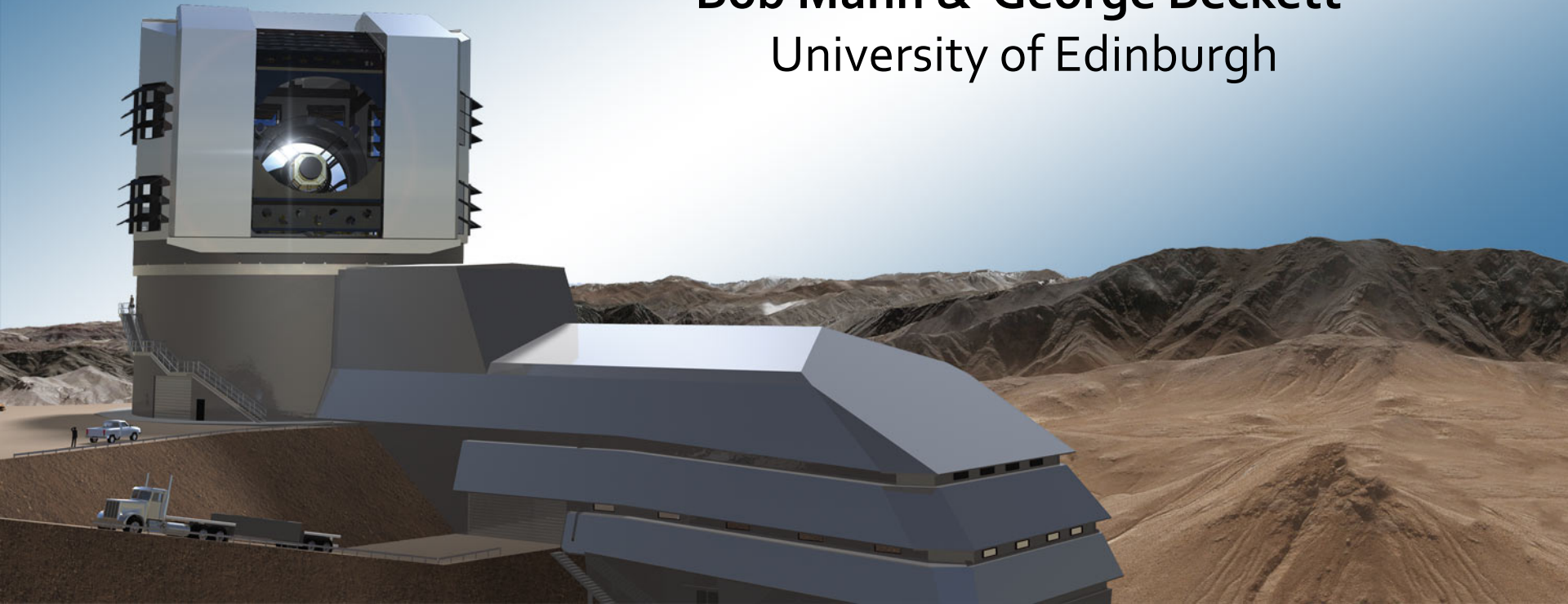


# LSST:UK and UKTO

Bob Mann & George Beckett  
University of Edinburgh



# Outline

- The Large Synoptic Survey Telescope
- LSST:UK
- LSST:UK computing requirements
- LSST:UK and UKTO

# Large Synoptic Survey Telescope

- Large optical survey telescope under construction in Chile
  - Ten year sky survey from 2022
  - US-led: NSF + DoE, plus foreign partners (UK & France largest)
  - 6.5m effective primary; 9.6 sq. deg FOV
- Étendue = mirror area x camera field of view
- If étendue is large enough, can go **wide, deep and fast** at same time
  - Different kinds of analysis from the same dataset
  - LSST will have a great impact across almost all of astronomy



# The LSST Science Book

- Contents:
  - Introduction
  - LSST System Design
  - System Performance
  - Education and Public Outreach
  - The Solar System
  - Stellar Populations
  - Milky Way and Local Volume Structure
  - The Transient and Variable Universe
  - Galaxies
  - Active Galactic Nuclei
  - Supernovae
  - Strong Lenses
  - Large-Scale Structure
  - Weak Lensing
  - Cosmological Physics



