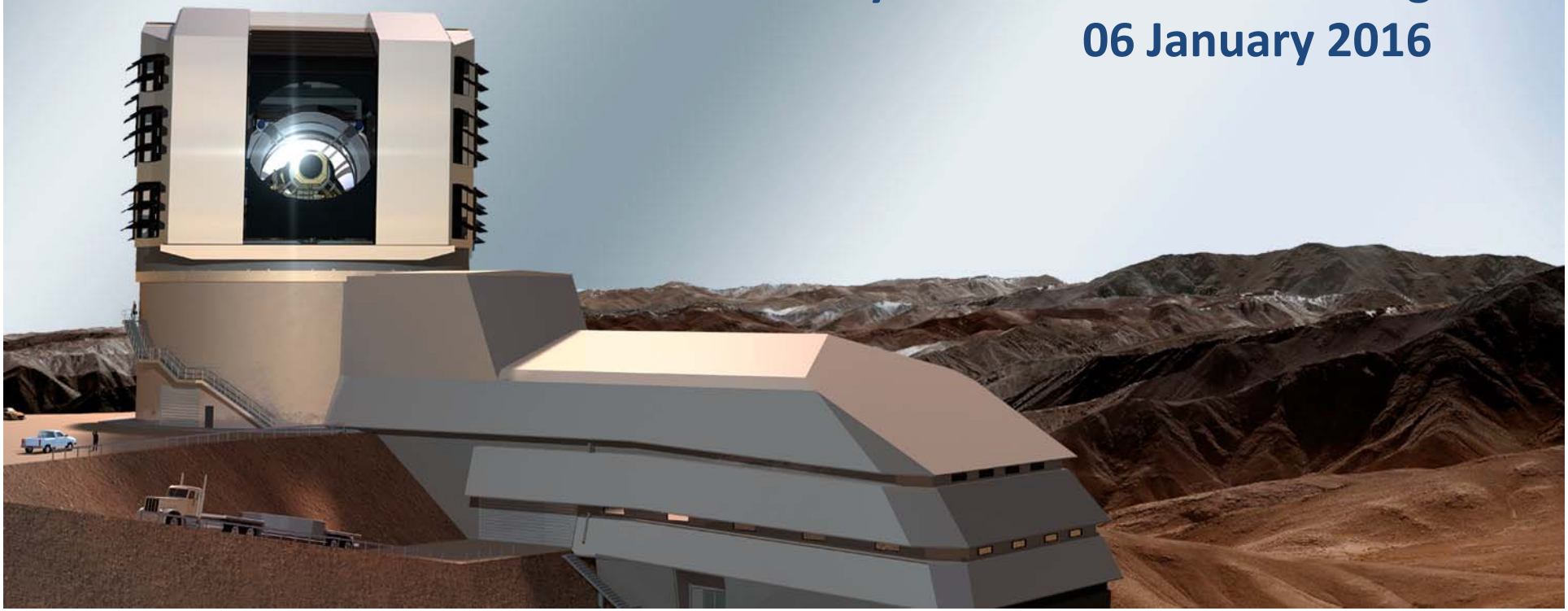
LSST and Synergies with the VO

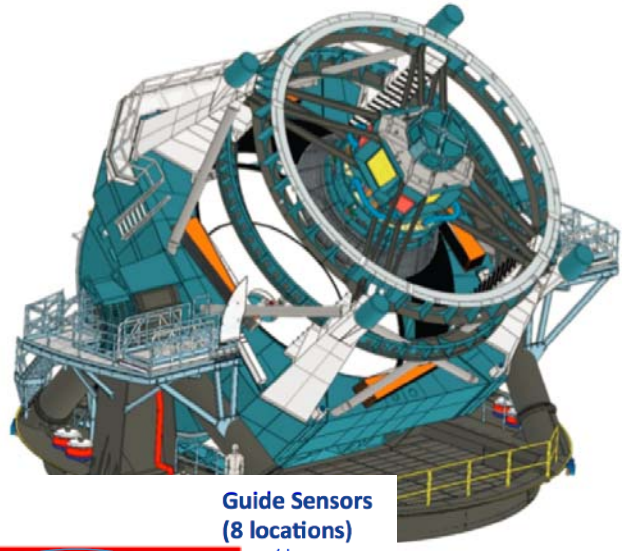
David R. Ciardi
Project Scientist: Science User Interface and Tools

