensive validation of detections, classifications, and alerts. In order to achieve the levels of accuracy required, new algorithms for detection and classification must be created, as well as innovative automated techniques for alert filtering and validation.

5.4.2.2.2 Activity summary and relationship to other research activities

In this research activity, we will evaluate current pixel-level algorithms for subtraction of image data, and will develop and refine new approaches to these problems.

We will investigate ways to construct an appropriate variance array in association with a stacked or subtracted image. We will develop metrics to allow useful comparisons between different algorithms and codes. The result of this effort will be a chosen set of optimized algorithms for the basic pixel-level operations. Once the optimized algorithms have been chosen, we will implement them in a prototype photometric pipeline, and use it to re-analyze precursor data sets, providing opportunities for new science (e.g. from deep co-added images) well before LSST sees first light.

In order to carry out this plan, we will require a suite of test data including idealized data produced by analytic simulations, time series of data produced from simulated images with specified variable components, and finally actual precursor data sets of real time series data. These test data will be used to characterize and compare the performance of the algorithms under consideration.

As described in the related research activity in section 5.4.2.3, the scalability and fault tolerance of the prototype pipeline will be evaluated in a parallel processing environment.

5.4.2.2.3 Tasks

5.4.2.2.3.1 Image test data suite

The main purpose of the test data suite is to enable the efficient, robust, and repeatable testing of numerous image processing algorithms. It is our intention is to design the test data suite and provide suitable documentation so that it will be a valuable open resource for the entire LSST collaboration and similar survey projects. The first step in assembling the test data suite is the creation of a completely idealized time series of data. These data include only user–specified objects, such as Moffet functions for stars, Sersic profiles for galaxies, and artificial noise. Various time series of data should be produced from convolution of these images with the expected spatial (and/or temporal) variation of the LSST point spread function (PSF), to model the range of expected image data from the LSST. Sets of images should be produced with the anticipated cadence and depth of the LSST science images. These idealized images will allow for fundamental tests of the algorithms where the truth is known, and will be used to optimize accuracy and completeness. At this stage, any inefficiencies and systematics can be localized to the algorithms themselves.

The second stage in the test data suite will be the creation of a time series of images created from user specified variable components (e.g. a star varying with a periodic light curve). This will gauge the response of each algorithm to variability of non-idealized objects. Importantly, once this stage has been thoroughly developed, it may also be added as a component to the prototype to measure real time detection efficiencies. While many on going surveys have recognized the need for such real time efficiency analysis, this has rarely been implemented in practice. This component will be a requirement for the LSST.

Finally, the test data suite will include time series of data from the following precursor data sets:

- SuperMACHO - 1.5 TB of images to V + R "<24 and a cadence of 0.5 images per night. ESSENCE - 1 TB of images to R,I "<23 with a cadence of 0.25 images per night.
- SDSS - 2 TB of 5 color data to r "< 22.5. Approximately 100 deg 2 with up to 40 epochs (and an additional 100 deg 2 with of order 10 epochs) separated by timescales of 3 hours to 4 years.

- DLS - 1.3 TB of data in B, V, R, z to "< 24 with a staggered cadence. This includes 3.7 deg 2 days at 1000 sec intervals, and 0.2 deg2 days at 105 and 106 second timescales.
- LONEOS - 4.5 TB down to 19th magnitude, with a cadence of 4 images per night, separated by 30 minutes. A 2.5 TB subset of these data includes 14,000 deg 2 with more than 40 epochs, and 1,700 deg 2 with more than 100 epochs. We will be able to use this data to explore the impact of spatial under-sampling upon the LSST data.
- MACHO – 7.3 TB down to V, R < 22. This includes 91,000 images spread across 40 deg^2 of the LMC, 3 deg2 of the SMC, and 45 deg2 of the GalactIc bulge. Sampling ranges from about 100 epochs to more that 1000.

These representative series of photometrically and astrometrically calibrated images sample a variety of imaging depths and band-passes, observing conditions, source densities and cadences. We expect the overall volume of the full, "portable" test data suite to be at least several TB and possibly much larger, as more data become available and our capability to absorb and process them increases.

5.4.2.2.3.2 Image subtraction/detection algorithms

5.4.2.2.3.2.1 Algorithm Overview

The difference image analysis of an input image can be split into the following modular tasks:

1. Determine the overlap between an input image and archived template image using the available WCS information; we shall investigate whether there is a need to resample at this point, or whether this can be done as part of the kernel convolution.
2. Find and apply the convolution kernel that best matches the PSFs of the template and input image, and subtract pixel–wise the template from the input image.
3. Detect and measure objects in the difference image.

Each stage will be designed to monitor explicitly the propagation of noise and inefficiency. Doing so will allow for correct thresholds and precise knowledge of the limitations of the dataset, an absolute requirement for drawing scientific conclusions from the data. In those stages where images are convolved, we will examine the effects of ignoring pixel covariance on the science goals.

5.4.2.2.3.2.2 Algorithm Details

We intend to register images based upon their relative WCS information. The WCS projection will either be used as a strict mapping, or as a starting point for re-mappings with another functional form. We expect the geometric distortions of the LSST telescope to be non–negligible, and potentially to vary with time. Therefore, we must accommodate for changing astrometric solutions for our input images. We must also explicitly correct for non–uniform sky area sustained by each image pixel due to astrometric distortion — i.e making our images "flat" in terms of sensitivity — and will investigate sky matching algorithms in the presence of the scattered light which plagues most wide angle imagers. To enable the appropriate co–addition and differencing of images, we will investigate re-sampling images to an equal–area projection (e.g Aito), which will conserve surface brightness and allow summing or differencing pixel values to measure fluxes. Several packages will be examined for this image re-sampling module, including Swarp, Montage, and internal software in use by the DLS, SDSS, and SuperMACHO collaborations. Using the test data suite, we will quantify properties such as accuracy, sensitivity to under-sampled data, and computational requirements for each package. A direct comparison of these algorithms on the same data set is not currently available.

The analysis of image subtraction will be centered around the Alard and Lupton (1998) method. While this method has revolutionized the field of variability studies, there are several possibilities

for substantial improvements. First, the convolution kernel can be adjusted to better represent the data. In general, the expected variation of the LSST PSF may require specialized kernel representations. We will build this flexibility into our algorithms.

Second, the treatment of the spatial variation of the kernel coefficients may be improved. Current methods to model spatial variation (polynomial functions xi °— yj) have not evolved since the first introduction of Alard and Lupton-based image subtraction, and have been shown to fail under a variety of conditions; we plan to quantify these failings and design improved alternatives.

The optimal detection and measurement of objects in a difference image is crucial for the scientific success of the LSST time–domain science. There are a number of existing codes including (DAOphot, DOphot, SExtractor, and photo). It is not clear whether any of them can fulfill all of the requirements on accuracy, bias, speed, and robustness. In particular, the majority of photometric codes were developed for direct images (and in some cases direct images dominated by stars),while difference images have substantially different properties – for example, the noise has very different behavior. For this reason, we consider the correct propagation of noise throughout the image convolution processes as an absolutely necessary step. Our object detection algorithms will be designed to accept and make use of input noise information.

5.4.2.2.3.3 Prototype image processing pipeline

We plan to build the prototype pipeline around existing codes, such as PHOTOPIPE (Super-MACHO), PHOTO (SDSS), the DLS pipeline, and the ESSENCE pipeline. LSST scientists are principals in these projects and have access to the source codes (in the case of SDSS, the code is in the public domain).

One of our aims is to allow all members of the LSST collaboration and eventually entire the scientific community to share algorithms and codes. With this in mind, we propose taking the algorithmic cores of these pre-existing pipelines (mostly implemented in C) and glue them together using a common set of data structures. We expect to be able to use the Pan-STARRS utility library psLib, and to use SWIG to bind our modules together using python. During the construction of the prototype pipeline, we will investigate the interdependence of the algorithms.

5.4.2.2.3.4 Transient alerts

The first line of defense against false alerts will be user definable alerts that will allow false alarm rates to be specified. We will test these concepts with the prototype pipelines being developed during R&D as described above. We will also examine transient alerting mechanisms already in use in the robotic and conventional telescope domains as well as participate in the activities of the VOEvent working group of the IVOA, with particular focus on defining a probabilistic approach to transient classification/notification. (Expand on this approach)

5.4.2.3 Optimal object association and moving object orbit determination in densely populated images

5.4.2.3.1 Rationale and overview

As described in the LSST Science Requirements Document (SRD) and elsewhere in this document, an important mission of the LSST is to provide information and alerts on moving objects.

In order to track moving objects with or without known orbits, the attributes (e.g., brightness, color, position, velocity, etc.) of the observed objects need to be stored in the catalog. Most of the attributes are time-varying. For example, continuous movement of an object causes a change in

its position and may also change its velocity. The brightness of an object may also change over time due to various factors.

For each catalog entry, we must store references to the observations that are linked to that object. At any point in time, there will be some objects with uncertain linkages to other objects. This will need to be represented in the catalog. Further, some of the stored quantities, such as orbital parameters and their uncertainties are a function of the linkages, and thus we will need to represent the hypotheses.

Thus key algorithmic challenges include the creation of data structures to support effectively searching for possibilities, and ordering the search, so that they can be identified as earlier as possible. The problem at hand is characterized by low dimensionality of the representation (e.g. the number of parameters for a given orbit is not large), but that there are many orbits to be found, there is significant correspondence ambiguity in the data, and many orbits have insufficient observations. This suggests developing data structures for effective representation of votes for orbit models given probabilistic estimated trajectories to help structure the search. Proposed models then would be evaluated probabilistically.

The LSST cadence implies a high update rate of the catalog and continuous addition of data obtained from new observations into it. Special care must be taken to organize the catalog to manage the ever-increasing amount of data for moving objects in such a way that the size of catalog does not grow too fast over time and the catalog lookup is fast.

Being able to correctly locate moving objects in the past and present times and in the near future can help track the objects and extract candidate orbits more efficiently and accurately. We must develop algorithms and indexing structures so that we can search objects from the catalog in a scalable manner. Object searches can be done by their positions at a time point or during a time interval, by their trajectories, by their time-varying patterns of change in brightness, and so on.

The notion of versioning the records of moving objects will help us organize the catalog in a storage efficient way. This will keep the catalog from growing too fast and keep the cost of lookup in an acceptable range. We must also investigate whether a simple computational or a parameterized model can help on this. For example, as long as the velocity of a moving object does not change, the position of the object does not have to be recorded in the catalog every time it changes its positions.

5.4.2.3.2 Activity summary and relationship to other research activities

We will address the question of linkage strategies for main belt asteroids and near earth objects and how these strategies impact the observing cadence. For this purpose we will not consider slowly moving sources nor contamination from high altitude satellites.

We will develop algorithms and data structures to deal with objects whose observations are difficult to link in different LSST images due to their motion or time dependent brightness. We will develop these methods in the context of a catalog for moving objects which supports a variety of standard and probabilistic queries. A key component of what is stored will be probabilistic constraints on orbits of the objects. Being able to effectively query and update such a catalog leads to non-trivial data management problems. We will apply our expertise in the domains of computer vision, statistical modeling, geometric algorithms, and spatio-temporal data management to provide core algorithmic support for the realization of such a system.

