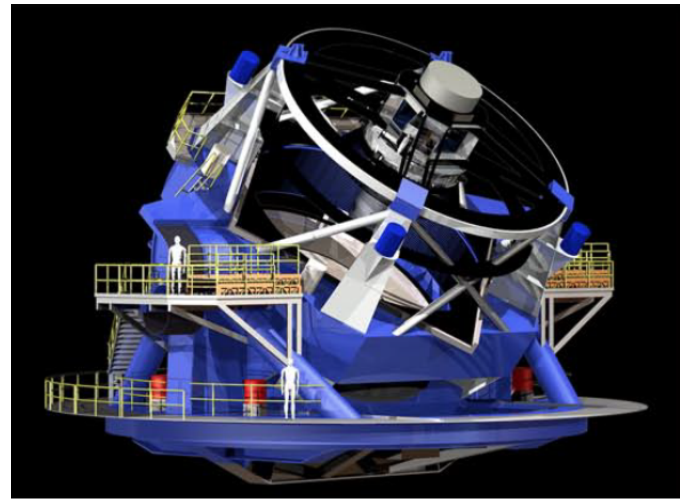
Present and Future Computing Requirements: Large Synoptic Survey Telescope

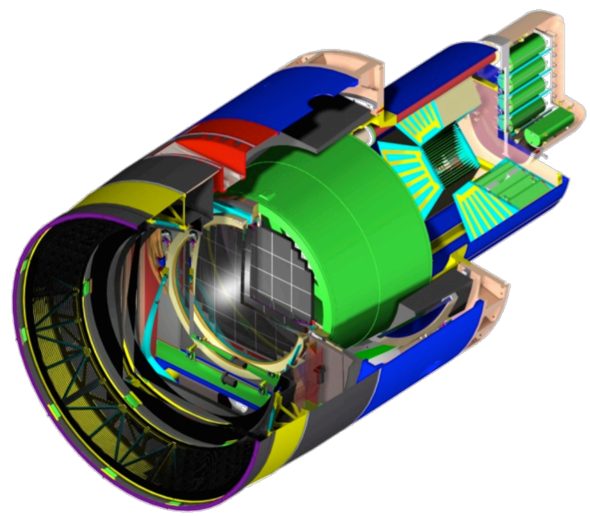
Andrew Connolly
University of Washington

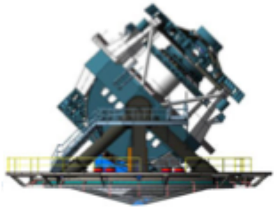
NERSC BER Requirements for 2017
September 11-12, 2012
Rockville, MD

Large Synoptic Survey Telescope

- Wide and deep
  - 27000 sq deg (wide)
  - 100 - 200 sq deg (deep)
  - 20,000 sq degrees every three nights
  - 9.6 sq degree FOV
  - 3.2 Gigapixels camera
- Broad range of science
  - Dark energy
  - Dark matter
  - Time domain universe



