



Large-Scale Astronomy data- management at NCSA

30 minutes)

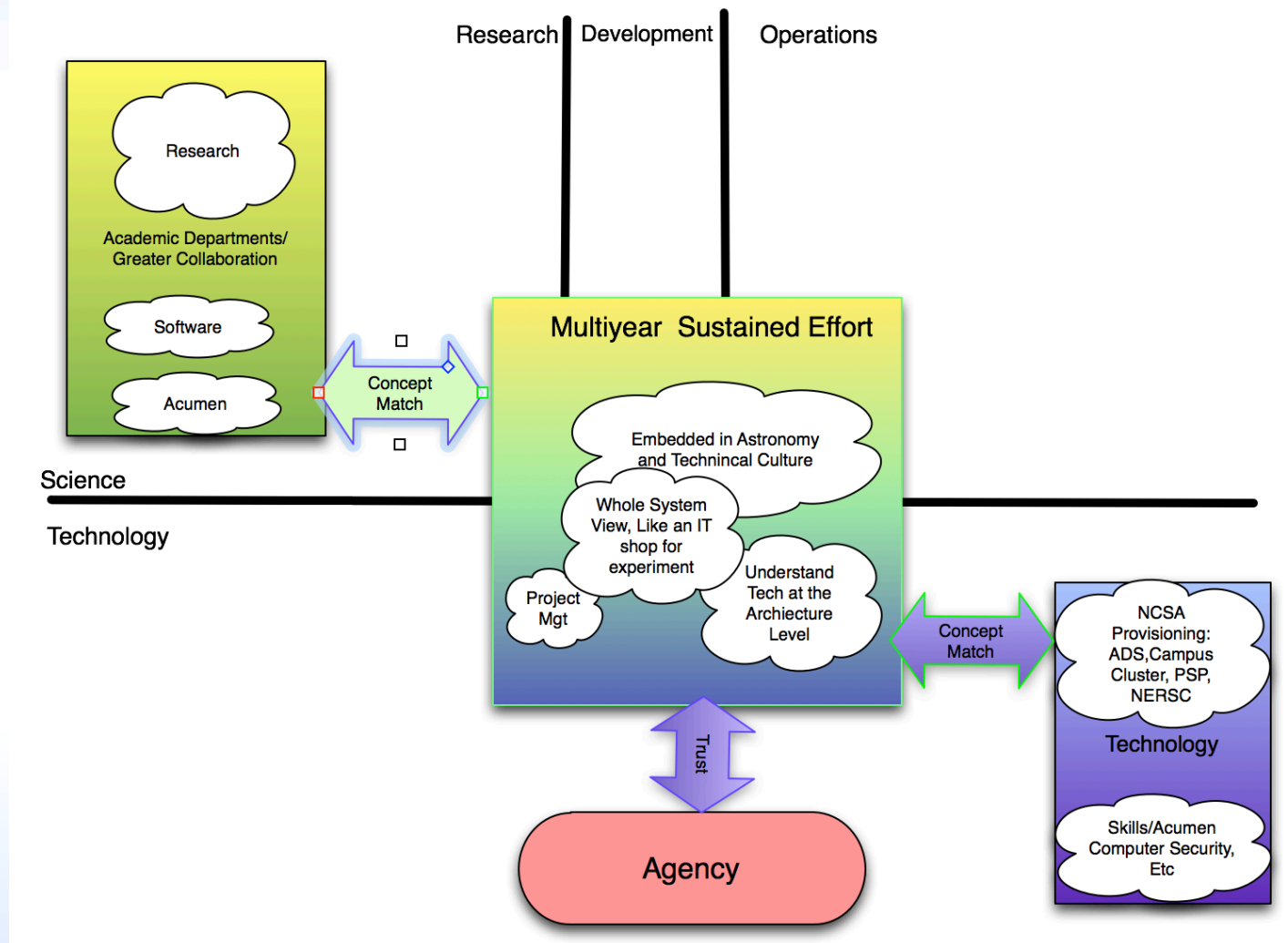


National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign

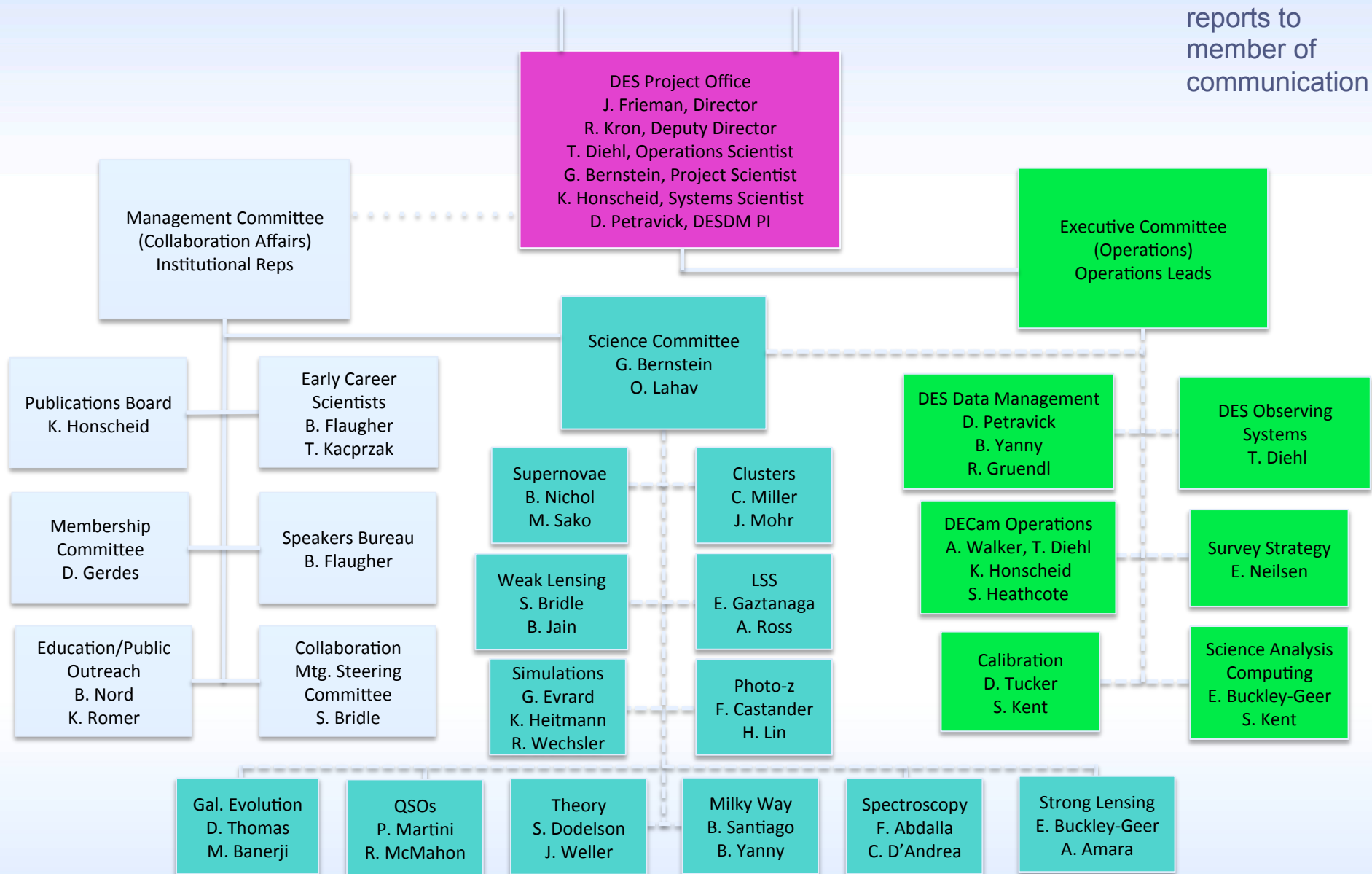
NCSA/U of I Astronomy

- Currently in Production
 - BIMA/CARMA millimeter radio array + Archive.
 - (U of I Astronomy)
 - South Pole Telescope.
 - (U of I astronomy)
 - DES -- Dark Energy Survey.
 - (NCSA)
- In construction:
 - LSST
 - (NCSA)
- Next door: ATLAS distributed T-2 (Physics)
 - NCSA supports the campus cluster where it reside