To provide the probabilistic components of the algorithms we will be guided in principle by a Bayesian approach. In particular we will work with domain scientists to understand how proposed trajectories and orbits (i.e. models) lead to probability distributions over observations. Bayes rule, together with any priors we wish to use, yields posterior probabilities of the models given the data. Brute force application of this inference process is generally not tractable, and thus a variety of approximations and data reduction strategies will be investigated. In particular, in this domain

we expect that verification of a number of models within reasonable certainty will be possible (e.g., some orbits will become well established), and representing these as known quantities as soon as possible will simplify the computation.

All analyses will be undertaken using in house simulations that will reach the effective depth of LSST (incorporating position and timestamps for known orbits but no magnitude information). The deliverables for the initial phase will be the code based on the algorithms described above, documentation for its use, a report on the efficiencies of the algorithms and their suitability for the LSST associations and a set of test data that demonstrate the properties of the algorithms.

Because future events can influence derived values (e.g. orbital parameters and their uncertainties) there are some special challenges to the maintenance of the catalog. It should be possible to accurately track the history of the changes in value, and the reasons for the old and new values. While this will largely be a new observation, the logic may be indirect, in that a new observation of a different object may reduce the probability that one of the links is correct. Consequently, this activity is closely related to the research activity described in section 4.4.2.6.

In the analysis of the prototype we will also need to study how implementation issues affect the algorithms and data structures. Such issues include implementation of processing and queries on parallel machines, size of memory caches, laying data on disk for fast access, and so on. Therefore, this activity is also related to the activities described in sections 4.4.2.2 and 4.4.2.8.

5.4.2.3.3 Tasks

5.4.2.3.3.1 Multi-hypothesis tracking algorithm

We will develop and deliver a multi-hypothesis tracking (MHT)algorithm based on quadratic tracks that is capable of analyzing asteroids at a density expected for LSST data. We will deliver ascoring metric for determining the efficiency of these approaches interms of the percentage of asteroids found and false positives. All tracks returned will be rank ordered based on the likelihood that they are orbital tracks.

5.4.2.3.3.2 Multi-tree data structures

We will develop and deliver data structures and algorithms based on amulti-tree (MT) approach to linking asteroids. These data structures will be designed to remove the need for close pairs of exposures(i.e. they will not require velocity information). We will characterize the efficiency of this approach as a function of astrometric accuracy and compare with the efficiency and speed on the MHT.

5.4.2.3.3.3 Cadence analysis

We will study the sensitivity of the MHT and the MT approaches to the cadence of observations (including the time separation between closepairs of exposures and on separations with timescales of several nights). We will deliver recommendations on the maximal spacing of observations expressed in terms of the fraction of asteroids recovered together with the number of false positives.

5.4.2.3.3.4 Bayesian priors

We will determine how well Bayesian priors (magnitude, accelerations and orbital parameters etc) improve the efficiency of the MHT approach. We will also develop data structures to incorporate associations into a database structure. These association structures will be designed to enable efficient searches for tracks within a database, merging of multiple tracks into a single orbit and for the dynamical growth of tracks with new data.

5.4.2.3.3.5 Data structures and query mechanisms

We will provide a methodology and prototype software for building a catalog of moving objects including orbit with error (if known), and probabilistic constraints on orbit (if uncertain). These probabilities will be updated as new events are observed. When sufficient events are linked, the orbits will become certain enough to either link or exclude all observed events, and the orbit will become known (with error). The catalog will be continually verified and updated for internal consistency and with respect to any available additional information. In particular, while the observation of an event will lead to integration of that event, the lack of an expected observations also needs to be considered. On occasion it will be necessary to automatically merge multiple objects, initially represented as distinct due to large uncertainty of their orbits, into one object. The catalog will support a number of standard and probabilistic queries as discussed further below.

To support the construction of the catalog, we will develop algorithms to provide probabilistic estimates for association of moving objects in sequence of images of the same part of the sky taken a close times according to a schedule designed by the LSST project to facilitate this. Because of this design, executing the association is expected to be relatively tractable, but a probabilistic approach will help extract maximal information from ambiguous cases, and, equally importantly, propagate it through the system. To the extent that it is helpful, we will consult the catalog of moving objects to help reduce ambiguity. In the worse case, linking ambiguity will persist as weaker than desirable constraints in the orbits of both objects, which will be resolved as additional events are observed (or not observed).

We will develop algorithms for determining the likelihood that catalog objects explain query observations, or a set of linked query observations (as might come from the process discussed in the previous point). Intuitively, this query answers: “what is this moving or changing object?”. Such a facility is clearly both useful to domain scientists, and to the construction and update of the catalog. The linking in (2) provides a fit for a trajectory, but to compute an accurate orbit the observations need to be linked with others at a substantively different time.

We will also provide algorithms for the determining the likelihood that a catalogued object will appear in a given image. We will further provide a mechanism for improving the performance of these queries in the case of a known sequence of multiple queries such as might occur due to a predetermined LSST viewing schedule. Intuitively, this query answers: “which moving objects will be seen here”.

The catalog we envision will provide the substrate for a wide range of standard and probabilistic queries. In collaboration with the domain scientists, we will develop algorithms and processes to support queries that are useful to the community with appropriate speed/accuracy tradeoffs.

5.4.2.4 Storage-efficient vs access performance-efficient data structures tradeoff analysis for provenance

5.4.2.4.1 Rationale and overview

The LSST will provide a raw image dataset unprecedented in sky coverage, depth, and time epochs. As this data is processed, extremely rich catalogs of objects will be created, as described in section 3.4.2. However, in order to be most useful, it will be important for the users of these catalogs to understand how the data were created: under what seeing and weather conditions, what was the system state at the time of capture, what were the operational configuration and circumstances, what images were used, what algorithms, and so on. Virtually any fact, from the precise pointing of the telescope and state of the focal plane at the time of capturing an image, to

the version of the firmware in the floating point processors on the computers is potentially relevant to determining whether the data are useful for a given application.

A critical key to efficient and effective archive use is the deployment of databases, and ingest rate demanded by the LSST is unprecedented, even without provenance information. In one experience with the CARMA telescope system, we found that the monitor data stream not only dwarfs the scientific (radio visibility) stream, but it also taxes the database's ability to load the monitor data into the database without falling behind. We will ensure that the provenance data collection will be loosely coupled with the "core" pipeline processing, to permit separation of the workloads as much as possible.

As a synoptic survey, the LSST project implies that all catalogs will have a dynamic nature to them. Even the catalogs of extra-solar, non-varying objects will likely be refined as the seemingly static sources are re-observed and errors are improved.
We must therefore design the LSST to manage multiple versions of catalogs, perhaps simultaneously.

There is a natural tradeoff between storing the minimal information needed for a given application, and calculating all derived data on the fly every time it is needed, versus saving some or all of the derived data for later use. We will explore this tradeoff in the context of provenance in this activity.

Many archives today, (including the BIMA archive) use XML in conjunction with databases to manage and export provenance/metadata information. Now that practical tools are emerging, RDF-based frameworks (e.g. OWL, RDF-S) are gaining prominence in the digital library world. For example, the National Earthquake Engineering Systems Grid (NEESgrid) has made successful use of an RDF-based metadata management system developed at NCSA. Such an approach may be a good match to managing high-level survey status information as well as describing catalog entity relationships (e.g. to object class types).

Designing the LSST data structures to track provenance sufficiently well that reprocessing can be undertaken in a routine way or that new applications can effectively utilize the data is therefore a requirement. Data provenance must be an integrated feature of the LSST data management architecture.

5.4.2.4.2 Activity summary and relationship to other research activities

We will design a flexible model encompassing all aspects of provenance/operation including site conditions, telescope and camera status and calibration parameters, and status of real-time pipelines and infrastructure. A prototype of the model will be created to verify that it can be saved at an appropriate rate in the provenance data structures, without adversely impacting the performance of the image and object pipelines. Suitable monitoring/mining techniques for these data will be designed to flag trends or problems, and to feed downstream pipelines and the upstream OCS. This design will be coordinated with the camera, telescope, and OCS teams.

We will track the IVOA's activities in this area and evaluate whether existing models are sufficient for our needs. In any areas where the VO will not suffice as currently envisioned, we will either influence the VO community to expand in the provenance area, or define additional mechanisms specific to LSST.

The NOAO Science Archive (NSA) architecture provides a good environment for experimenting with new metadata management techniques as well as for understanding how metadata will be used to drive services and processing.

The Data-to-Knowledge package developed at NCSA is a flexible toolkit for data mining. We will evaluate its use on actual astronomical data to determine how to derive provenance-related information from the data in ways not necessarily planned a priori during design.

Note that this activity does not include data quality assessment (DQA) for downstream data products except where there is a need for input to the real-time systems. One example would be input to the observing scheduler needed to prioritize fields. If a given field has too few high quality observations needed for weak-lensing science the scheduler would have to have this information. Determining whether this aspect involves direct assessment of images or linkage to downstream DQA is to be part of the design process.

5.4.2.4.3 Tasks

5.4.2.4.3.1 Provenance model

Define the parameters needed to characterize the provenance data. Assess how deeply to go into the telescope, camera, and data management infrastructure and operations information to capture the essential state needed for re-processing and new applications.

We will focus on the portion of the LSST data model that describes the major products of the survey and the state of the survey at any given time. We will examine the issue of multiple or "pending" object classifications. We will use these issues to build requirements for a general metadata management system that we can then prototype.

5.4.2.4.3.2 Feedback model

We will design/develop a scheme to provide system status and data quality assessment with real-time feedback to the OCS. We will incorporate event monitoring, trends, re-calibration alerts.

5.4.2.4.3.3 Provenance prototype

Prototype the provenance data structures and processing. Integrate with pipeline prototypes and evaluate performance relationships. As part of this prototyping effort we will look at the current state of commercial and open-source databases and their ability to scale to large volumes and rates. We will explore ways to take advantage of bulk data capabilities, indexing, and stored procedures to achieve sufficient performance. We will also look at approaches to table schemas that balance disk storage overhead with good performance, particularly when it comes to handling versions.

Database replication for the benefit of archive mirror

5.4.2.5 Petabyte scale database structure and query optimization

This activity will perform a tradeoff analysis of storage-efficient vs access performance-efficient data structures for temporal vs spatial, data quality assessment, and community science dimensions Rapidly changing database technology and uncertainty whether any database technology will in fact be able to meet the LSST needs, it would be unwise to lock in any of the existing systems with the hope that the vendor will sufficiently scale the system to LSST needs. Some of the LSST requirements are unique to the scientific community (data immutability is one example), so it is unlikely any vendor will change their product to take advantage of LSST-specific features unless it is proven to be commercially viable. This leaves LSST with no choice but to explore a system architecture using preferably open-source database engines, to achieve efficient and cost effective access to trillions of objects added each night and petabytes of data.

To deliver a robust, scalable, and high-performance database system, a significant effort needs to be put into researching exiting database engines and prototyping based on the most promising candidate(s). It is also essential to understand as much as possible how the system is going to be used (load, common queries, type of queries, access patterns).