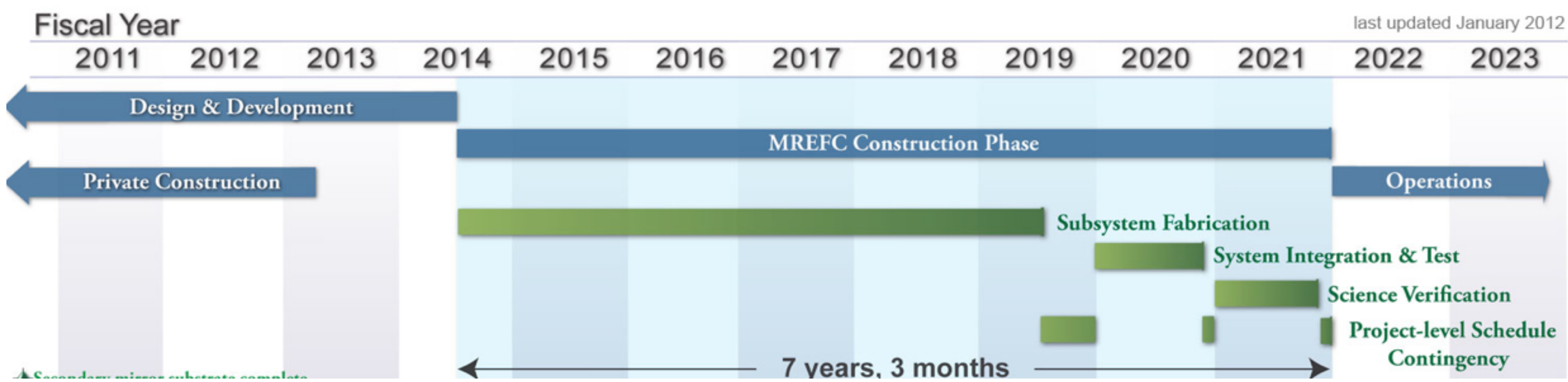




Data flow from the LSST

- One “Visit” is a pair of back-to-back exposures
  - 2x15 sec exposure; duration = 34 seconds with readout
- Data volume associated with this cadence
  - one 6-gigabyte image every 17 seconds
  - 15 terabytes of raw scientific image data / night
  - 100-petabyte final image data archive
  - 20-petabyte final database catalog
  - 2 million real time events per night every night for 10 years

LSST timeline



- Estimate: commissioning in 2020
  - Primary/Tertiary mirror being polished
  - Secondary mirror blank in hand
  - Sensor program delivered first prototype sensors
  - Processing pipelines under construction, hand-in-hand with simulations of Operations, Images, Catalogs

Capabilities limited by systematics

- 1% absolute photometry, 0.5% relative photometry
  - A factor of two improvement on current ground-based wide-field surveys such as the SDSS and 2MASS
- Astrometry uncertainties of 0.01 arcsec
  - A factor of 3 improvement over constraints in current surveys such as the SDSS and 2MASS
- Correlated shape residuals of $<10^{-5}$ in the PSF
  - Requires 3 orders of magnitude improvement in PSF interpolation (as a function of wavelength and position on the focal plane), and 2 orders of magnitude improvement in our ability to calibrate shear.

Priorities for 2017

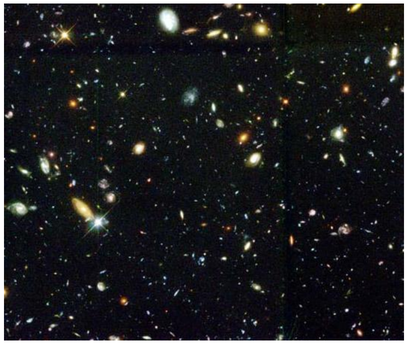
- Data Processing (LSST project)
  - Data analysis at 10% of the operational data rate
- Science Algorithms (Dark Energy Science Collaboration)
  - Weak lensing: PSF and shape characterization
  - LSS: Systematics from stellar density and sky brightness
  - Supernovae: Photo-z and classification
  - Clusters: absolute mass calibration
  - Strong lensing: external mass characterization
  - Cross-cutting: deblending of galaxies

Over 50% of DESC high priority tasks are related to systematics and require simulations (astro-ph/1211.0310)

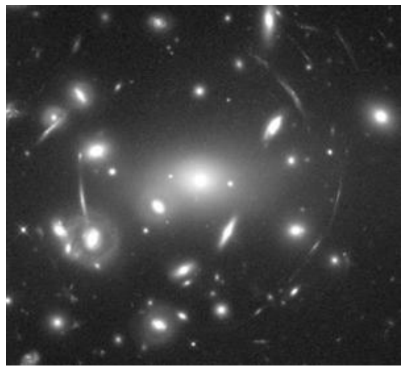
Computational Tools

- Cosmology simulations
  - See talk by Salman Habib
- Image simulations
  - Photon-based simulation of the LSST system
- Data analysis pipelines
  - Including the science analysis pipelines (level 3)

Image Simulations: The Life of a Photon

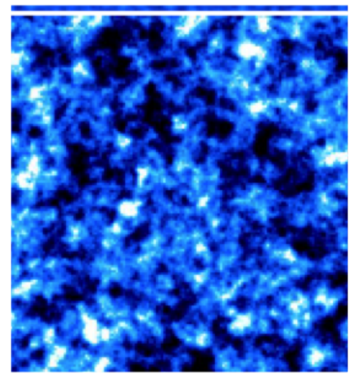


Light emitted by virtual universe.