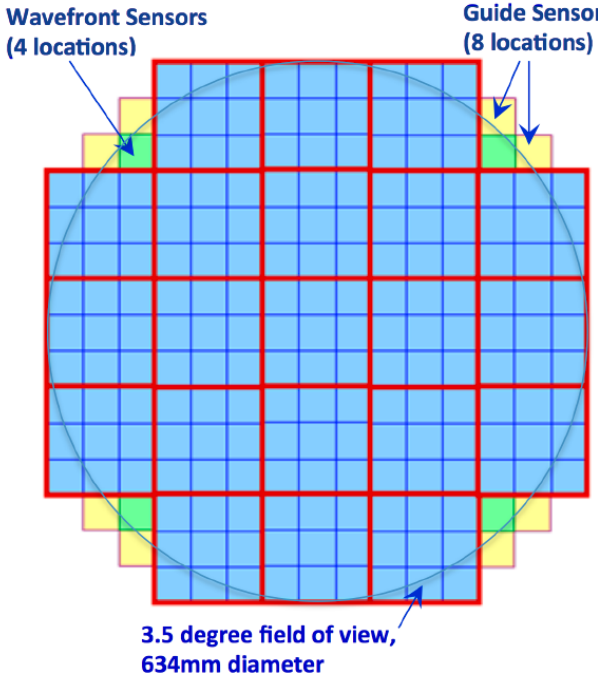
US Virtual Observatory Alliance Annual Meeting
06 January 2016



LSST In A Few Slides

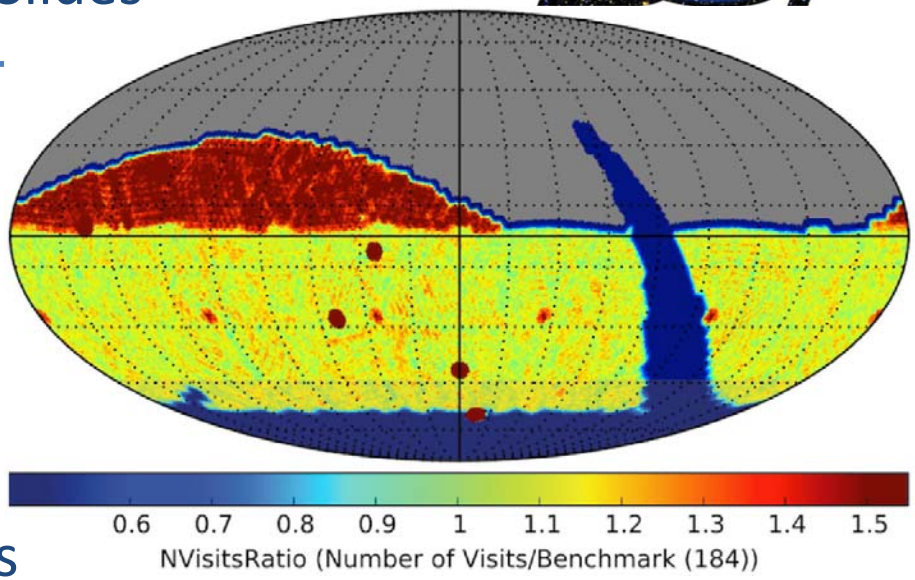
- The telescope and Camera
  - 8.4m (~6.5m effective)
  - 9.6 deg^2 field of view
  - 3.2 gigapixel camera
  - 189 4K × 4K CCD
  - 0.2'' per pixel
  - 6 bands (ugrizy; 320-1050nm)
  - 2s readout; 5s slew & settle