A high-level roadmap to building a high-performance system includes:

1. Prototype pipelines/applications/queries that stress the data base/file system in the ways we anticipate
   - Identify database system and archive requirements.
   - Choose a “promising” database engine.
   - Build a basic prototype, focus on schema.
2. Create the overall data access architecture
3. Enhance the prototype: tune & scale it, identify & remove limits, make it fault tolerant.
4. Validate using LSST science related queries.
5. Investigate other database engines, do data challenges with the vendors of each technology
6. Make a 'final' choice.
7. Build a production system.

5.4.2.5.1 Identifying database requirements

Database system and archive requirements will be identified using a space of data products, users, and science cases/missions. The process will include:

- developing a representative set of use cases
- mapping the use cases into the database queries on the data objects
- understanding database load
- understanding access patterns

Predicting precise load and access patterns is close to impossible, our goal is to come up with a good approximation.

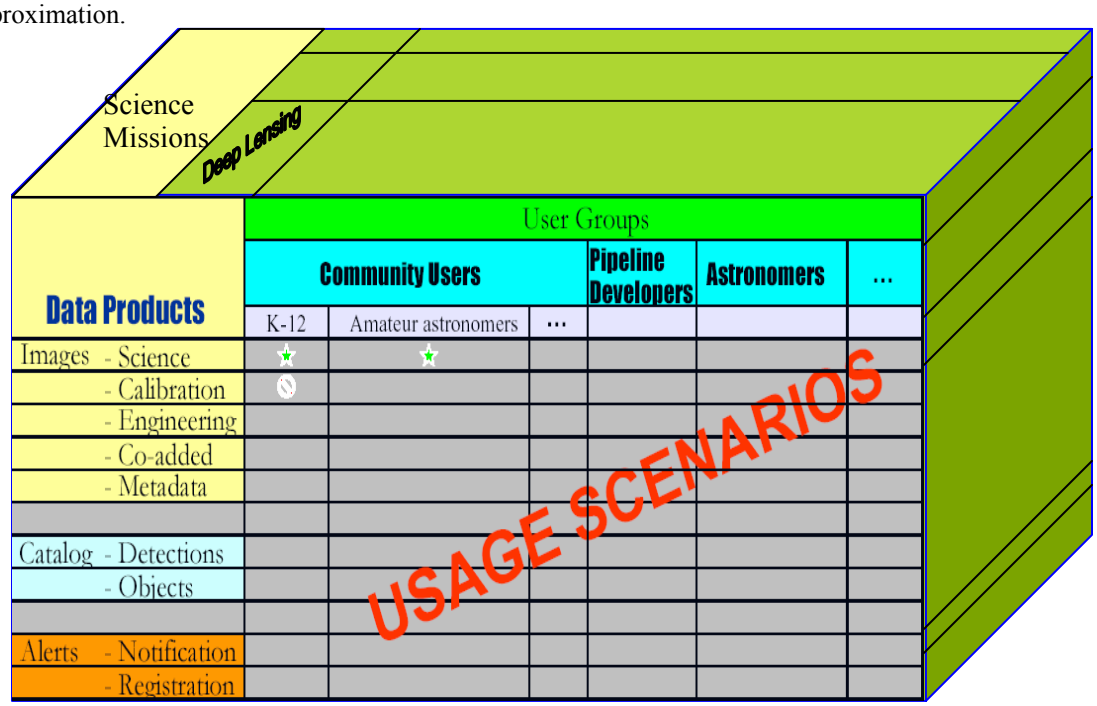


Figure 5.4.2-1 Illustration of a 3-D concept for handling queries.

Figure 5.4.2-1 demonstrates the idea of utilizing the 3-D space to identify and enumerate queries that could cover the space of a science mission, data objects, and users. Examples of science missions are shear calculation of weak lensing and alerts for micro lensing and other interesting phenomena. User groups cover the anticipated user of the LSST database and archival system. Data products correspond to data objects that will include raw images, derived images, detected objects, etc. We already started doing the

analysis and enumerating queries detailed use cases and queries in a text form that will later be translated into SQL queries.

5.4.2.5.2 Choosing a database engine for prototyping

Two major database technologies exist today: relational and object-oriented. Given the minuscule market of object oriented databases (< 0.5 billion $), and lack of a vendor with a predictable long term future, initial focus is put on relational technology. Object oriented databases will be seriously considered only if we are unable to prove that a relational database is capable of meeting LSST needs, or, if OO databases significantly grow in the next couple of years, which seems rather unlikely.

One of the most "promising" relational database engines and certainly a very convenient one for prototyping is an open source database, MySQL. It is preferred over non-open-source systems for many reasons, including:

- allows maximum flexibility, and opens possibility of tuning for LSST needs
- simplifies system maintenance at collaborating sites
- cuts costs
- avoids problem with vendor bankruptcy.

MySQL was picked over its main open-source competitor PostgreSQL due to its popularity and performance. The MySQL software is also well modularized, which should simplify code changes in case we need that.

5.4.2.5.3 Building a prototype

We started building a prototype by defining data model and schema needed to support the expected queries and data accesses, reusing the previous work and experience from other astronomy projects.

It is natural for the schema to evolve overtime at this early stage. We will take this into account and will:

- make sure the schema is sufficiently extensible,
- create an interface to isolate users from the schema details.

The basic structure of the database schema is derived from the well defined hierarchy in the data "Image→Detections→Objects→Classified Objects". This sequence occurs naturally in time-domain astronomy (e.g. SuperMACHO Database), and lends itself as the basis for the table structure. The extraction of Detections from the Images usually takes place at the latter stages of the image reduction pipelines, once the necessary preliminary stages (etc.) have been completed. The detections are unified into Objects by clustering algorithm, which may run either concurrently with the data ingestion into LSST DB, or at specific recurring time intervals on a static instance of the database (e.g. between the observing runs). Finally, the object classifications are deduced using various analysis codes (light curve, motion, etc.) based on the existing object separation.

5.4.2.5.4 Enhancing the prototype

To achieve scalability and performance it is essential to use data partitioning and index partitioning and query parallelization.

MySQL allows table partitioning through "MERGE Storage Engine": a collection of tables can be viewed as a single virtual table. Further tests need to be done to determine its scalability, robustness and performance.

Query parallelization in MySQL is achieved by running parallel processes on a multi-CPU server machine. Ability to parallelize across multiple machines, not only across multiple CPUs within a single server is essential when dealing with petabytes. It is a feature continuously requested (already) by MySQL customers running larger installations today, so it is conceivable it will be part of the product by the time LSST starts production. We are in contact with the key-people at MySQL, and are considering putting some effort into improving the parallelization in MySQL if this becomes a deciding factor.

To address fault tolerance and high performance issues, we expect to reuse extensive BaBar's experience with managing large data volumes. Many sites, including CERN (LHC experiment), FermiLab (D0 experiment), BNL (RHIC experiment) and IN2P3 (AMS experiment) are using, or are planning to use the data access server developed at SLAC for BaBar. The server, called xrootd was build to address exactly the issues discussed: providing ultimate fault tolerance and performance for a system with petabytes of data, hundreds of servers and thousands of disks. We will investigate how to incorporate xrootd into the LSST system, in particular how to use it underneath the MySQL database engine, as this would immediately allow us to use all of its features including automatic load balancing, redirection and Mass Storage bindings together with MySQL.

5.4.2.5.5 Validating a prototype

The prototype will support LSST science related queries. These queries will be used to test the functionality and scalability of the LSST database prototype. It will be used to validate all aspects of the system, including :

- schema design,
- database system design,
- scalability at least up to expected average and peak data volumes and rates,
- performance,
- fault tolerance,
- functionality and coverage of the identified use cases.

A set of simulated LSST images has to be developed to help with these tests. This does not mean that we need to wait for these images to be developed. We can simultaneously use some of the precursor data to perform these tests. We believe that there are enough work and research ideas to explore before we start doing a full scale scalability tests that require the use of realistic LSST simulated images.

5.4.2.5.6 Investigating other database engines

The effort related to investigating other database engines widely depends on the outcome from the first prototype. The possible scenarios include:

- the prototype proves to be very solid and is likely to meet LSST needs
- the prototype is good, but likely some other database engine would be a better choice
- the prototype proves that chosen solution is not sufficient: e.g. problems with scalability, and/or fault tolerance, and/or performance. It may be a problem with the chosen engine (MySQL), the chosen database technology (relational) or bad decisions made during building the prototype based on mischaracterized requirements.

Depending on the outcome, we may decide to:

- adopt the solutions chosen during prototyping, and closely monitor how the chosen engine evolve, and what other database vendors offer
- rethink used approaches/architecture

- migrate the prototype to a different relational database engine, or use object oriented database technology

5.4.2.5.7 Making final choice

The final choice will include choosing database engine and the database architecture. The decision needs to be made by the end of the year 2007. Waiting much longer would not give us sufficient time to develop a solid production system, while making the decision much earlier would be premature.

5.4.2.5.8 Building a production system

Using our findings from the prototype efforts we expect to deliver a first fully functional version of the production system in 2010 (?).

5.4.2.5.9 Matching observations to hypothesized orbits and building the moving object catalog

As new observe rations are made, and new links among them are proposed, our understanding of orbits and the associated data structures need to be updated. At the heart of the process is matching moving object observations to existing candidate orbits in real time. As a consequence of doing so, orbits will be updated, and their probability of existence will be revised.

We will maintain a data structure of orbits so that associated observation hypotheses for a given image, and reciprocally, moving objects for an observation, can be quickly retrieved. Each orbit is associated with a probability of existence based on how well it fits the data, and how many observations that it predicts did not occur. The parameters of the orbit will also have associated errors computed from fitting errors and error propagation from the data (itself due to a large variety of sources including viewing conditions, sensor noise, and frame differencing artifacts). The catalog will support queries with associated confidence levels, and only moving objects that have greater certainty than the requested level would be retrieved. The retrieved orbits (if any) are able to explain the query with the specified certainty.

The observations themselves are assumed to have been identified as candidates for moving objects, based on frame differencing observations. However, as that process is projected to be imperfect, we will further associate with each observation a probability that the object is in fact a moving object. This gives us flexibility in how conservative the frame differencing operation needs to be.

From the above, we can compute the probability that a particular orbit is the cause of a particular observation. We then propose to compute one or more good matchings between these sets of points under the assumption that there is a one-to-one correspondence between observations and orbits, with allowance for outliers. We will first consider variants of bipartite graph matching such as max-weight matching, with a natural weight being the logarithm of the probability of each candidate match. These types of algorithms are more attractive than greedy matching because they are able to optimize a global error measure. This approach tolerates some amount of noise, due to the fact that outliers can be handled gracefully and the error is accounted for globally.

The output of the algorithm consists of set of matches, i.e. pairs of orbits and corresponding observed points. Of course, we update the relevant data-structures accordingly. The algorithm also produces a list of outliers, namely observations that we were unable to explain with any existing orbit with high enough confidence. These unmatched observations then become the focus of additional effort to be linked, as well as being available as the target of database queries.

