



DARK ENERGY
SURVEY

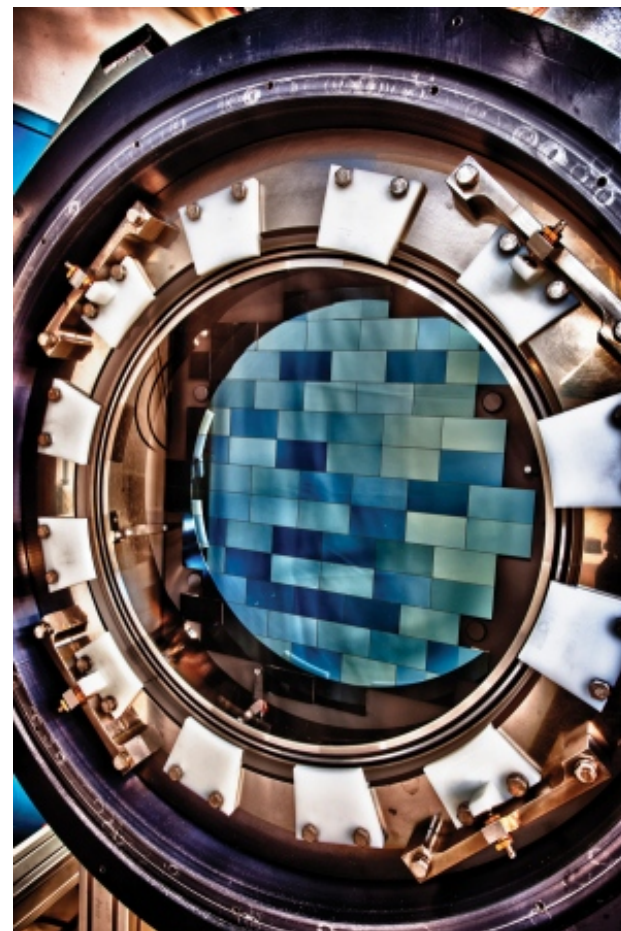
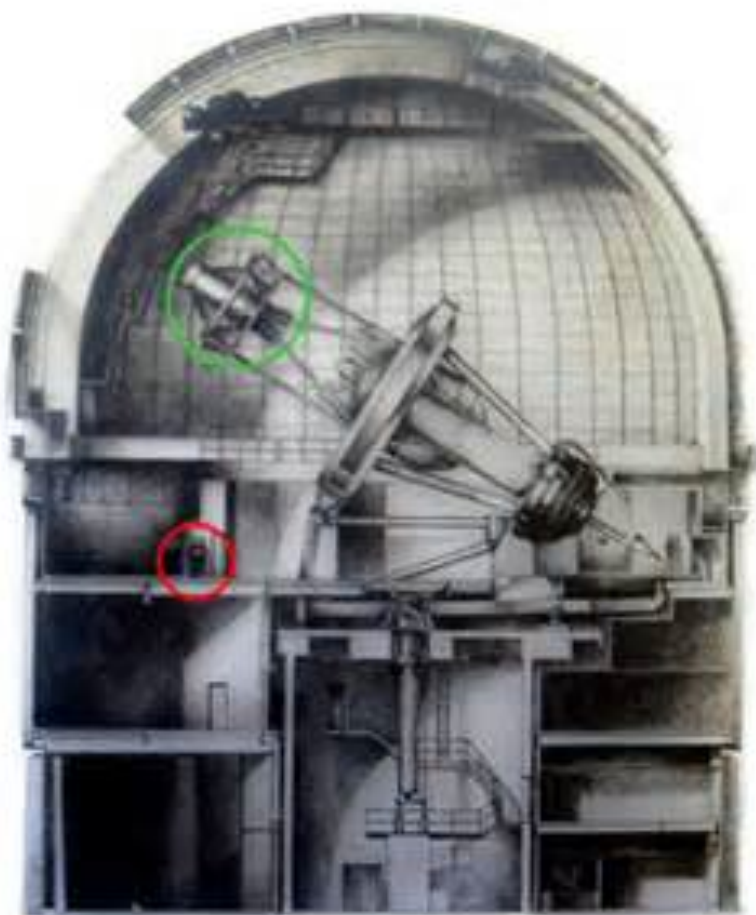
DARK Energy Survey Data Management