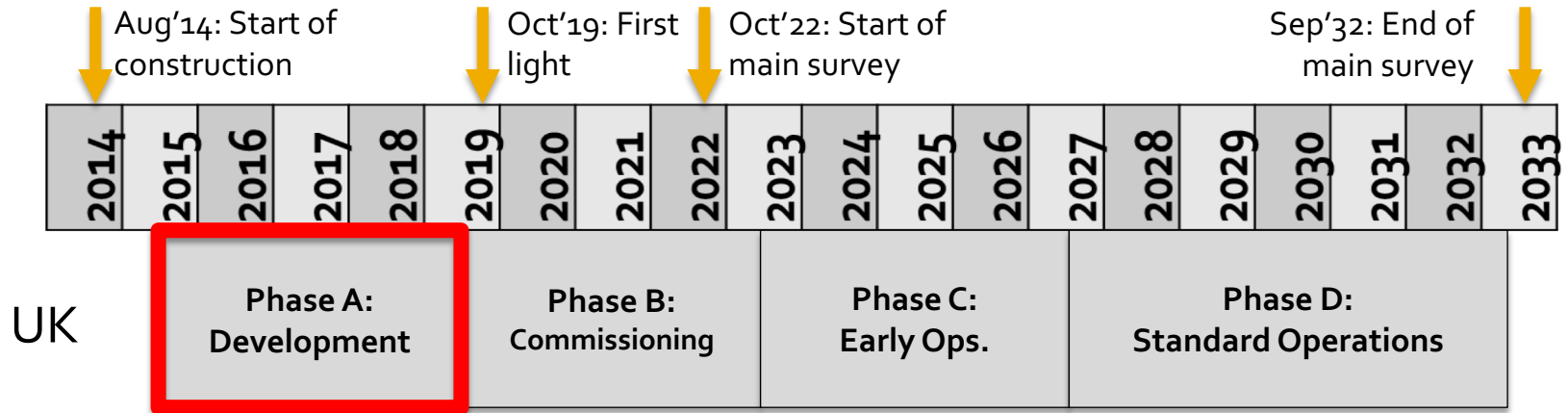
**arXiv:0912.0201**

# LSST:UK Consortium



Every astronomy group in the UK

# LSST:UK Science Centre Programme



- Four-phase programme
  - Forecast budget of ~£32M (not incl. capital for Data Access Centre)
- 2015: STFC awarded first tranche (£17.7M)
  - £15M operations contributions – data rights for 100 PIs
  - £2.7M for Phase A R&D
- 2018: Phase B proposal due next month

# Computing centres

- Archive Site
  - Level 1: Nightly difference imaging: → alert stream
  - Level 2: Annual direct image pipeline: → data releases
- Data Access Centre
  - Ingest data releases
  - Provide user access to data
  - Run Level 3 data analysis code
- Archive Sites
  - NCSA
  - CC-IN2p3 (Lyon) [L2 only]
- Data Access Centres
  - NCSA
  - Chile
  - Lyon
  - UK
  - Others?

# LSST:UK Computing Requirements

- Following examples all inexact by factors of several
  - Difficult to get quantitative estimates from many users while still planning their science
  - Unclear which science analyses will be international
  - Unclear how different Data Access Centres will coordinate activities and share workload



# Development Activities [Phase A]

- Largest requirements likely to be supporting Dark Energy Science Collaboration
  - Data Challenge 2: Summer 2018—Spring 2019: UK contribution of ~20M core hours
  - Data Challenge 3: ~50M core hours

# Contribution to Commissioning [Phase B]

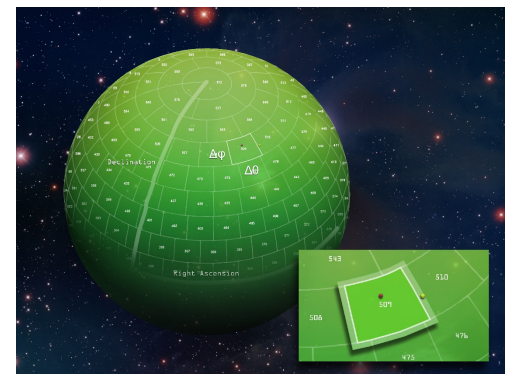
- Modest volume of images generated: ~few PB
- Repeated analysis with different parameters
- Expect ~700 cores over ~2 years
- Storage of 2PB increasing to ~12PB

# Transient Alerts [Phases B-D]

- Each night, LSST will issue millions of potentially interesting transient alerts
- LSST:UK proposes to analyse full stream in real-time, providing alert classifier
- Dedicated Transient Alert service
  - Starting with precursor survey (ZTF)
- Storage
  - 100 TB/year (some as buffer, some persistent)

# Hosting survey data [Phases C and D]

- Sky Survey produced as 11 Data Releases
  - Catalogues and image files
  - In 2022, DR1 will be ~10PB
  - ...
  - In 2032, DR11 will be ~70PB
  - Will need to have two most recent online
- Database cluster (Qserv)
  - 130 Tflops in 2022
  - ...
  - 330 Tflops in 2032



# Data Analysis [Phases C and D]

- UK astronomers will have range of storage & compute requirements
  - Estimate 10% of Archive Site capacity
  - 20 Tflops in 2022, ..., 140 Tflops in 2032
- Some Level 3 analyses produce significant data products
  - 5 PB in 2022, ..., 35 PB in 2032
- Not including Science Collaboration requirements, such as DESC (50M core hours/year from 2019—2022, then ???)