It should be clear that the machinery developed to solve the above problem will provide the capability to efficiently satisfy a large number of queries related to moving objects. For example, "Which objects are we going to see at a certain region of the sky, at a certain frame of time, with confidence higher than some threshold". We need to develop very general capabilities of this nature for the ongoing improvement of the current understanding of the existing orbits. We propose that these capabilities are also exported as services, both to other data consumers in the LSST project, other astronomers, and the public at large.

5.4.2.6 Extensible pipeline and catalog construction set prototyping

5.4.2.6.1 Rationale and overview

The LSST construction and operations period (2008 – 2022) will surely see advances in scientific algorithms relevant to the LSST pipeline processing and discoveries that are not now determined. In order to take advantage of these advances and discoveries without undue cost, the LSST pipelines and catalog structures must be mutable and extensible.

This is particularly important in the context of community science. We can enable the use of "core" LSST pipelines that are created within the controlled environment of the LSST project simply enough; it is a much more challenging prospect to permit other pipeline modules, created outside the LSST project, to be used in conjunction with the core pipelines. This requires a formal architecture for extensibility, with documented interfaces and services supporting construction, configuration, deployment, execution, and monitoring.

5.4.2.6.2 Activity summary and relationship to other research activities

The primary activity is the creation of a Pipeline Construction Set that allows pipeline modules to be practically "plugged in", especially over a network connection. Similarly, permitting new data types and associated meta-data definitions to be added to the catalogs during operations. We will draw on technologies in existing astronomical pipelines as well as the grid environment to extract and scale mechanisms supporting extensibility.

An important goal for a pipeline processing framework is to provide a uniform execution environment where a science pipeline application does not need to know where it is being run. If this can be achieved, it will be easier to deploy existing science software.

We feel it is important that this environment be flexible for a variety of approaches to pipeline frameworks. The most important reason for this is that it is likely that different pipeline frameworks and processing models will be best suited to different types of algorithms. A good example is the question of using shared memory versus distributed memory machines: tightly-coupled processing on shared chunks of data benefit from the former, which helps minimize message-passing overhead. In contrast, data parallel problems are more simply mapped to a Beowulf-type cluster; however, they need to be supported by good parallel file systems when the input and output data gets large. Another reason to remain flexible is so that it will be easier to create a hybrid framework that draws on the best components from multiple source frameworks.

We will define a core set of services and practices that allows the LSST system to execute a set of science codes, these include:

- a common way of executing a pipeline application and passing in its inputs
- transparent access to data
- common logging techniques and capture of output streams (e.g. standard out)
- common exit strategy, including
    - capture of exit status
    - declaration of output products

- robust, automatic archiving of pipeline data products
- workflow management: the logical chaining/coordinating of pipeline applications
- process monitoring, error detection and recovery
- transparent support for security.

The LSST is supporting multiple approaches to pipeline processing, most notably a "data-centric" approach as described in section 5.4.2.3 and a "pipeline-centric" approach using traditional tools such as Condor and PBS. We will create an assessment of all relevant application algorithms, how they best map to particular programming models (e.g. data-parallel vs. tightly-coupled message passing), and the corresponding requirements on infrastructure (e.g. shared vs. distributed memory, parallel file systems vs. distributed data-parallel access). We will test moving object detection/orbit determination algorithms against different programming models and their corresponding platform types using parallel computing resources. It will be important that we understand the needs of all of the applications and middleware being developed: not only can we fold the requirements into our pipeline environment design, we can be sure to fill any missing middleware services that might otherwise be out of scope for a particular middleware framework (e.g. security).

Note that while this capability is highly desirable, it is essentially a cost control measure, not a matter of technical feasibility. The degree of flexibility/extensibility will be analyzed and characterized for various architectures, such that cost curves can be estimated based on formal UML specifications of proposed extensions. This will permit the cost-effectiveness of such extensions to be evaluated prior to incurring the full cost of development.

5.4.2.6.3 Tasks

5.4.2.6.3.1 Existing pipeline extensibility analyses

The ability to leverage existing solutions is a major focus of the pipeline environment effort and stands as another motivation for flexibility. In this effort, we plan to study existing pipeline frameworks for applicability to LSST data processing. The SM/Essence Pipeline is quite well-understood and widely noted for it's use of well-verified software modules from IRAF and other modules (e.g. DoPhot) to do much of the standard processing that will be necessary for single-frame processing of LSST images. However, this pipeline has shown that there are modularity issues with IRAF that make it difficult to use in conjunction with other non-IRAF frameworks.

The BIMA Image Pipeline system is interesting to LSST in the way metadata is used to drive processing, its use of python to glue existing software programs, and its techniques for handling error recovery.

The ESO Astro-Wise and Common Pipeline Language (CPL) technologies offer promise for LSST in terms of full object orientation and the consequent benefits to module reuse. With ESO's recent moves into open access, we will evaluate and prototype with these technologies as well.

5.4.2.6.3.2 Grid technologies for LSST pipelines

We will also explore existing grid-based middleware tools that could be leveraged to support LSST pipelines. We will examine the use of existing workflow management systems, including Chimera/Pegasus and OGRE, in light of the LSST processing needs. We will design the environment to address all of the capabilities listed above, and we will enable these capabilities incrementally, drawing on existing tools as available and implementing them otherwise. We will give priority to the capabilities most important to existing pipelines we wish to test and the pipelines being developed by LSST.

We will deploy at least one of the pipeline systems (or perhaps a hybrid) within the Simulated LSST Grid (SLG) described in section 5.4.2.3. This will demonstrate simple, single-frame calibration using existing optical/IR software. This will include flat fielding, bias subtraction, and simple photometric and astrometric calibration using the USNO-B catalog.

By 2006, we also expect to have available to us application modules from other optical surveys, possibly those supporting Pan-STARRS or the Dark Energy Camera (DEC). These surveys, with their high pixel count frames, would be a reasonable match to the processing needed by LSST. By 2007, we expect that much of the prototyping of both pipeline frameworks and applications will mature enough to incorporate into the SLG to demonstrate basic processing of single frame calibration (and more) based on simulated LSST data.

5.4.2.6.3.3 User-defined pipeline modules

Finally, we recognize that supporting user-supplied modules well is a non-trivial effort and producing a full-working prototype of this capability is highly desirable. We will set the foundation for a common pipeline environment, characterized by a set of APIs to services and a set of practices for interacting with the LSST grid, is a necessary foundation for supporting user-supplied code. This user-supplied code, wrapped in a kind of "pipeline shell," could be executed on a subset of existing data in the archive or plugged directly into the automated chain of processing used to create the standard LSST products, thereby applying it to future data.

In addition to the standard pipeline environment, this capability will also require a mechanism for "sandboxing"—in which the rest of the standard data and processing are protected from the user-supplied module—as well as verification—in which the module is verified to be compliant, within limits of its resource use, and otherwise non-malevolent in its behavior. The latter might be simply achieved by benchmarking it against a large canonical subset of the LSST archive. Since the archive will not exist at the time of first light, simulated catalogs will be used for this purpose during R&D.

5.4.2.6.3.4 Development of IRAF-based user-configurable pipeline infrastructure exposed through VO services

The VO framework holds the potential to provide a significant resource to the LSST project, both in facilitating community access to the data and in developing some of the tools that will be needed for analysis, visualization, and data mining, by internal and external scientists. One of the challenges of this very early stage of VO development is taking advantage of legacy software, such as IRAF.

IRAF represents a considerable investment, and more importantly, it is deemed trustworthy by the community. However, IRAF was developed as a complete environment, and as such, it was not originally designed to be partitioned into reusable modules for incorporation into other pipeline frameworks. In fact, IRAF shares this characteristic with most other existing pipelines, and is therefore representative of this class.

This need not be a complete obstacle to using some portions of IRAF in an LSST context, if approached from the standpoint of using modern interface technology to expose interfaces that permit using IRAF from a remote server. Consequently, LSST is supporting the incorporation of IRAF into this type of model for community data access and analysis that is consistent with the vision of the Virtual Observatory.

IRAF integration with the VO on the server side is being accomplished by exposing existing IRAF tasks as web-services to the VO. This is a specific case, but will also be a step to developing the techniques needed to deploy almost *any* compiled or scripted task as a VO service.

While the web-service container being developed is general enough to interface any IRAF procedure (some of which may be quite complex) it is still fundamentally geared towards the

single-processor environment commonly found on the astronomer's desktop. Web-services themselves are atomic in that they expect to immediately return a result and the complexity of a large client application rapidly increases as the number of services being managed also rises. To be truly useful in the distributed processing environment needed for LSST data one wants the science application to make use of services which:

- automatically take advantage of parallel hardware
- allow for asynchronous processing of data and polling for results
- permit users to develop more complex applications by scripting within the web-service environment or easily incorporating new, compiled routines
- limit the transport of data over the network by staging intermediate results
- minimize the impact of XML data inflation in VO-compliant data transport

Enhancing the web-service adapter currently used to manage multiple back-end IRAF sessions can provide a crude form of parallelization, but won't fully integrate IRAF services in a Grid workflow environment. Development of new service functionality is also limited to some extent by the IRAF environment. IRAF-based services currently require tasks to be written using CL scripts and/or SPP compiled tasks, meaning that basic VO tools such as VOTable parsers and URL handlers need to be developed. This inherently limits the role that IRAF and similar pipelines can play in the LSST.

While this approach is unlikely to be the most effective model for internal construction of new pipelines, it represents a fundamental step in the transition to the new paradigm for community interaction with data – which LSST is driving. Providing a migration path for existing IRAF users is an important step in providing broad community access to LSST data.

5.4.2.7 Scalability and fault tolerance in multi-teraflop data management pipelines and multi-Gbps data transmission

5.4.2.7.1 Rationale and overview

The LSST's high data volume necessitates that the pipelines keep up with the generated image output, or risk overwhelming computational and short-term storage resources. As such, it is particularly important that the LSST pipelines execute as reliably as possible. In addition, as scientific algorithms evolve, re-processing the LSST data will occur, probably several times over the course of the survey. This necessitates that the LSST pipelines be highly scalable.

5.4.2.7.2 Activity summary and relationship to other activities

Current development trends in parallel processing middleware technology will be studied with regards to possible long-term impacts on LSST project development efforts. Tradeoff studies will be done to determine the true cost (time and money) in application development and hardware complexity.

Probable failure modes and rates for LSST data and science pipelines will be investigated and mitigation strategies will be devised. Solutions will likely be composed of multiple layers of fault-tolerant software and hardware technologies. Both off-the-shelf industry standard and emerging technologies will be utilized in pipeline prototypes.

We will use a system simulation platform to address a range of questions regarding how data and processing should be distributed across the various LSST sites and to address issues related to fault-tolerance. We will build into the environment methods to simulate various hardware, network, and system software errors that we will use to test our pipeline's ability to recover from them.

This activity will be based on stellar image precursor data provided in related research activities described in sections 5.4.2.1 and 5.4.2.1.

5.4.2.7.3 Tasks

5.4.2.7.3.1 Off-the-shelf middleware for reliable parallel processing in clusters

Many powerful algorithms for parallel-processing image-analysis applications become practical once the time penalty for transmitting large amounts of data across a cluster is negligible.