Notes on Astronomy Section Organization



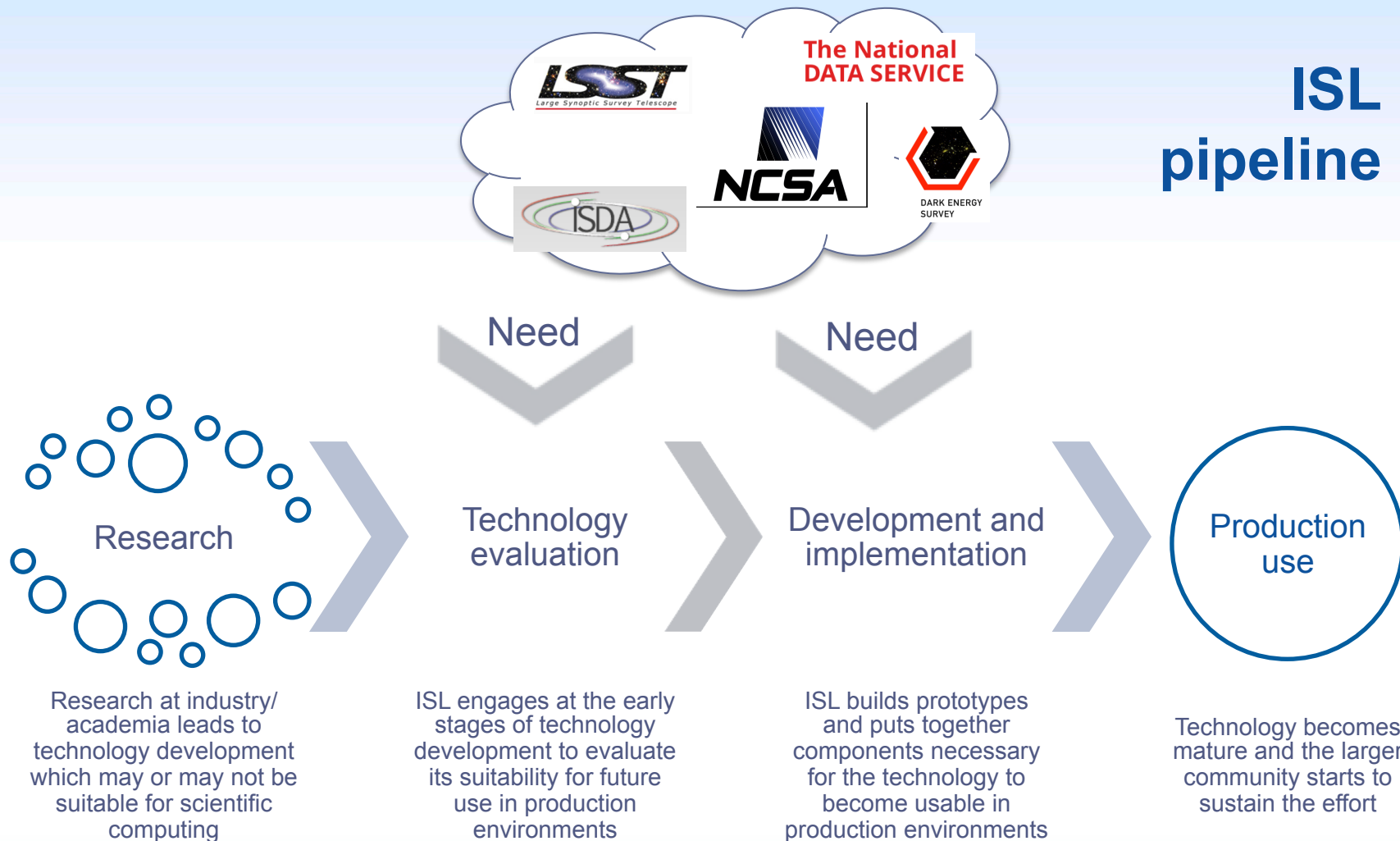
reports to
member of
communication



Experience (Infrastructure)

- NCSA Supercomputing program:
 - XSEDE national cyber structure
 - Blue waters – 200,000 cores. (25 PB disk, ½ ExB tape)
 - NCSA Private Sector program cluster.
- DES and experience in the Astronomy Group:
 - High Energy Physics (LHC, Run II, etc.)
 - DES use of facilities at
 - NERSC
 - FERMILAB (Open Science Grid)
- LSST
 - LSST security Officer (Alex Withers) is at NCSA, ex Brookhaven Lab

ISL pipeline



“ISL actively seeks new partnerships with external collaborators in Technology and applications areas related to data-intensive computing.”