# Summary (used in BEIS bid)

CPU Power	2018/ 2019	2019/ 2020	2020/ 2021	2021/ 2022
Required (k core-years)	2.4 (DESC) + 0.1 (DEV)	5.7 (DESC) + 0.7 (COMM) + 0.7 (DEV)	5.7 (DESC) + 0.7 (COMM) + 0.7 (DEV)	4.5 (DB) + 5.7 (DESC) + 0.7 (GEN)
Storage Capacity	2018/ 2019	2019/ 2020	2020/ 2021	2021/ 2022
Required (PB)	0.5 (DESC) + 0.5 (DEV)	4 (COM) + 2 (DEV) + 0.5 (DESC)	12PB (COM)+ 2PB (DEV)	10 (DR1, 2) + 5 (GEN) + 0.1 (TRANS)

Should become clearer  
over next year or so

<sup>1</sup> Where estimates are in TFlops, a conversion of 35 core years/ Teraflops has been used.

# LSST:UK and UKTO

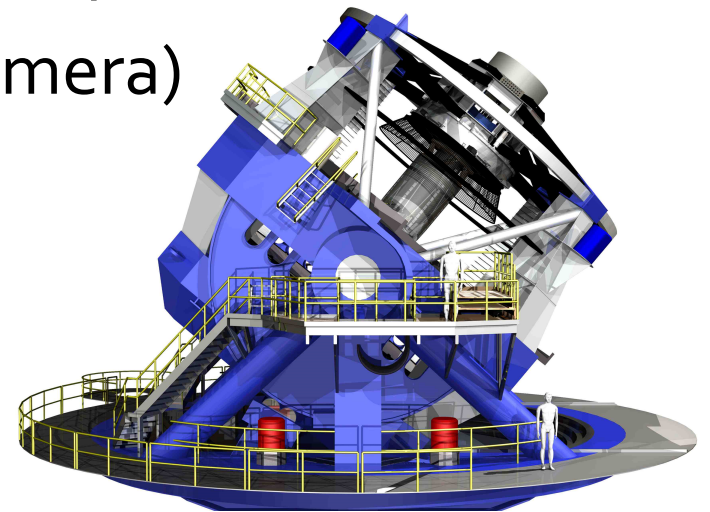
- LSST:UK is very supportive of goals of UKTO
  - Expect UK DAC to be implemented within UKTO
  - Can see lots of expertise we can benefit from
  - Can contribute expertise in some areas – e.g. DBs
- Have run test projects on GridPP and looking forward to working with database testbed
- But...contribution we can make to generic UKTO work really limited by available staff effort





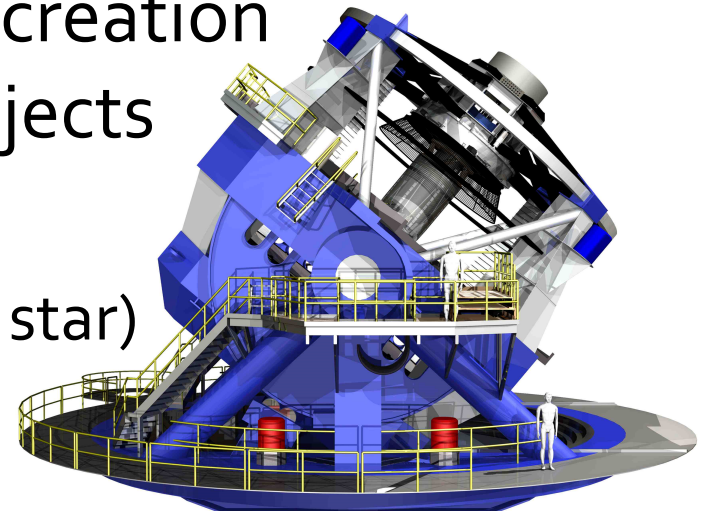
# Large Synoptic Survey Telescope

- Costs
  - Construction budget of \$640M
  - Operational budget of \$370M
- Funding
  - US National Science Foundation;
  - US Department of Energy (camera)
  - International investment (\$100M of operational costs)



# Large Synoptic Survey Telescope

- Will run 10-year survey, starting in 2022
- Will image 30k sq. deg. of sky every three nights
  - Observe each 'patch' more than 800 times
- Stacked images help identify very faint images
  - Expect 24Bn new galaxies and 14Bn new stars
- Frequent, regular visits allow creation of time-series for dynamic objects
  - Near-earth objects
  - Supernovae (explosive death of star)