Contributions to the design of data structures and software interfaces for various LSST pipelines will be based on hands-on experience gained by writing prototype fault-tolerant parallel-processing image-analysis applications based on precursor data sets. Computer-aided software engineering practices for the development and deployment of high-performance parallel-processing applications will be reviewed.

A 64-bit fault-tolerant Beowulf computer cluster will be utilized for the purpose of developing and benchmarking the performance of fault-tolerant parallel-processing image analysis applications on a state-of-the-art cluster. This machine will be used to investigate various middleware approaches for developing robust pipelines and fault-tolerant high-performance-computing applications; special emphasis will be given to the exploration of the usefulness and limitations of state-of-the-art middleware solutions like Open Message Passing Interface (Open MPI), Condor, and OPUS.

LSST partners will be full participants in the Open MPI project (http://www.open-mpi.org). Open MPI is a fault-tolerant implementation of the MPI 2 industry standard that will debut 2005. The current main partners of the Open MPI project are (1) the Advanced Computing Laboratory at the Los Alamos National Laboratory, (2) the Innovative Computing Laboratory at the University of Tennessee, and (3) the Open Systems Laboratory at the University of Indiana.

LSST will attend and fully participate in the Open MPI development efforts by attending the quarterly Open MPI meetings and contributing code. We will also investigate the applicability of Open MPI's component framework architecture for the development of reusable software modules for LSST parallel-processing applications.

5.4.2.7.3.2 Data-centric pipeline tools

The pipeline development tools described in 5.4.2.7.3.1 will allow us to implement prototype pipelines based on industry standard, off-the-shelf middleware, but will not implement many of the ease-of-use and optimization capabilities that affect scalability and are being explored in the research community. We will also examine these "cutting edge" concepts.

We will use GridDB (a project at UC Berkeley) to evaluate the potential of new workflow management tools for developing image analysis pipelines. In this paradigm, data passed between the pipeline algorithms are stored in database tables and as a result clients are able to create, manage, and interactively access the results of the pipeline algorithms using query language interfaces. This enables a flexible and powerful support for ad hoc and real time continuous queries.

In addition, we will utilize this paradigm to make execution of data pipelines on cluster computers transparent to the users. We will work on data partitioning, retrieval, and I/O scalability of such pipelines. We will explore how to support data and algorithm provenance, modular architectures, and real-time queries over streaming images.

We will also define an architecture that will support data memoization and support checkpoints and restart from previous runs. We will explore using the data provenance concept to implement pipeline-aware scheduler that will maximize resource utilization based on collected statistics from

previous runs. These capabilities will be incorporated into operational prototypes as they become available.

5.4.2.7.3.2.1 Pipeline prototype environment at LLNL

We will implement a prototype LSST image processing pipeline on the 1152 node MCR Linux cluster at LLNL. The logical structure of the prototype system is shown in Figure 5.4.2-2. The data generator module will create sets of LSST image files (~200 4Kx4K images) for each camera exposure using either precursor survey data or the LSST simulator. These files are read by the image manager module, stored in the image archive as required, and allocated to the pipeline modules. Allocating processor resources will be handled by the SLURM resource manager (through the ISP API) while storage will be managed by the LUSTRE parallel file system.

The initial pipeline implementations will use image processing and difference image analysis algorithms from the SuperMacho pipeline. As new LSST-specific algorithms are developed, they will be incorporated into the pipeline processes. Standard interface definitions will be developed and documented to make this possible.

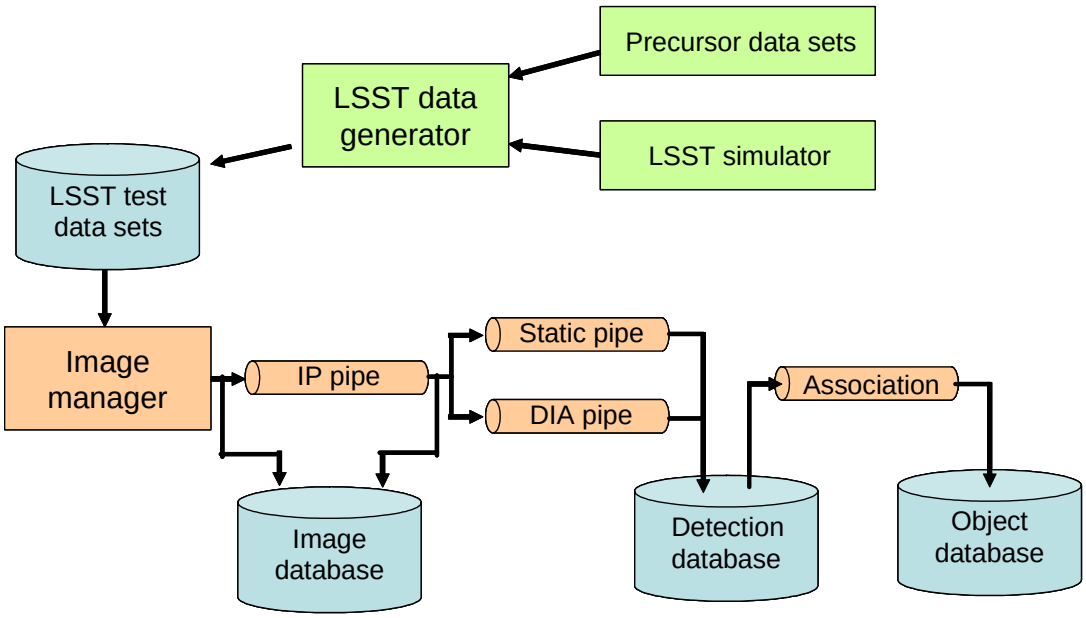


Figure 5.4.2-2 Structure of the LSST pipeline simulator

5.4.2.7.3.2.2 Pipeline prototype environment at BNL

We will also host a prototype data center at Brookhaven National Laboratory based on the RHIC/ATLAS Computer Facility. We will use available resources at that facility to host a modest amount of computing power and storage space that will be used by the collaboration to test software for data access and analysis. We will benchmark the capabilities of this system in order to inform the decision making for the LSST data centers.

The prototype data center at Brookhaven will build on the experience of the RHIC and ATLAS Computer Facility (RCF and ACF). The RCF currently provides data storage and compute facilities to hundreds of scientists at RHIC, and the ACF will do the same for the ATLAS experiment at CERN in the near future. The scale of the facility is presently several thousand CPUs used to analyze several hundred Terabytes of data delivered to the facility over Gigabit Ethernet networks.

We expect to secure several racks of machines which are being retired from the RCF as a testbed platform for LSST code development and benchmarking. Approximately 40 suitable machines will be available by the Elaboration phase of the proposal (2Q05). While these machines are not the present state of the art, the large number of them means we can configure them as platforms that can be used to test how the individual pixel level algorithms perform when running in parallel on precursor data from the test data suite created in related research activities described in sections 4.4.2.1 and 4.4.2.5. These tests will determine how much of the prototype photometric pipeline, and eventually the full prototype environment can be designed to utilize parallel processing capability.

A significant amount of disk space will be required for storage of data from precursor surveys, and caching for making data available for processing. We will examine the tradeoffs between using the inexpensive IDE disks in the servers compared to higher performance disks available to the network. RCF has recently purchased Network Attached Storage (NAS) systems from Panasas 1 and we will monitor their experience with these systems. Other possibilities will be considered as storage technology continues to evolve.

The software environment and network management will be provided by experienced RCF staff. The machines will run Scientific Linux 3.x.x (which is a free build 2 of Red Hat Enterprise Linux 3). For source code management we will utilize a CVS repository accessible to the entire LSST collaboration. We presently envision managing the building of LSST prototype software as RHIC software is managed, with automake tools and nightly rebuilds. The automake tools have made it straightforward to upgrade operating system versions, or change platforms, and a system of nightly rebuilds enforces continuing progress and consistency. We will investigate integrating automake with a system such as eUPS that manages a projects' need to version libraries.

Many databases will be needed for LSST development for purposes ranging from organizing processes and files to telescope environmental data to databases of detected objects. Initially, we will set up PostgreSQL and MySQL database servers which can be used for experimentation, benchmarking, and development. Many groups at the RCF have been using Condor 3 for process management for submitting and monitoring a large number of parallel analysis jobs. A similar Condor setup on the BNL prototype LSST cluster will be used to carry out the initial tests of parallel pipeline processing.

Grid tools that have been developed and used by RHIC and ATLAS groups at the RCF and ACF and are being integrated into the Open Science Grid initiative in place at Brookhaven. These tools will be evaluated for interoperability with the TeraGrid toolset.

5.4.2.7.3.3 Pipeline performance analysis

We have a large set of processors to work with at multiple LSST partners and we can experiment with configurations of the system scaling that emulate different architectural approaches to the top-level system design. For example we can allocate different sets of processors to the mountain-top functions, the base-camp functions, and the archive center functions. By varying the available processors in each set and the communication bandwidth available between each, we can explore variations in the system architecture.

We will evaluate the implementation of a prototype system on the MCR and BlueGene/L systems at LLNL, as well as systems at NCSA. In order to define performance characteristics for image pipelines running on large-scale cluster architectures, we will run test data sets of varying sizes through the prototype pipeline to validate the performance model. We will establish system requirements for meeting real-time alert requirements. Specific issues that we will investigate include

- Data communication bandwidth requirements
- CPU and memory scaling requirements

- Optimal approaches to data partitioning and application parallelization
- Database interface performance
- Fault-tolerance mechanisms

5.4.2.7.3.4 Simulated LSST grid on TeraGrid

We will create the Simulated LSST Grid (SLG) is to provide a distributed testing environment for LSST software. In the simulation grid, each separate platform represents a different data center in the actual LSST data management grid. We currently expect this grid to have the following data centers (described in the LSST Data Management system architecture):

Mountaintop, where the camera control and data acquisition systems operate,
Base Facility, where data ingest and quick processing occurs,
Archive Center, where the bulk of the data processing occurs and where data products are loaded into the archive and made available to the public.
Mirror Sites, where the archive is replicated (in full or in part) and where some specialized processing may occur.

By deploying a simulation grid in a distributed environment, we can test and account for the effects of the network on the data management system. More specifically, we can actively test alternatives for distributing storage and cycles across the grid. We can also look a full-range of hardware and software failure modes that includes the effect of the network.

We will deploy the SLG on the national TeraGrid facility which features nine geographically distributed, high-performance computing sites connected by high-capacity fiber (up to 30 Gb/s between sites). These sites currently contain disk capacities of the order that we expect to need by first light. Thus, the TeraGrid sites are a good match to the expected LSST sites. By building on existing grid-based technologies, the SLG environment can be made general and not specific to TeraGrid, allowing for deployment on other systems.

We will prototype mechanisms that allow developers to add new components to the SLG distribution in a simple but robust way that accounts for software dependencies. Next, we will create a packaging mechanism that allows users to download, unpack, and automatically configure the distribution for a new installation. There are a number of robust packaging utilities in use today that our system can be based on. This includes the Pacman package (http://physics.bu.edu/~youssef/pacman/) which is currently used by a number of grid communities.