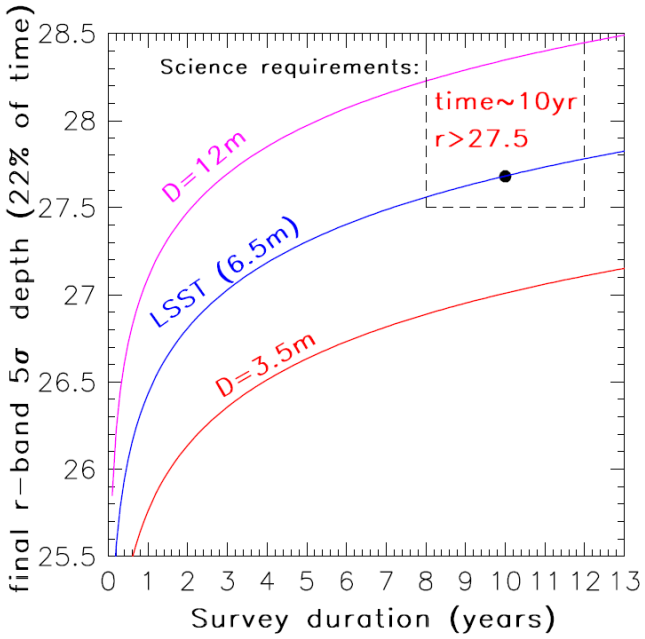




LSST In A Few Slides

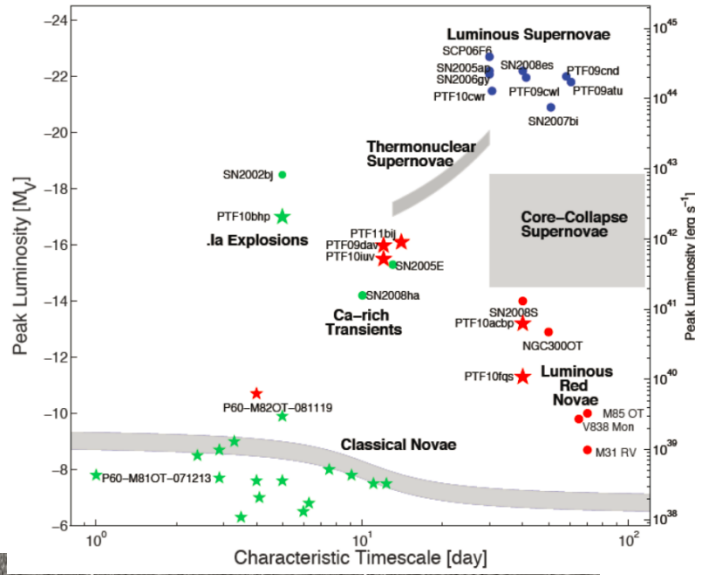
- Deep, Wide Fast survey
  - 18000+ deg2
  - 10 years
  - 2 15s back to back exposures
  - 30s exposure/visit
  - ~825 visits
  - r≈24.5/visit; r≈27.5 total

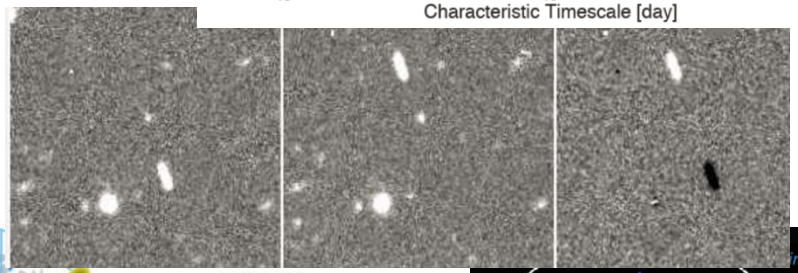


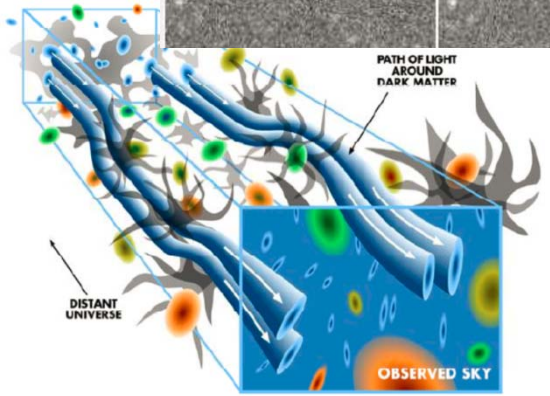


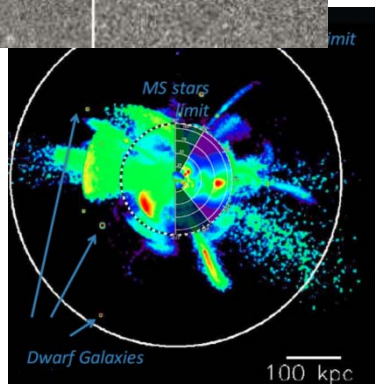
LSST In A Few Slides

- Time domain science:
  - Novae, black-hole binaries, GRBs …
  - Source characterization
  - Instantaneous discovery
- Census of the Solar System:
  - NEOs, PHAs, moving objects
  - Solar system & planet formation
- Mapping the Milky Way:
  - Structure and accretion history
  - Properties of all stars within 300 pc
- Dark energy and dark matter:
  - Strong Lensing
  - Weak Lensing
  - Supernovae