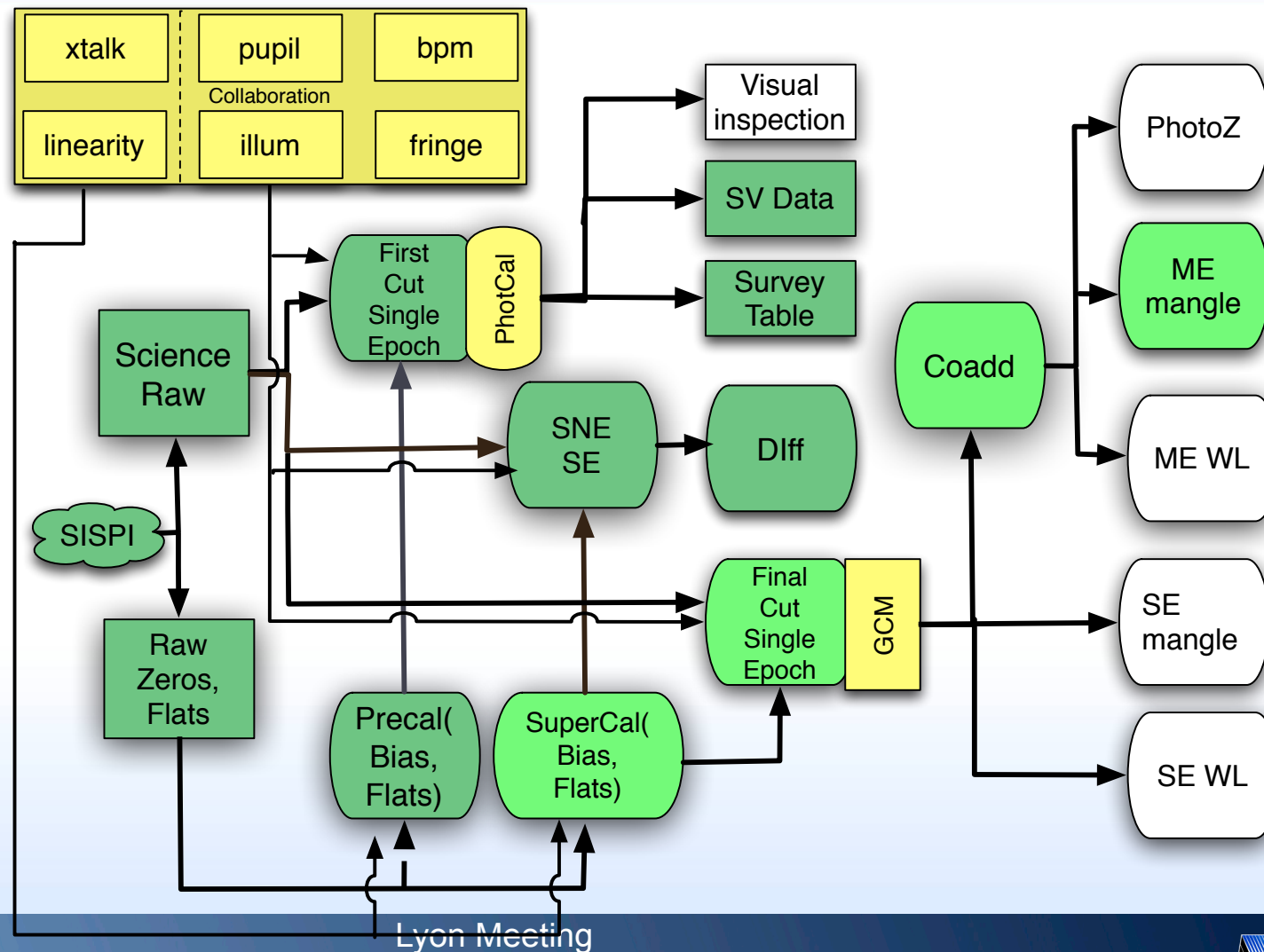
DES Overview

- Science: probe cosmic acceleration with 4 techniques:
 - Clusters, Weak Lensing, Large-scale Structure, Supernovae
- Two multiband imaging surveys:
 - 5000 deg² grizY: positions, shapes, photo-z's for 200 million galaxies, 100,000 clusters
 - 30 deg² griz time-domain survey: light curves for 3500 type Ia supernovae
- Collaboration-built DECam: 3 deg² FOV, 570-megapixel imager and 5-element optical corrector for Blanco 4-meter telescope at CTIO.
- Data management (DESDM) system to produce science-ready data products.
 - Five-year (525-night) survey began Aug. 2013.
 - Observing Seasons run Aug.-Feb.
 - Nearly 2/3 through 2nd season