We expect that any “official” LSST software would be deployed automatically; however, the packaging mechanism can provide a means for testing existing but not-yet-adopted software (i.e., software in development) within the data management environment. We will then create a simple web portal that allows privileged users to deploy new software. The packaging mechanism will be key to extending our comprehensive testing beyond a contained environment like the TeraGrid. It will allow partners to bring the LSST software environment to their own computing platforms. Not only will they be able to test their modules integrated with the rest of the LSST environment, the deployment provides a mechanism for bringing local computing facilities used by our partners for development into the LSST grid. Three such important platforms include the Beowulf cluster being used to develop OpenMPI applications as well as the clustered prototyping testbeds that will be used at LLNL and BNL.

5.4.2.8 Technology trend analysis, projection, and validation strategy

5.4.2.8.1 Rationale and overview

The design of the LSST data management system architecture is influenced by the computing, storage, and communications technology we expect to be available to implement it, starting with construction in 2007 – 2008 and continuing through the principal survey period until 2022.

Past experience has shown that premature "locking in" of these technologies invariably leads to increased life cycle cost and even premature obsolescence. It is our intent to select these technologies at the last possible point in time that still permits achieving the planned construction schedule. We anticipate this will be in the late 2007 time frame for the initial operational configuration.

Advances in the cost/performance ratio of these computing, storage, and network technologies accrue from multiple sources:

- Raw performance at the component level due to improved design and manufacturing processes and technology, e.g. semiconductor CPU feature density and clock speed improvements vs. cost (i.e. Moore's Law)
- Net performance at the subsystem/system level due to advanced parallel architectures, e.g. cost of teraFLOP/s improvements in cell and multi-core processor architectures vs. conventional single CPU and cluster architectures
- Advances in software algorithms and data structures that more efficiently utilize hardware resources, e.g. IPv6 communications protocol and the cell object

The infrastructure section of this proposal contains diagrams that depict the anticipated trends in computing, storage, and networking as analyzed by the LHC PASTA group between 1996 and 2002 (reference). This group considered primarily the first two elements above, While useful in itself in bounding the cost to a first order of magnitude, it falls short of the full effect that must be considered in light of the remaining element, and of the impressive advances that have occurred since 2002.

In particular, while there is some evidence that the speed of advances in CPU technology is slowing (reference) as individual micro-processors achieve performance close to the theoretical limit of semiconductor physics, there is counter-evidence that the latter two elements are continuing or even accelerating in ways that more than compensate.

For example, the BlueGene/L supercomputer (joint IBM/DOE/NNSA effort) is an example of the second element above. To achieve ultra-performance computing BlueGene/L takes a radically different approach from traditional supercomputers by not relying on increased power in individual CPU nodes. Utilizing a cell-based design methodology, BlueGene/L is a scalable architecture in which the computational power of the machine can be expanded by adding more building blocks, with no introduction of bottlenecks as the machine scales up. By utilizing system-on-a-chip (SOC) design technology and low-cost/low-power embedded microprocessors, BlueGene/L achieves a theoretical peak computational rate of 367 teraFLOP/s through extreme scalability. With more than 2^{16} (65,536) dual processor nodes and 16 TB of memory (16 x 2^{40} bytes or 256 MB DDR SDRAM per node) this will be the fastest machine yet built as of 2005. BlueGene/L has three main communications networks: a three-dimensional torus for nearest-neighbor calculations on grids; a global tree network for broadcasts and reduction operations; and a barrier network for synchronizing the complex algorithms in scientific calculations envisioned for the machine. (reference)

Similar advances are occurring in the evolution of the "Cell Processor" a joint venture between IBM, Sony, and Toshiba. (Expand and reference).

5.4.2.8.2 Activity summary and relationship to other activities

In this research activity, we will track and analyze the current characteristics, and then predict the expected characteristics, of computing, network, and storage hardware over the 2007 – 2022 time frame. We will use existing sources (e.g. the PASTA report) as well as draw on unparalleled partner expertise in major computer, storage, and communications system specification, procurement, and support, to survey and analyze cost vs. performance trends in applicable technologies. This analysis will then be used to establish the critical tradeoff parameters needed to optimize the total life cycle cost, availability, and reliability of the LSST data and processing resources.

These parameters will be used to create system architecture models and simulations that will explore the effects of trading off between them, i.e. transmitting and storing derived data products versus adding computational resources needed to recreate the derived data products in multiple locations. We will explore emerging system architecture and technology alternatives to the computer, storage, and network systems. We will then map the alternatives onto the technology trend parameters to define lifecycle cost, availability, and reliability models.

As described in section 5.4.2.3 in a related research activity, we will also analyze how different application software and middleware technologies exploit the hardware infrastructure to improve scalability and reliability.

As described in section 5.4.2.6 in a related research activity, we will also identify how to most effectively leverage layered data system architectures in order to insulate the design as much as possible from underlying platform dependencies. With this insulation, as the technology evolves, we will facilitate re-hosting on new platforms with minimal effort and cost.

5.4.2.8.3 Tasks

5.4.2.8.3.1 Top-level architecture definition and performance requirements

Develop broad functional requirements for each data management subsystem. Define the top-level architectural approach to data management. Establish baseline performance requirements including computational performance, storage, and network bandwidth.

5.4.2.8.3.2 Technology analysis

Analyze technological trends in all of the dimensions described in the overview section. Create year by year projections for all years from 2007 – 2022.

5.4.2.8.3.3 Computer system conceptual designs

Develop a set of system point designs based on projections of technology. Define computational node design, interconnect architecture, storage system design, and parallel processing environment. Define network interfaces and bandwidth requirements. Since these components interface with the camera, telescope, observatory control system (OCS), and the down stream data flow, the designs will be coordinated with the camera, telescope, and OCS teams.

5.4.2.8.3.4 Lifecycle cost, reliability, and availability model

Create a tradeoff framework/model for predicting life cycle costs, reliability, and availability that is parameterized along each technology dimension and permits inputting of multiple system designs. The model will cover system design and procurement, facilities, operations, maintenance, and upgrades. It will discuss tradeoffs on different approaches to procurement and support, for example, full-service vendor vs. self-integration approaches. The model will be developed such that it can be upgraded as the system design evolves.

6 Project Organization and Management

6.1 Introduction

The LSST Corporation (LSSTC) is a non-profit 501(c)3 Arizona corporation with headquarters in Tucson, Arizona formed solely to design, construct, and operate the LSST. LSSTC membership currently includes fourteen US institutions: Brookhaven National Laboratory, Harvard-Smithsonian Center for Astrophysics, Johns Hopkins University, Las Cumbres Observatory, Inc., Lawrence Livermore National Laboratory, National Optical Astronomy Observatory, Research Corporation, Stanford Linear Accelerator Center, Stanford University, The Pennsylvania State University, University of Arizona, University of California at Davis, University of Illinois at Urbana-Champaign, and the University of Washington.

The US Department of Energy will be a major partner along with the National Science Foundation in the construction and operation of the LSST. DOE participation is based on their fundamental interest in LSST probes of dark matter and dark energy. The DOE laboratories have assumed responsibility for design and construction of the LSST camera system and focal plane data acquisition system; this represents about a third of the project hardware costs. SLAC will take the lead and act as the interface with the DOE Office of Science. While the plan is for DOE to fund the camera project, there will be participation from non-DOE organizations. Beyond the camera, the involvement of the DOE laboratories brings to the LSST project fundamental enabling technical capabilities, honed from extensive experience with numerous previous and ongoing large experiments.

6.2 LSST Project Management

6.2.1 Management Structure

Management of the LSST project is based on proven project management practices. The guiding principles of the management plan include:

- The LSSTC Board of Directors, led by the LSSTC President, sets policies, approves project organization and selected high-level technical requirements (such as the LSST Science Requirements Document), and has primary fiduciary responsibility for the project.
- An LSST Director and Project Manager, each reporting directly to the LSSTC Board. Together they supervise all the scientific and engineering teams
- A Change Control Board (CCB) and a Science Advisory Committee (SAC) with well-defined roles and responsibilities
- A management structure and tracking system based on the Work Breakdown Structure (WBS)
- A formal Risk Management process to characterize budget, technical, and/or schedule risks, to assign risk numbers to each WBS element, and to track changes in risk as progress is made
- Rigorous, formal program reports and reviews, including whatever reviews and reports are required by the NSF and DOE

Figure 6.2.1-1 below shows the overall management structure. As described above, the Board is the primary governing body. The Director and Project Manager work in collaboration to manage the project. The Science Advisory Committee and Change Control Board maintain oversight and endorse major changes in technical scope and direction of the project. Both the Director and Project Manager are members of both committees. Disagreements, if any, between the Director, Project Manager, and/or either the Advisory or Change Control Boards will be resolved by the Board of Directors.

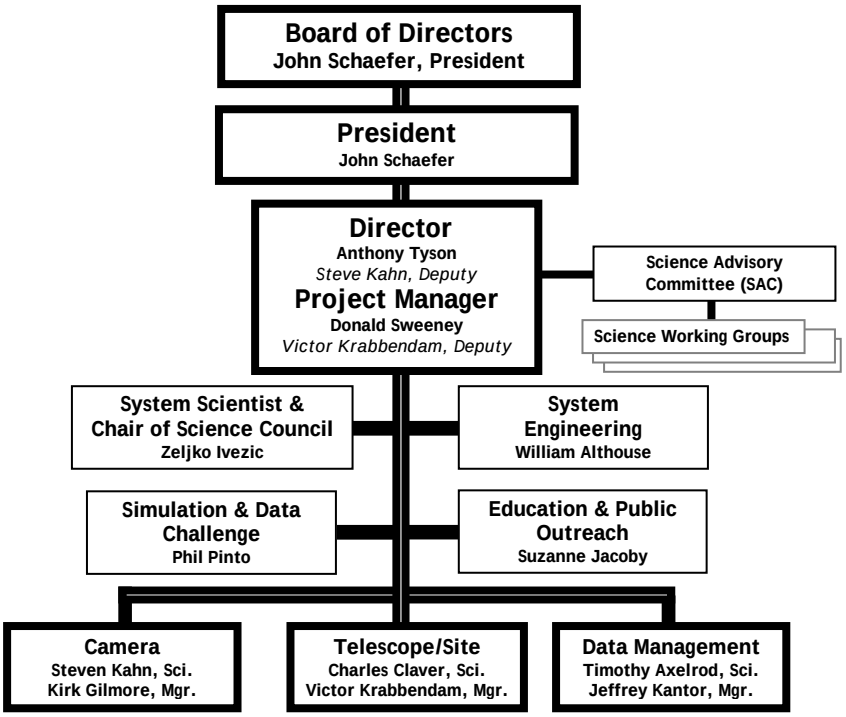


Figure 6.2.1-1 LSST Management Structure

As shown in Figure 6.2.1-1, three subsystem teams labeled Data Management, Camera, and Telescope/Site will execute the primary WBS tasks. The composition of the subsystem teams is intended to unite the scientific and engineering interests of the program. Each subsystem team is led by a scientist and a project manager. Each subsystem project management team will:

- Have responsibility for performance and delivery of specific WBS components
- Manage appropriate staff, budget, and deliverables
- Develop and own the Risk Management Score Card for their respective WBS tasks
- Represent their subsystem to the Project Manager and to the Change Control Board.