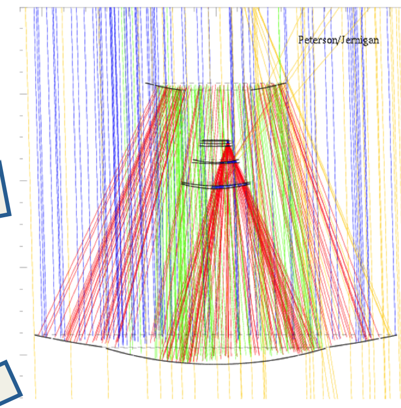


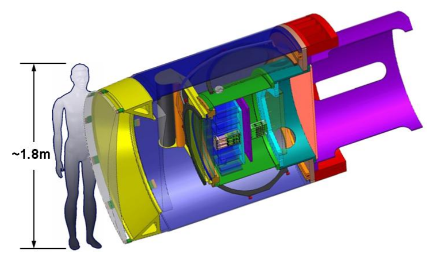
Distortions due to gravitational lensing.

Distortions due to atmosphere



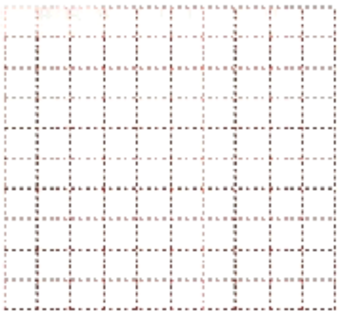
Refraction and scattering due to telescope.

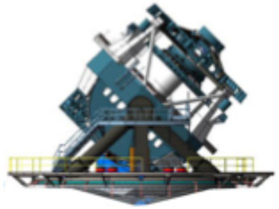




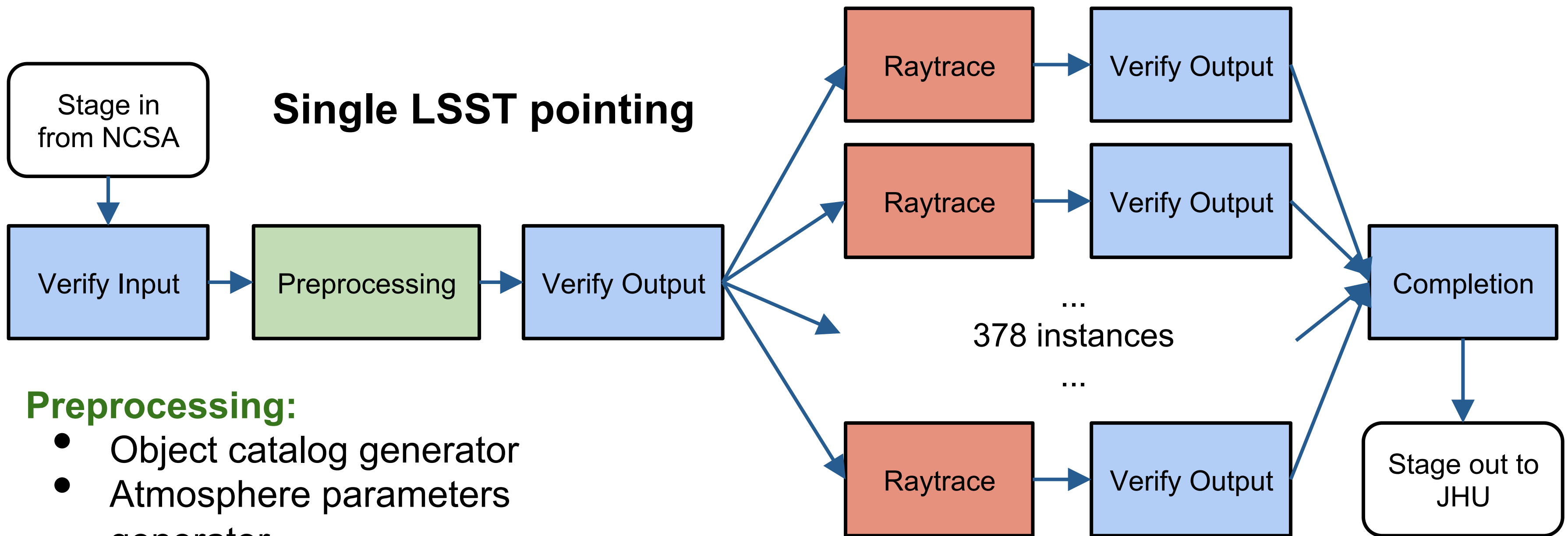
Detector effects (photon-to-electron conversion, saturation, readout noise, electron-to-data conversion, etc.)