Don Petravick



DARK ENERGY
SURVEY

DECAM and the 4m Blanco Telescope



10/15/13

Dark Energy Survey -- CHEP13

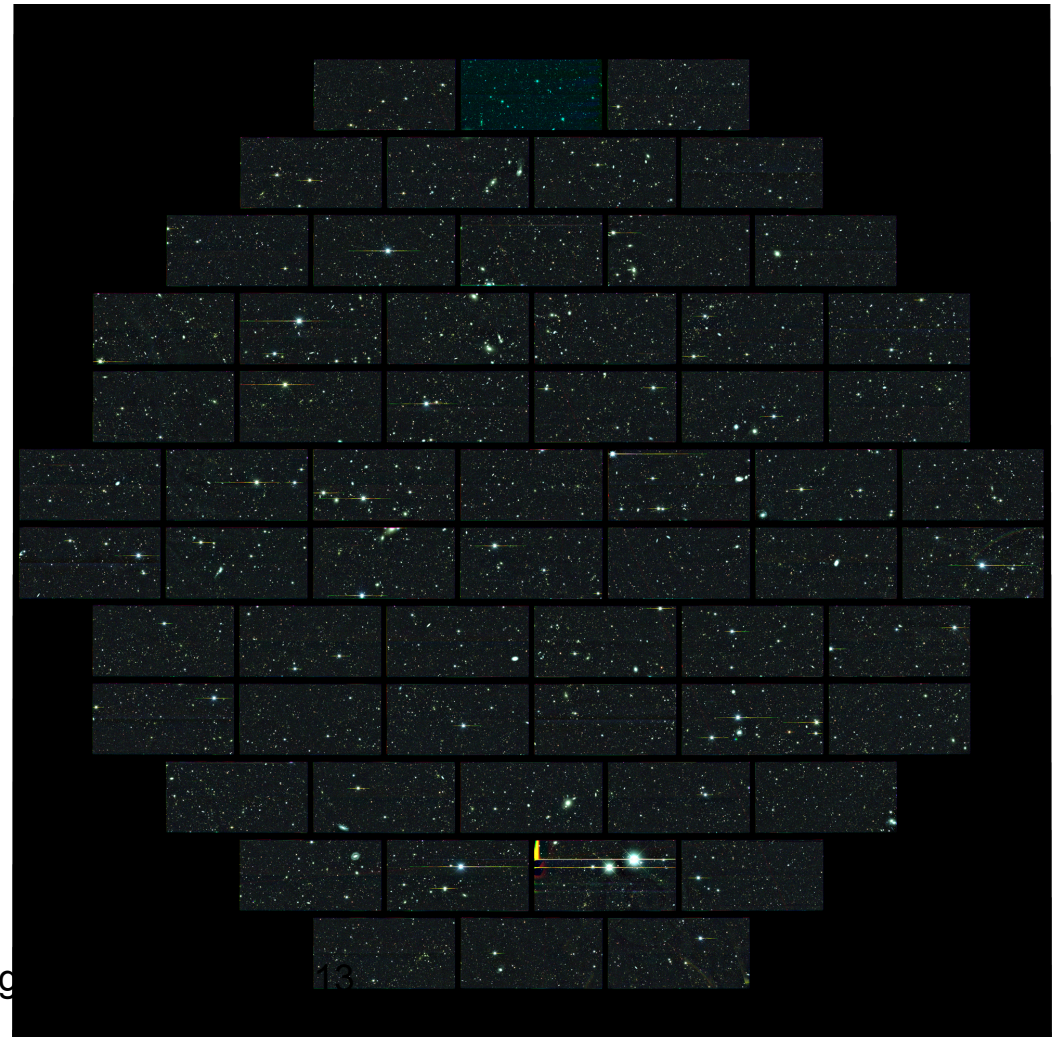


The Dark Energy Survey

DARK ENERGY
SURVEY

- Stage III DE project using 4 complementary techniques:
 - I. Clusters
 - II. Weak Lensing
 - III. Large-scale Structure
 - IV. Supernovae
- Two multiband surveys:
 - 5000 deg² *grizY* to 24th mag
 - 30 deg² time-domain *griz* (SNe)
- Built new 3 deg² FOV camera and Data management system
 - Survey 2013-2018 (525 nights)
 - Premier facility instrument for astronomy community

Composite image of deep SN field





DES Science Summary

DARK ENERGY
SURVEY

Four Probes of Dark Energy

- **Galaxy Clusters**

- ~100,000 clusters to $z > 1$
- Synergy/overlap with SPT
- Growth of structure and expansion history

- **Weak Lensing**

- Shape measurements of 200 million galaxies
- Growth of structure and expansion history