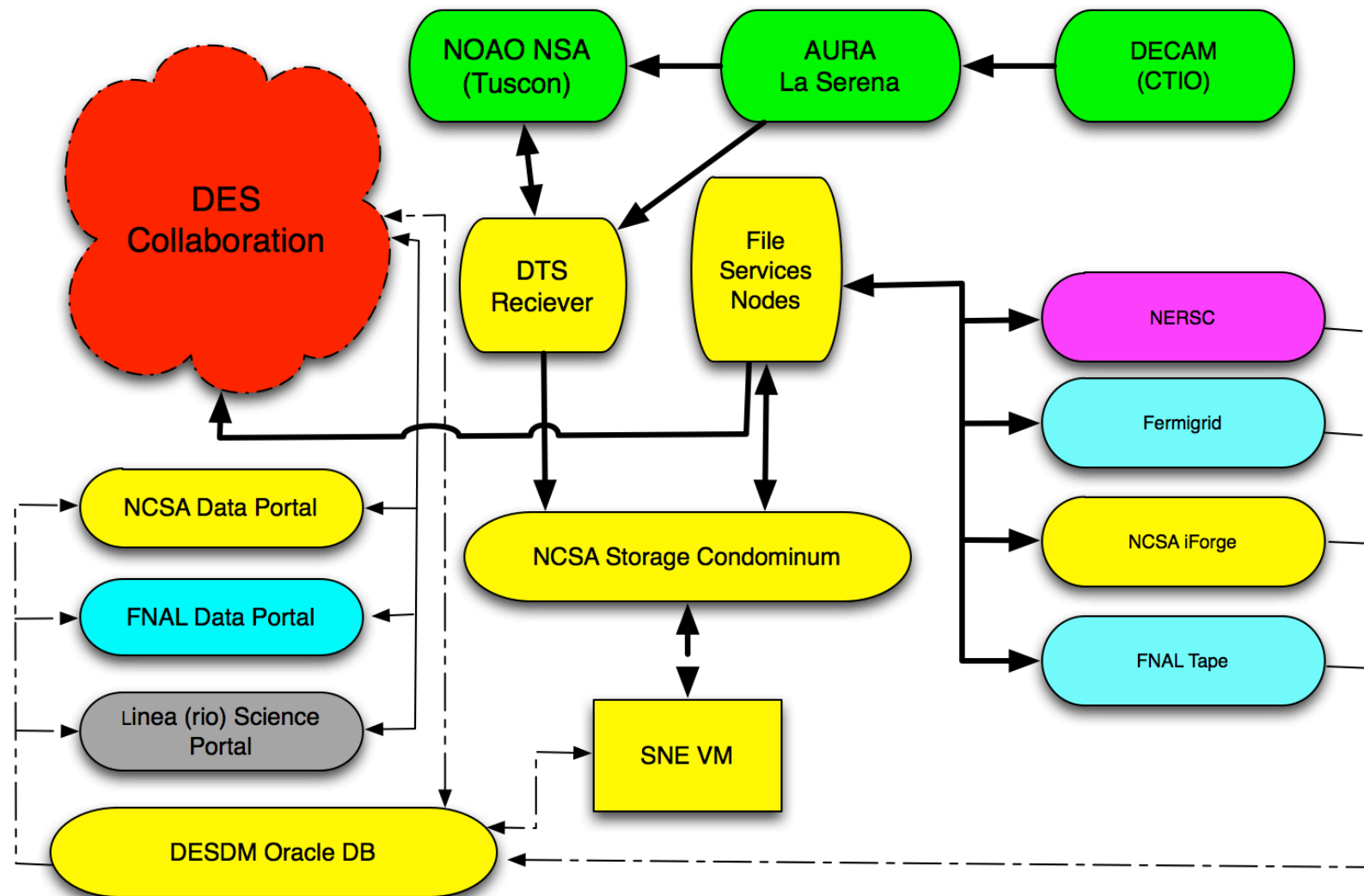
DES Context

- What is the physical origin of cosmic acceleration?
 - Dark Energy (DE) or Modification of Gravity?
 - Determine DE Equation of State parameter & its time evolution
- The requisite measurements are challenging
 - Multiple methods needed, control of systematics paramount
- Dark Energy Task Force (DETF, 2006):
 - Stage III: **DES** (imaging), BOSS/eBOSS (spectroscopy), (HSC, KIDS)
 - Stage IV: **LSST** (imaging) 2022-, DESI (spectroscopy) 2018-, (Euclid, WFIRST)
- Imaging and spectroscopic surveys provide complementary probes.

Production process (too simple)



Current DES



Dark Energy Survey Production

- Building blocks:
 - Central storage (~2PB) in NCSA central storage condominium.
 - 125 core, 100 (usable) TB table space
 - Grid Model for Bulk Computing (next slide)
 - Cluster for central services (~10 nodes)
 - Job management (condor)
 - File transfer (http/s, gridFTP, r sync).
 - Quality assurance
 - Ongoing development
 - Running less mature pipelines

Bulk Computing Platforms

Capablilty	Description	Challenges
Blue Waters	200,000 core extreme machines, 25PB lustre file system	Limited outbound connectivity -- just relaxed
Fermigrid (OSG)	Collection of servers, CVMFS file system for software; large number of cores available	Single-core allocation model (being relaxed), weak central file system
NERSC	Batch system that allow DOCKER containers	More capacity needed
Private Sector Program	600 available cores w/ competent GPFS file system	No longer large enough for all processing
U of Illinois Campus Cluster	condominium holding investigator-owned computers, idle cycles available	96 core dedicated capability -> 500-600 dedicated to DES in 2016
XSEDE	National Cyber infrastrucutre oriented towards simulation	Weak file systems for simulation, queue wait times, turn around need for nightly processing