Flow of photons





ImSim Computational Flow



Preprocessing:

- Object catalog generator
- Atmosphere parameters generator
- Atmosphere screens generator
- Cloud screens generator
- Telescope optics parameters

Raytracing:

- Photon raytracer
- Cosmic ray adder
- Background radiation adder
- CCD readout modeler

- Each CCD and exposure can be treated independently
- Therefore, we have *378-way parallelism per pointing.*

Computational Strategies

- Simulation codes use fast ray-trace algorithms
- PhoSim assumes geometric optics and simulates ~10^{10} photons per CCD (5×10^5 photons/s). Operations include FFTs and fast intercept calculations
- No message passing between nodes (currently)
- Computational challenge is the parallel startup IO: 200MB of data (~2000 files) required per task and the requirement that all tasks complete.
- Evolution of code is expected to include multi-threaded approach, possibly GPUs.

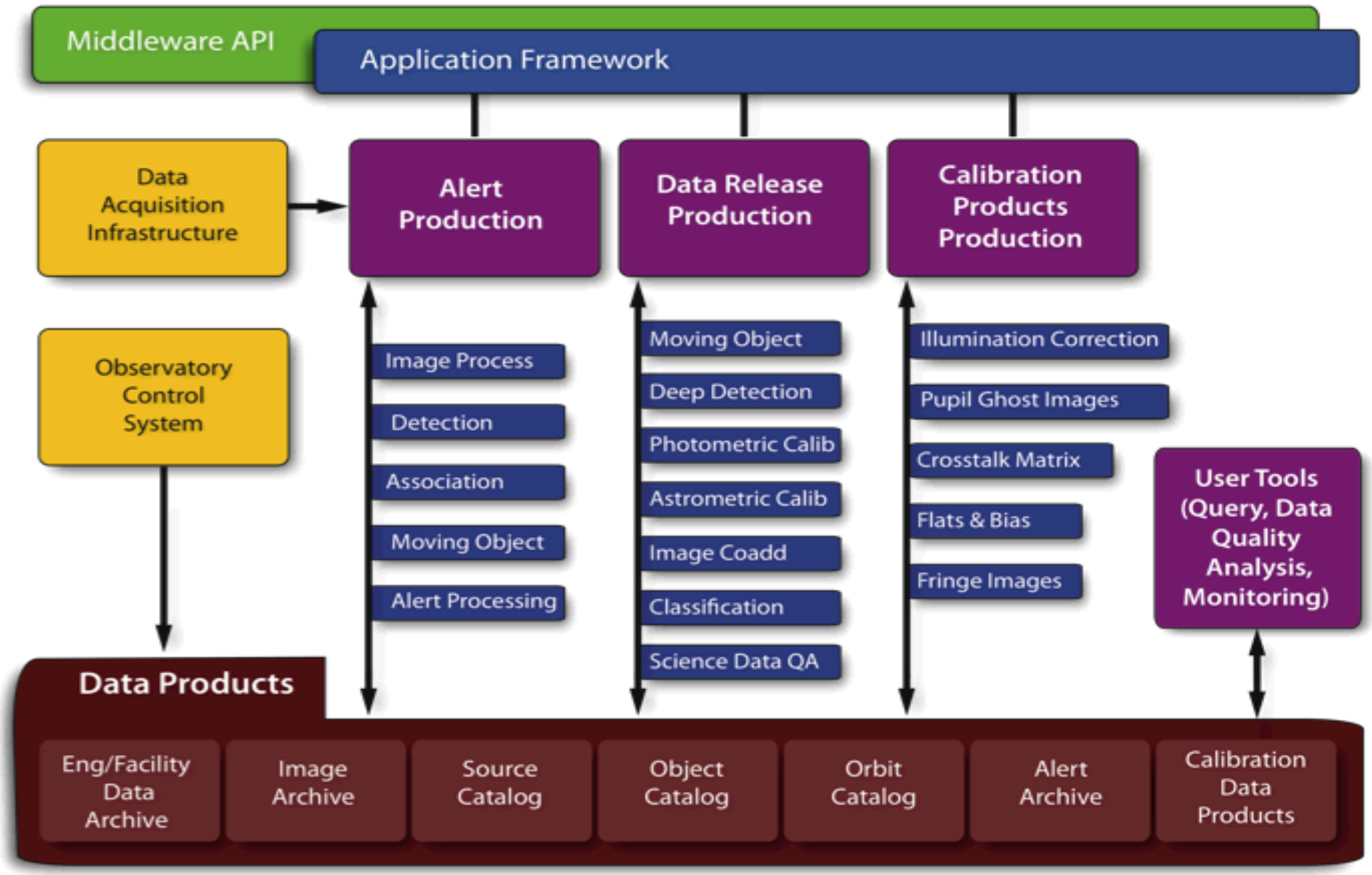
Current HPC usage

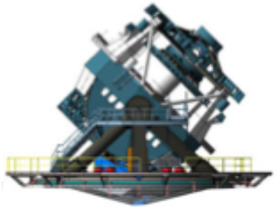
- Image Simulations
  – Babar, Condor, Open Science Grid, Google Exacycle
  – Currently run at 2000 cores, but have scaled to 50,000 cores using Google Exacycle resources
  – 2012: 8M CPU hrs (4M Google, 4M Babar (SLAC), Condor, Open Science Grid)
  – 2013: NERSC request for 4M CPU hrs (Carver)
  – Concurrency and run times: 2 larges scale (4M CPU hr) runs and at 4K cores (for ~30 days)
  – Run in scavenger mode

Current HPC usage

- Read 200TB (large data runs; 1.6TB unique data)