- **Large-scale Structure (BAO)**

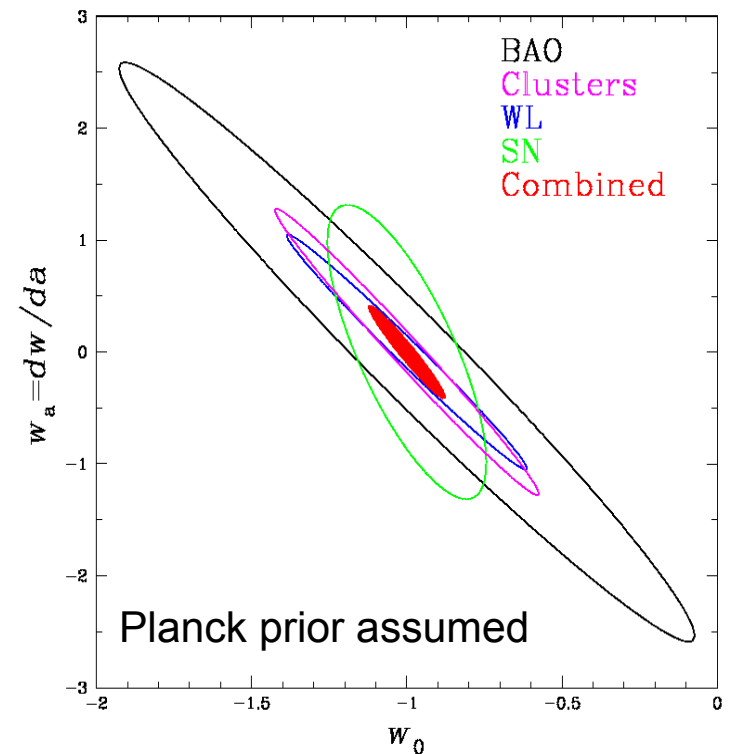
- 300 million galaxies to $z = 1$ and beyond
- Expansion history

- **Supernovae**

- 30 sq deg time-domain survey
- ~4000 well-sampled SNe Ia to $z \sim 1$

10/15/13 Expansion history

Forecast Constraints on DE Equation of State

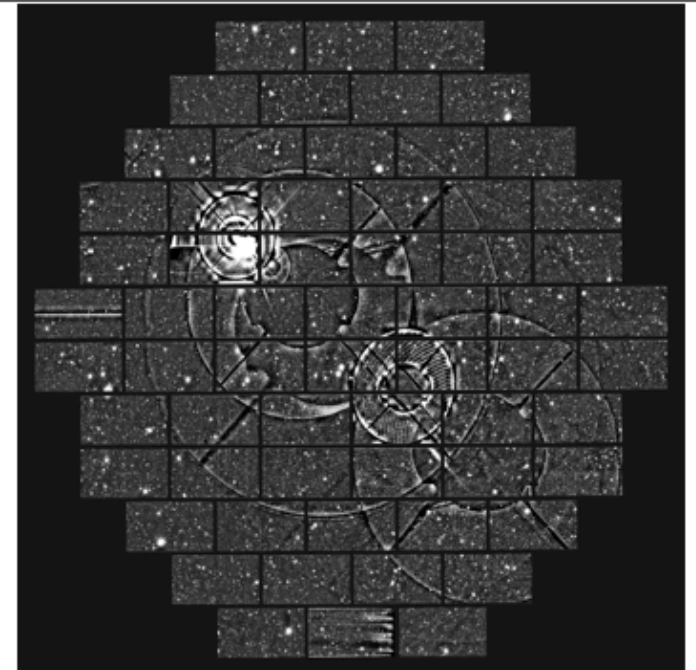
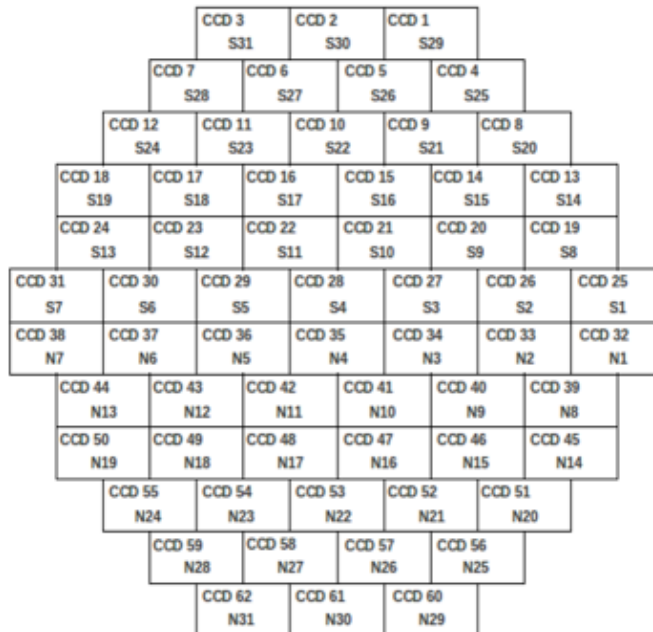