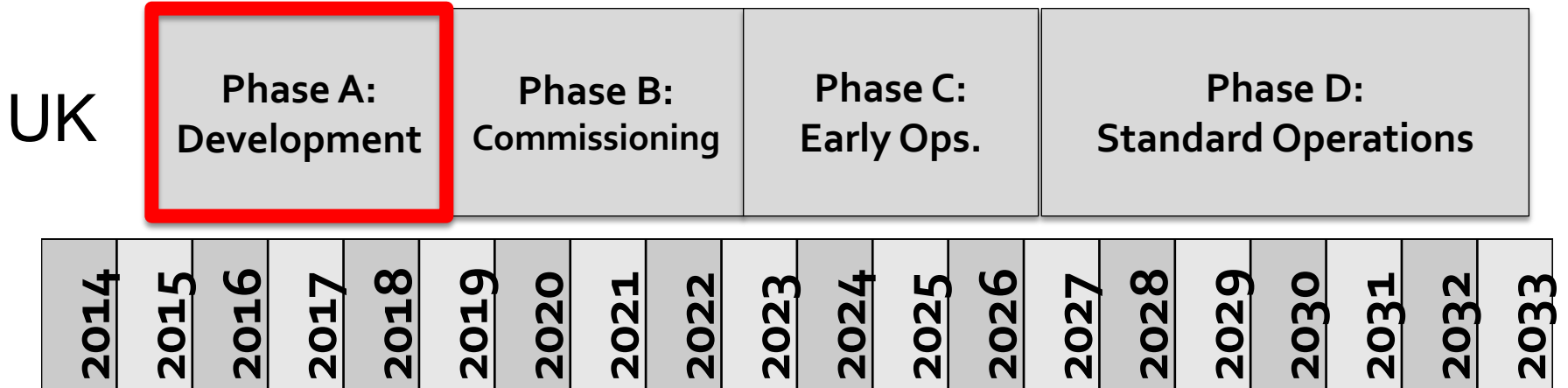
# Data Management

- Data is proprietary for two years – restricted to LSST PIs
- Survey data hosted in Data Access Centres
  - in Chile – hosting and local computing
  - at NSCA, USA – processing, hosting, and local computing
  - at IN2P3, in France – processing and hosting
- UK will host regional Data Access Centre
  - Hosting and local computing
  - possibly supporting more than UK astronomy

# Data Classification

- Four classifications of data
  - Raw data – images taken by camera
  - Level 1 Products – difference imaging data w.r.t. reference
    - Released nightly, detail objects that have unexpectedly changed brightness or position
    - $\sim 10^6$  events published each night (metadata and postage-stamp image)
  - Level 2 Products – image products (reduced and calibrated) plus catalogues
    - Released annually, as incremental Data Releases
    - D/R grows from 1.6PB in Yr 1 to 31PB in Yr 10
  - Level 3 Products – datasets derived from Levels 1 and 2 products
    - Not part of survey output: created by community
    - Estimate 10% of computing and storage requirements for Level 3 products

# Timeline for LSST



- 1 August 2014: start of construction project
- October 2019: telescope First Light
- October 2022: start of main survey operations
- September 2032: end of main survey

# LSST:UK Phase A



- £2.7M for Phase A programme (July 2015—Mar 2019)
- Edinburgh is coordinator
  - Bob Mann is PI; MGB is project manager and technical lead
- LUSC-DAC: Data Access Centre (6 staff years, Edinburgh)
  - DAC design – DAC testbed, Data Challenges, support for LUSC-DEV
- LUSC-DEV: (16 staff years, Man, Cam, QUB, Soton, UCL, Oxf)
  - Weak lensing – analysis of galaxy intrinsic alignment, shape classification
  - Milky Way – star/galaxy separation, tidal stream detection
  - Transients – alert handling, classification
  - Solar System – NEO, light-curve analysis
  - Sensor characterization – image analysis systematics

## LSST Data Volume and Products



Raw Data

Level 1

Level 2

- Two 6.4-gigabyte images (one visit) every 39 seconds
- ~1000 visits each night, ~300 nights a year
- Up to 450 calibration exposures per day
- >15 terabytes of raw data in each 24 hour period
  
- Can detect >2 million real time events per night, for 10 years
- Changes detected, transmitted, within 60 seconds of each observation
  
- A catalog of ~38 billion observed objects (24B galaxies, 14B stars)\*
- A catalog of ~5 trillion detections (“sources”) and ~32 trillion measurements (“forced sources”) \*

(\*) current baseline estimate

For more see Section 3.5 of LSST SRD (<http://ls.st/srd>)