- Write 12TB (for large data runs; 6.4GB per simulated focal plane)
- Runs at 1GB per core
- Store ~20TB of data

Data Analysis Pipelines





Data Analysis Pipelines

- Unit of operation is a single CCD (as with generation)
- Pipeline framework: a script that sequences various processing algorithms
- Orchestration layer: an event system built around the Java Event Framework
- Pipelines run in parallel (Condor's “glide-in”): data is fed in as data-parallel chunks. Complex workflows achieved by combining multiple pipelines together.

Data Analysis Pipelines

- Core algorithms written in C++. Framework written in Python (swig binds to C++)
- For real-time performance or when parallel threads need to exchange data we use an MPI harness (mpich2). Otherwise pleasingly parallel.

Computational Strategies

- Pipeline codes primarily perform pixel operations.
- Use standard libraries: FFTW, numpy, Boost, Mpich2, MySQL databases, Eigen, cfitsio, astrometry.net, Minuit2.
- 1TB memory, 250 nodes, 4GB per core
- Current scaling challenge is the middleware scaling to >1000 cores due to IO, memory requirements and fault tolerance.
- Evaluating GPUs for compute intensive operations (e.g. image warping).

Current HPC usage

- Data Analysis Pipelines
  - NSF XSEDE resources: Abe, Lonestar, Gordon, and Trestles
  - 2012: 500,000 CPU hrs (primarily Gordon)
  - Scaled to 1000 cores
  - Concurrency and run times: 2 larges scale runs (500,000 CPU hr)
  - Read 50TB (loaded 1.5 billion rows into database)
  - Runs at 2-4 GB per core