Factor 3-5 improvement over
Stage II DETF Figure of Merit



DARK ENERGY SURVEY

exposure: DECam_00229257
 object: DES survey hex -532-562 tiling 3
 band: g
 coords: 305.75853, 57.02553
 exposure time: 90.0
 airmass: 1.23000001907
 Firstcut assessment: good



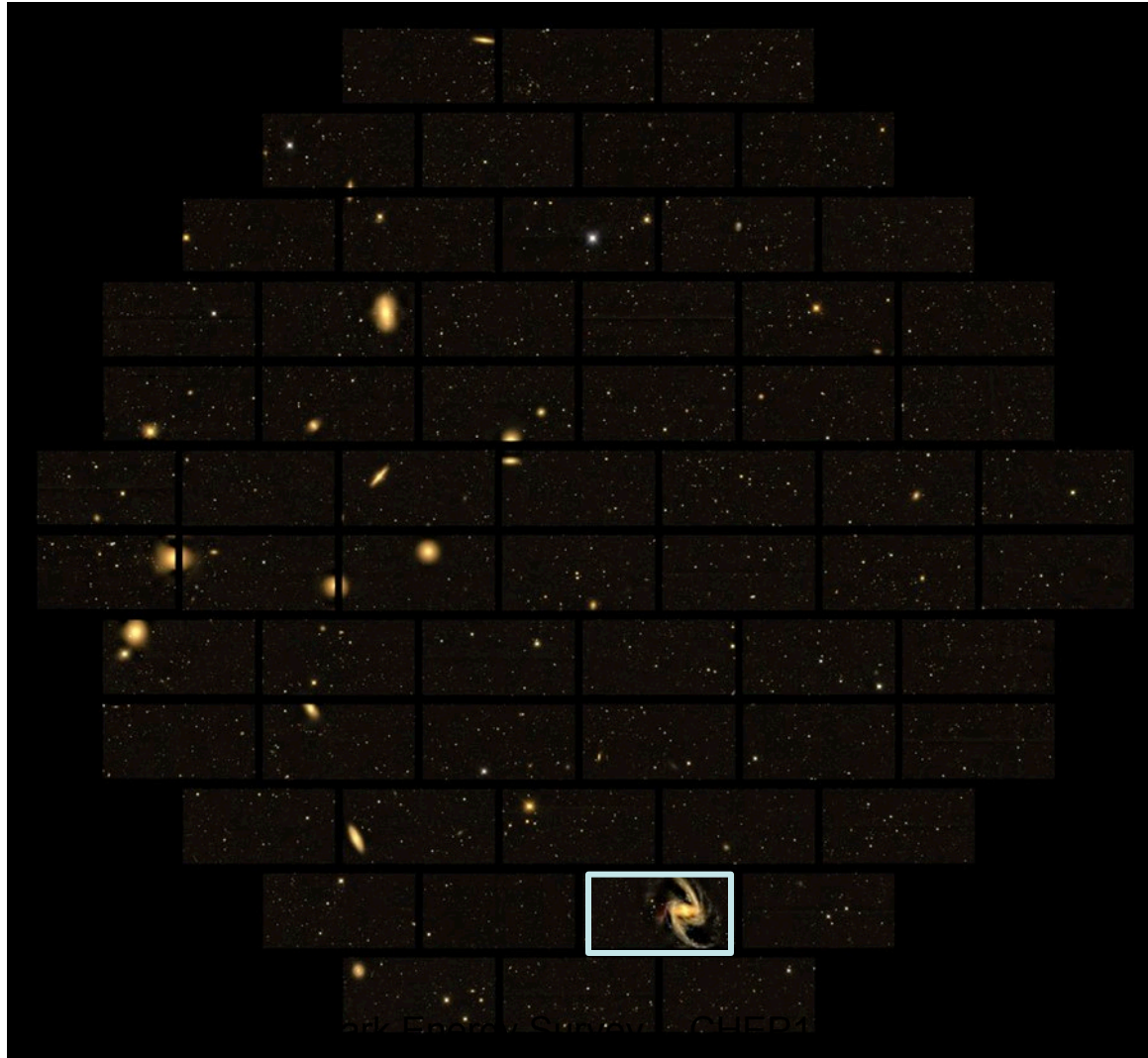
10/15/13

Dark Energy Survey -- CHEP13



DARK ENERGY
SURVEY

DECam Images