LSST Data Products

Level 1

- A stream of ~10 million time-domain events per night, detected and transmitted to event distribution networks within 60 seconds of observation.
- A catalog of orbits for ~6 million bodies in the Solar System.

Level 2

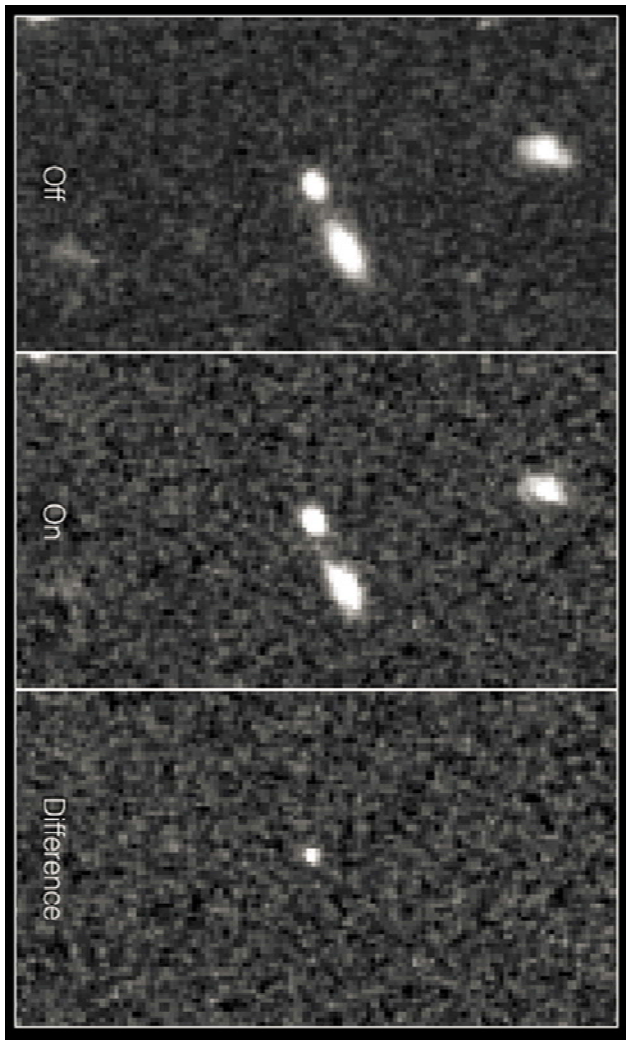
- A catalog of ~37 billion objects (20B galaxies, 17B stars), ~7 trillion observations ("sources"), and ~30 trillion measurements ("forced sources"), produced annually, accessible through online databases.
- Deep co-added images.

Level 3

- Services and computing resources at the Data Access Centres to enable user-specified custom processing and analysis.
- Software and APIs enabling development of analysis codes.

Level 1: Alerts

- State-of-the-art image differencing pipeline
- Alerts issued within 60 seconds of observation
  - 10,000,000/night (average), 10,000/visit (average), 40,000/visit (peak)
- Each alert includes:
  - Position
  - Flux, size, and shape
  - Light curves in all bands
  - Variability characterization (e.g., low-order light-curve moments, probability the object is variable, previous id)
- Cut-outs centered on the object (template, difference image)



Level 2: Annual Releases

- Calibrated and consistent catalogues & images
  - Objects, detections, detections in difference images, coadds, etc
  - Enable static sky science and time-domain science which is not time sensitive (e.g. statistical investigations of variability)
  - Made available in annual Data Releases
  - Two releases in the first year
- Complete reprocessing for each release
  - Every DR will reprocess all data
  - Including reprocessing of level 1 data
  - Forced photometry
- Projected catalog sizes:
  - 18 billion (DR1) ➜ 37 billion (DR11) separate objects
  - 750 billion (DR1) ➜ 30 trillion (DR11) individual measurements
  - Cumulative ~500 PB image and ~50 PB catalogue data