6.2.2 Integrated Project Management Control System (PMCS)

The LSST Project will implement a Project Management Control System (PMCS) to monitor and assure compliance with cost and schedule baselines. The PMCS will be implemented by a project controls specialist who is responsible for schedule, cost, and financial performance monitoring, variance analysis, and monthly project status reporting.

The PMCS will:

- Establish and maintain an integrated cost and schedule baseline
- Provide for the orderly and systematic authorization of work and project budget
- Develop and publish timely management reports which display cost, funding and schedule status with respect to baseline plans
- Measure actual and forecasted cost and schedule status against the performance measurement baseline to determine current performance and forecast future performance
- Maintain a clearly documented audit trail of all changes to the performance measurement baseline through the Work Breakdown Structure (WBS)
- Identify potential problem areas in sufficient time to implement the proper management actions.

The WBS ensures that project management control flows down from the Project Manager to all subsystem managers. A Control Account Manager for each subsystem will be under the direct authority of the Project Manager and is required to report monthly to the Project Manager on cost, schedule, and performance measurement.

The PMCS system formally maintains the project's cost and schedule baselines, providing timely performance measurement data and reports. These data, and the corresponding reports, provide the Project Manager and subsystem managers with the necessary visibility to analyze progress and identify any significant problems and issues in order to establish and implement corrective action.

6.2.3 Configuration Management and Change Control

The LSST Configuration Management and Change Control process to be implemented during the development phase of the project is defined in a formal project policy approved by the LSST Board of Directors on February 28, 2005. This section recaps that policy.

Configuration change is desirable and inevitable during LSST development phase. However, even during this early stage, the baseline must be documented, fixed, and respected. A deliberate and controlled process for adopting change must also be implemented to coordinate internal components of the project, involve interested technical members of the project, and present LSST externally as a focused, well-managed project.

Proposals for new alternatives or configurations for the LSST are strongly encouraged. Alternative or prior configurations can be called early point designs or proposed designs but they must not be confused with the official project specifications or configuration. Official configuration or specification changes can only be implemented following the process below.

- Mature documents will be selected and placed under configuration management by the Project Manager after consultation with the Director and System Scientist; documents will be selected that affect the overall scope, costing, performance, schedule or interfaces of the project; a systems approach will be used
- Proposed changes to controlled documents must be approved by the Change Control Board; the Board members include the Project Director, the Project Manager, the System Scientist, the System Engineer, and the Project Managers and Scientists for Telescope, Camera, and Data Management; the Project Manager will facilitate and chair the meetings
- Change Control Board deliberations shall include consideration of the technical, budgetary, and schedule impact on the project
- A change is approved by consensus of the Change Control
- Proposed changes can be submitted to the Change Control Board by any project group or individual member

- The Change Control Board shall consider proposed changes in a timely manner
- The Project Manager shall ensure that all changes are documented and communicated as appropriate.

6.2.4 Contingency Management

One means of mitigating risk is the allocation of contingency (reserves). Contingency consist of any unallocated resource available to the Project, including the usual budget reserves but also schedule float and technical margins. Contingency application is one element in the hierarchy of risk management tools. The first level is to implement corrective action to maintain baselines. The second level invokes the allocation of technical margins (reserves). If that is an insufficient or inappropriate solution, then cost and/or schedule reserves may be applied. Finally, de-scoping may be used as a last resort, and if used, will be coordinated with the LSST Board and the federal Program Officer.

Once baseline configuration items are established, the Change Control Board (CCB), chaired by the Project Manager, will consider and recommend disposition of requests for changes to system-level designs and interfaces, as well as proposed drawdowns on project cost, schedule and/or technical reserves. Budget, schedule and technical reserves are under control of the Project Manager. Subsystem internal re-allocations may be made within the defined resources available to the subsystem, documented through the established CCB processes and approved by the subsystem scientist and project manager without further consideration by the CCB. Changes that affect anything outside the subsystems (Project-level requirements, Project deliverables, other LSST subsystems) must be documented, and a request submitted to and approved by the CCB.

6.3 Camera Project Management

6.3.1 General Camera Management Flowdown

The LSST overall management process will require authorization by the Project Manager for major procurement expenditures. This flows down through the relevant subsystems, e.g. LSST -> Camera -> Sensors, accordingly, with managerial approval at each stage. The LSST Camera organization chart shown in Figure 6.3.3-1 has been constructed along the lines of a WBS model, where the individual subsystem managers are responsible to the Camera Manager for the development and delivery of the relevant components, who is then responsible to the LSST Project Manager for the development and delivery of the camera as a whole. As the project develops from the R&D to the construction phase, there will be many cost, schedule, and performance issues that will have to resolved by a coherent process. It is essential that the entire Camera Management is involved in the process to resolve these issues in each case, since the recommended actions may have implications for other subsystems within the camera.

Here Camera Management refers to a Camera Management Council (CMC), which is chaired by the Camera Manager and includes all of the subsystem managers as well as the Camera Scientist. For a given proposed procurement, this group will review the cost and schedule implications of the details of the plan put forth by the relevant subsystem manager. If it is consistent with the existing budget and schedule as it exists at that time, it would be approved without issue. If, however, the proposed procurement represents a variance against plan, then the groups as a whole will evaluate its implications for the other subsystems. Most if not all of these issues can be settled by consensus within this Management Council.The Management Council does not in general get into technical debates about the merits of the proposed approach. Technical responsibility has been delegated to the individual line managers, along with the responsibility for implementing the program once the expenditure authority has been approved.

The purpose of this "camera level" coordination is mainly to ensure compatibility with the other subsystem developments. The CMC will discuss implications if, in fact, the studies cost more than originally anticipated, or if the schedule gets extended for various reasons, or if it is clear that the studies will not accomplish what was originally intended. That is what we meant by "camera team approval".

The Camera Management Council meets periodically to assess and resolve issues as they arise. The Management Council will be the primary body responsible for developing the schedule and budget for the camera both during the R&D phase and during construction.

The camera managementt organization also has other functions including a Camera Steering Committee. This group is intended to be the proper forum for discussion of political and policy issues, e.g. controversies between institutions, the interaction with DOE, etc. It is not meant to be a technical or financial decision making body.

6.3.2 Institutional Camera Management Flowdown

As shown in the organizational chart, Figure 6.3.3-1, each institution is responsible for development of a camera sub-system.

- SLAC: overall camera project management; camera mechanical design; focal plane assembly; camera integration and test; front-end DAQ; supporting science activities, modeling and analysis.
- BNL: sensor and FEE development; integration of sensors with FEE and BEE; support for focal plane assembly and test, and camera integration and test; Sensor metrology; collaboration in the front-end DAQ; modeling and analysis.
- LLNL: mechanical and optical engineering; participation in the assembly and test of the optical elements and filters; support for camera integration and test.
- Harvard: electronics engineering; support for BEE/sensor integration and test.
- UC Santa Cruz: camera control software; support for camera integration and test; supporting science activities, modeling and analysis

6.3.3 Work Breakdown Structure

The detailed Camera Work Breakdown Structure (WBS) is also part of the camera management structure illustrated in Figure 6.3.3-1. The top-level of the WBS is divided into eight parts and is directly related to the management structure as shown. The technical challenges and approaches associated with each WBS element are discussed elsewhere in this proposal. The camera cost and schedule is based on the WBS and includes all activities for the R&D phase regardless of the source of funds. Camera Project expenses will be tracked according to cost accounts based on the WBS.

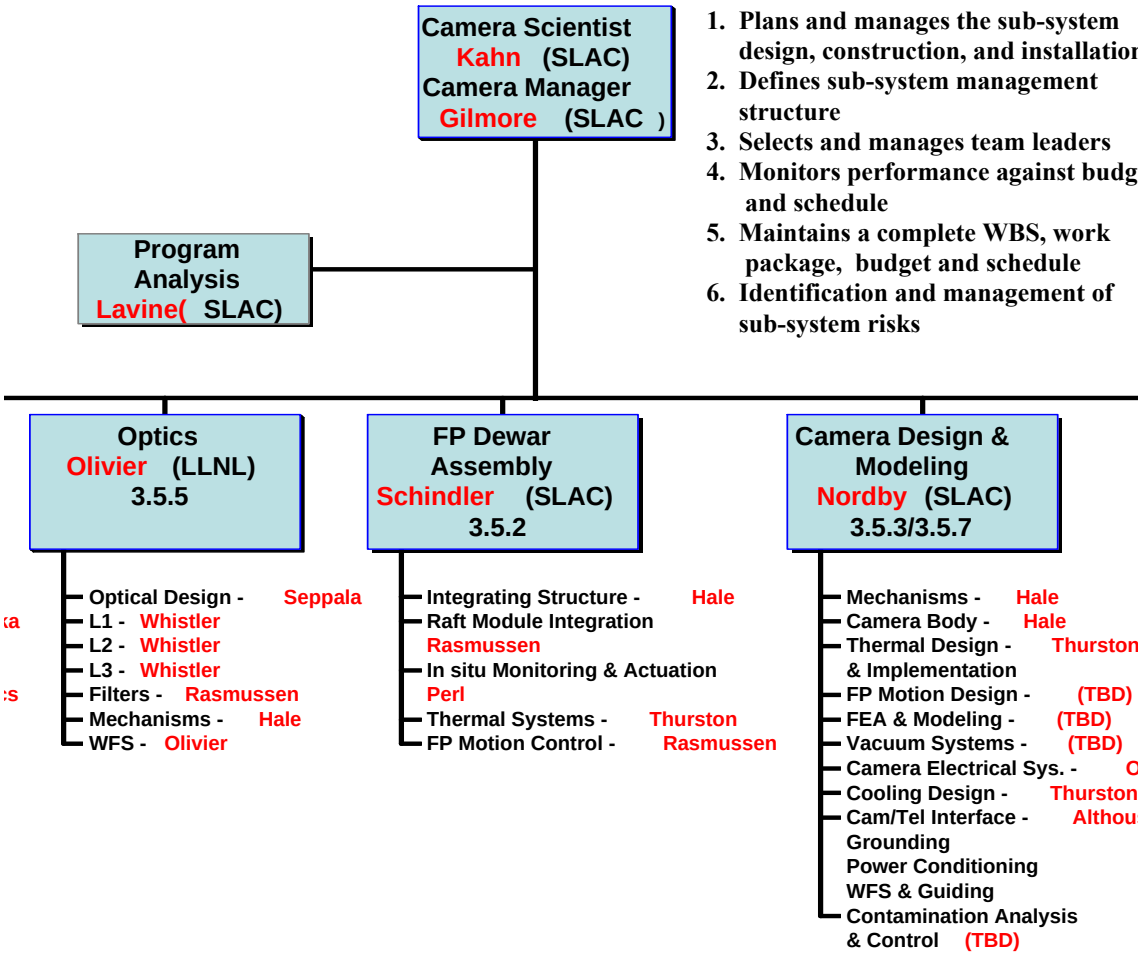


Figure 6.3.3-1: Camera Organization Chart

7 Cost and Schedule

7.1.1 Cost and Schedule Analysis

This section addresses the budget plan for preconstruction R&D and project engineering during the next two years, and the long-range planning process for developing the Performance Baseline for Earned Value Management during project execution.