HPC Requirements for 2017

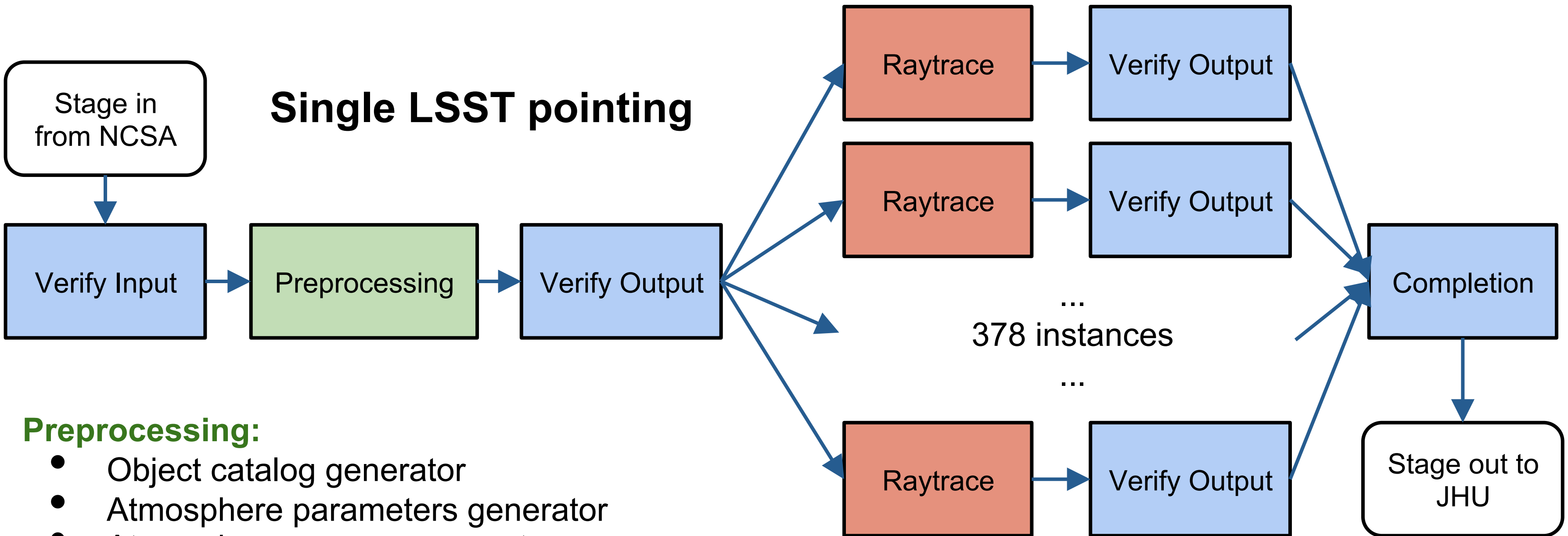
- Goals:
  - Generate and process LSST data at 10% of operational data
    - 30,000 visits (1 month of LSST data)
    - Biannually 60M CPU hrs (generate) 20M CPU hrs (analyze)
    - 80,000 cores 1GB per core (generate) 10,000 cores (analyze)
    - 400TB of data
    - Goal is the development and scaling of weak lensing measurements (e.g. multifit)
  - Rapid algorithm development for DESC
    - 200 focal planes (1.2TB) in 24 hrs
    - Once a month 200K CPU hrs x 3 concurrent (generate)
    - 20,000x3 cores 1GB core

Summary

- To scale to the size of the LSST development
  - Requires ~80,000 cores (data generation at small and large scale) with 2-4GB per core
  - Expect to generate and serve ~500TB of data (will require the analysis to be run remotely)
  - Shared memory is advantageous for the analysis
  - Many core systems would benefit image generation (requires broad adoption given cost of rearchitecting)
  - 32x increase in capabilities would enable generation of LSST data (and processing) at 3x the LSST data rate.

Backup

ImSim Computational Flow



Preprocessing:

- Object catalog generator
- Atmosphere parameters generator
- Atmosphere screens generator
- Cloud screens generator
- Telescope optics parameters

Raytracing:

- Photon raytracer
- Cosmic ray adder
- Background radiation adder
- CCD readout modeler


Optics
+Tracking
+Diffraction
+Detector Misalignments & Perturbations
+Lens Misalignments
+Mirror Misalignments Perturbations, & Micro-roughness
+Detector
+High Altitude Atmosphere
+Mid Altitude Atmosphere
+Low Altitude Atmosphere
+Pixelization
+Saturation & Blooming