Level 3 and Workspace Environment

- Products created by the community and made available both privately and publicly
- Use-cases not fully enabled by Level 1 and 2:
  - Reprocessing images to search for SNe light echos
  - Characterization of diffuse structures (e.g., ISM)
  - Extremely crowded field photometry (e.g., globular clusters)
  - Custom measurement algorithms
- Enabling Level 3:
  - User databases and workspaces
  - Non-LSST data: catalogs and files (e.g., images, spectra, time series)
  - Enabling user computing at the LSST data center
  - For processing that will greatly benefit from co-location with the LSST data
  - Sized for ~10% of total compute budget
  - Making the LSST software stack available to end-users

Science User Interface and Tools

- No way we can anticipate all the needs/wants of the community
- LSST will build a portal that fulfills the needs of the general user (e.g., searching, image visualization, table manipulation, plotting, workspace etc.)
  - Enable users to do as much data discovery as possible
  - Facilitate data analysis by providing visualization and exploration tools
- Components re-usable by others to use and to build tools that meet their own special needs
- Manage work flow and data collections for users
  - Workspace and computing resources
- Supply software building blocks so others can build their own environment

VO Areas of Interest

- LSST is working on how best to provide access to the data and how to best take advantage of the VO protocols
- SCS, SIAP, and TAP access will almost certainly be enabled within the interfaces and workspace
  - May use thin layer to backend
  - Enable VOTable output
  - ADQL has geometric features specific to PostgreSQL which need attention
  - Will need to write an ADQL adaptor for QservSQL

VO Areas Under Consideration

- Time Series Protocol
  - Working specifically with IVOA to help define an appropriate protocol
  - Adaptation of spectra or simple time series protocols may not fully meet needs
- CAOM
  - Not really a standard, but model could help make footprint services relatively easy to implement
  - Currently does not handle catalog data (!!!)
  - Considering making tables CAOM capable or having a CAOM-compatible view
- Datalink
  - A given query has potential to return hundreds to thousands of images and data
  - Linking images – sources – time series (one-to-many and many-to-one) is crucial
- VOSpace
  - Need a workspace environment for users to store results, run code, share results, enable Level-3 data
  - Workspace science and technical requirements still under development

VOevent: Under Consideration

- We have a funded LSST Corporation proposal to engage the community about how to best evolve
  - VOevent and VOevent Transport Protocol
  - Annotations
  - More complex data
  - Needs to feed brokers that are being developed (e.g., Antares) as well as internally developed broker
- That funding starts now and carries through 2016

Alert System Questions

- Science Needs of the community
  - How are unique events identified and labeled?
  - How are events characterized?
  - How are events presented in a simple and digestible form?
  - How is multi-wavelength time-series information packaged?
  - How are annotations and additional data, necessary for the characterization and classification of events, supported?
  - What is the scientific workflow of event follow-up and how does the alert system support such a workflow?
- Current Infrastructure
  - Can the current event infrastructure be adapted to address the scientific needs of the community?
  - Is the current alert distribution infrastructure scalable to the needs of LSST?
  - What infrastructure is missing or poorly defined?

Time Scale for Alert System Study

Date	Task
November 2015	Initial Kick-off Planning remote/in-person meeting
January 2016	Remote 1: Assessing the Scientific Requirements
February 2016	Remote 2: Assessing the Technical Requirements
March 2016	Remote 3: Assessing the Current Infrastructure 1
April 2016	Remote 4: Assessing the Current Infrastructure 2
May 2016	Remote 5: Planning for the Technical Workshop
June 2016	Pasadena, CA: 3-day Technical Workshop
Jun – Aug 2016	Post-workshop proof-of-concept alert system development
July 2016	Remote 6: Post-workshop planning for final system development 1
August 2016	Remote 7: Post-workshop planning for final system development 2
September 2016	Final delivery of white paper and proof-of-concept alert system

Probably will push this schedule back by 1 – 2 months but gives an idea of the schedule we are trying to set

LSST Time Scales

- Construction: 2014 – 2019
- Commissioning: 2019 – 2020
- Science Verification: 2021 – 2022
- Operational Survey: 2022 – 2032

- Our engagement of the VO community will continue to grow as we move into the implementation and construction phase and need to develop solutions for the problems we encounter