A. Budget Plan for Preconstruction R&D and Project Engineering

The R&D budget plan for the LSST camera system includes manpower at SLAC and several other institutions funded by DOE (BNL, LLNL, Harvard, UIUC, etc.). The budget for materials and services (M&S) will be managed and disbursed by SLAC through financial plan transfers and procurements. Planned resources and budget authority for preconstruction R&D and Project Engineering in FY06 and FY07 are summarized in the Table 7.1.1-1. DOE construction funding is anticipated to start in FY08. Budget authority for an Engineering Test Camera is identified in FY07; this is an important consideration for risk reduction that will require financial resources prior to the start of construction of the LSST camera.

Table 7.1.1-1 Resources Required for Preconstruction R&D and Engineering

	FY06	**FY07**
Camera Collaboration Manpower (Total)	$5.5 M 30 FTE	$7.6 M 42 FTE
SLAC Manpower	$2.0 M (12.7 FTE)	$2.8 M (18 FTE)
SLAC M & S	$1.5 M	$5 M
Engineering Test Camera (Increment)	–	$3 M

B. Long Range Planning Process

Long range planning for the LSST camera project so far has included development of a conceptual model for the scope, schedule, cost and resources required to construct the camera system and integrate it with the LSST telescope and data management systems. That conceptual model currently is serving as the basis for defining the work scope, developing a project schedule (and acquisition strategy), setting subsystem budget targets, and prioritizing the camera R&D effort.

The conceptual planning model was developed in a top-down approach, grounded by the science requirements and technical requirements of the LSST program, and by a notional time line consistent with key program milestones such as the start of construction funding in FY08 and "first light" in 2012. This model includes significant resource contributions from the multiple DOE institutions participating in the program. As a "top-down" model, it involved order-of-

magnitude estimation of scope, cost, and resource requirements. The estimates, made in 2004–2005, were based on specific analogy, parametric scaling, level of effort, and expert opinion, not on project engineering design.[2]

A more detailed and more accurate scope-schedule-cost model currently is being developed and refined incrementally. In mid-2006 it will become the baseline for the Preliminary Project Execution Plan.[3] When validated and approved by DOE (CD-2) it will become the time-phased Performance Baseline in the Earned Value Management System.

The more detailed scope-schedule-cost baseline for the Project Execution Plan is grounded in the maturing reference design concept for the LSST and its camera system. That baseline is being developed in "bottom-up" fashion using detailed, activity-based cost estimates and unit-cost data where available.[4] The baseline is being captured in a project management information system that is based on a product-oriented construction WBS, augmented by a detailed activity schedule. Integration of the costs with schedule is being accomplished by assigning the estimated costs and resources to the scheduled activities so that the time phased budget "rolls up" through the tree of WBS elements.[5]

The same management system is being implemented simultaneously for each major system of the LSST project --- the Camera System, the Telescope System, and the Data Management system --- with central coordination by the LSST Project Management Office. The project management baseline for the camera system will be integrated with the baseline for the entire LSST project. Iterative review and refinement of the project execution baseline for all LSST systems are proceeding in parallel. Three iterations of the baseline refinement cycle are planned to conclude in April, June, and Fall 2006.

The time line for implementation, review and refinement of the Preliminary Performance Baseline is being driven by several critical events planned or anticipated in 2006:

- The SLAC Director's Review of the Camera Project in March.
- The Particle Physics Program Prioritization Panel (P5) review of the Camera Project in April.
- The LSST Project Manager's Review of the LSST Project Execution Plan in June.
- Submission of the NSF construction proposal for the LSST Telescope and Data Management Systems in November.
- Assessment of the Camera Project by DOE in the Fall.

7.1.2 Camera Development Schedule

The basic project timeline is shown in Figure 7.1.2-1 and reflects the LSST project effort for the three major subsystems, camera, telescope and data management. Figure 7.1.2-2 shows the detailed effort for the development of the camera.

---

[2] Office of Management Budget and Evaluation, U.S. Department of Energy, "DOE Cost Estimating Guide for Program and Project Management" (DOE G 430.1-1X, April 2004), Section 2.4–2.5. Available on-line: http://oecm.energy.gov/Portals/2/g4301-1x.pdf.

[3] Office of Management Budget and Evaluation, U.S. Department of Energy, "Project Management for the Acquisition of Capital Assets" (DOE M 413.3-1, April 2003), Section 5.6. Available on-line: http://www.directives.doe.gov/pdfs/doe/doetext/neword/413/m4133-1.pdf.

[4] "DOE Cost Estimating Guide for Program and Project Management," op. cit.

[5] The scheduling and cost estimating database architecture is being implemented using project management software from Primavera Systems (Bala Cynwood, Pennsylvania) and the "ProPricer" cost database by Executive Business Services (Temecula, California).

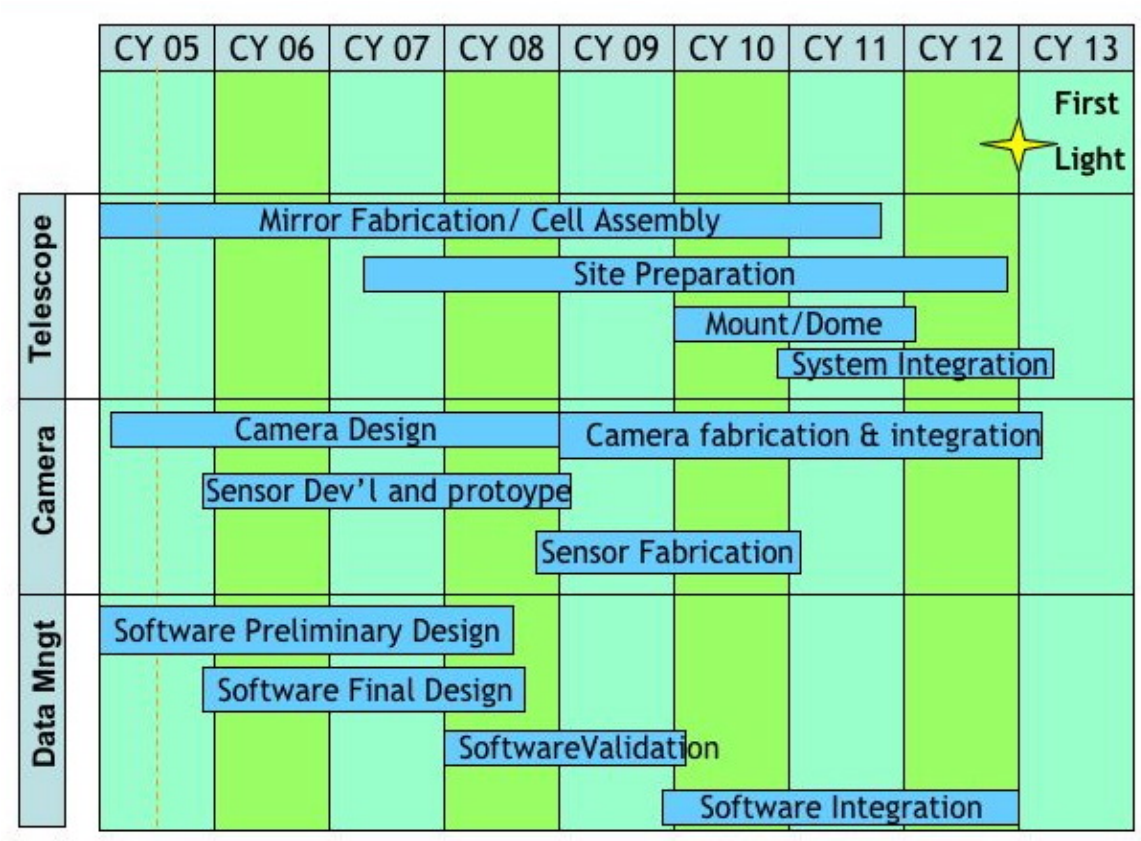


Figure 7.1.2-1 LSST Project Timeline

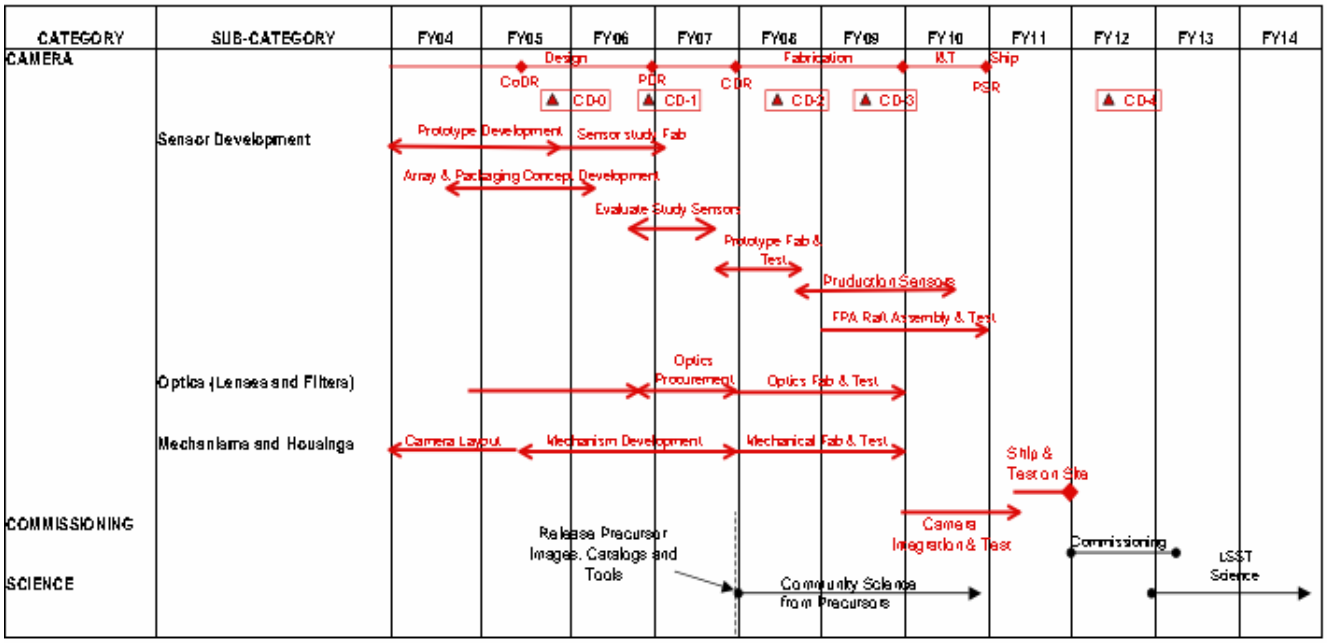


Figure 7.1.2-2: LSST Camera Development Schedule