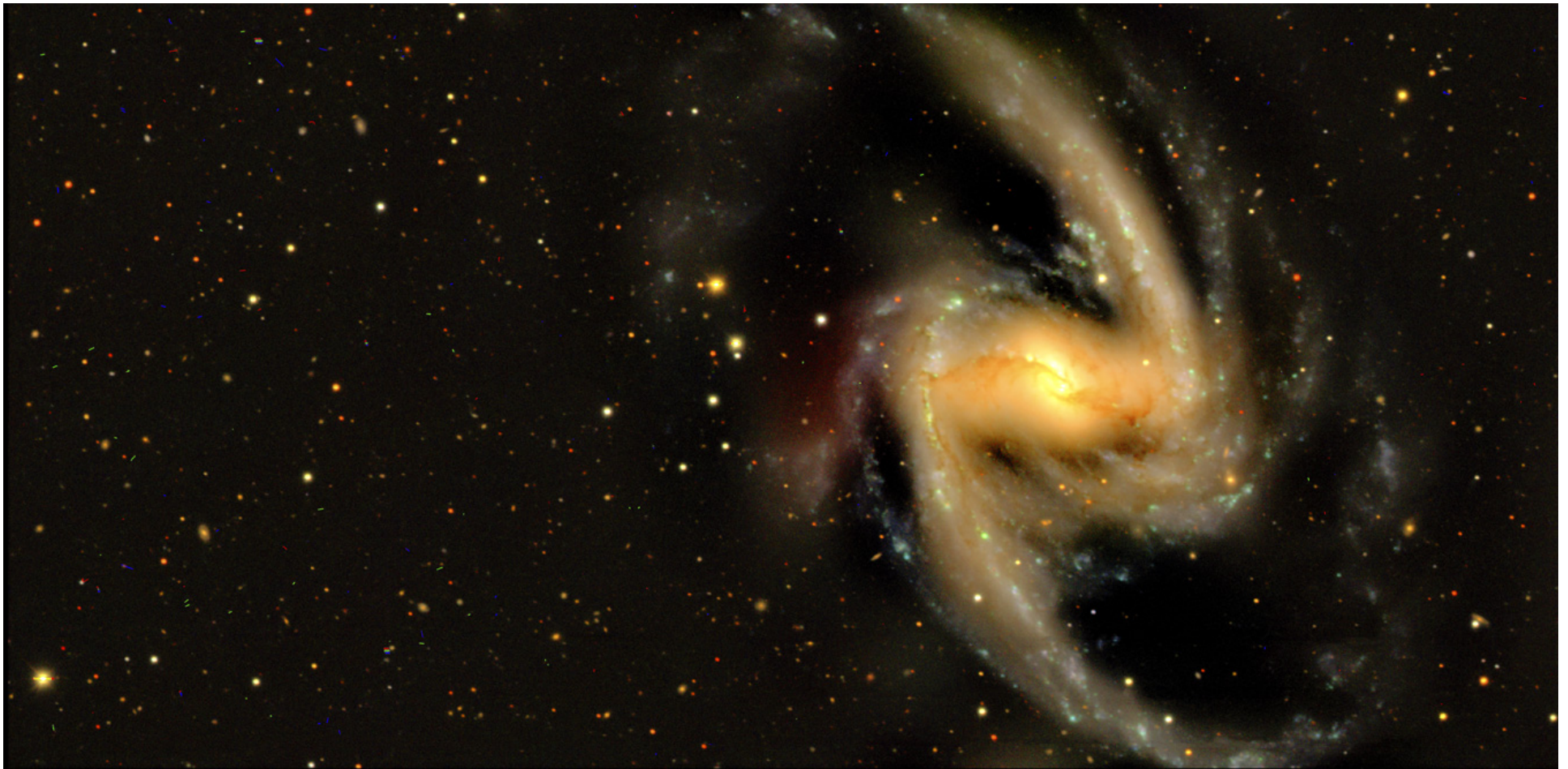


10/15/13



DARK ENERGY
SURVEY

DECam Images



10/15/13

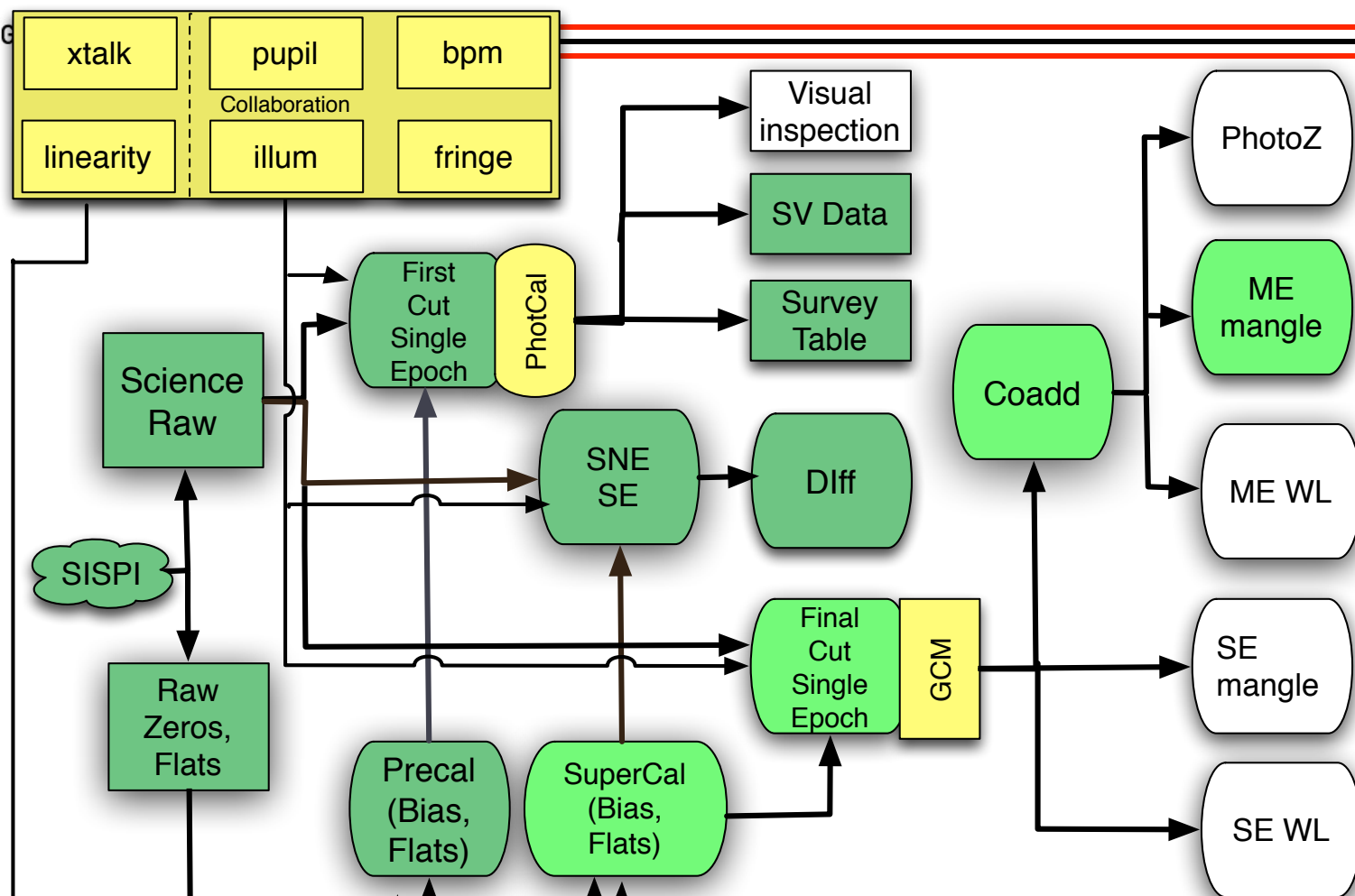
NGC 1365 in the Fornax Cluster

Dark Energy Survey - Dec 13



Pipeline Status

DARK ENERGY
SURVEY



10/15/13

Dark Energy Survey -- CHEP13



Data Ingest from CTIO

DARK ENERGY
SURVEY

- Data from CTIO are transmitted to NCSA over the normal packet-switched research networks.
- Transfers are moderated by the NOAO Data Transport System, which manages bandwidth from the observatory.
- Bandwidth is allocated for the typical one-minute cadence for science exposures. These data arrive and are ingested within minutes.
- Calibration exposures occur before dusk, on a faster cadence prior to dusk.

10/15/13

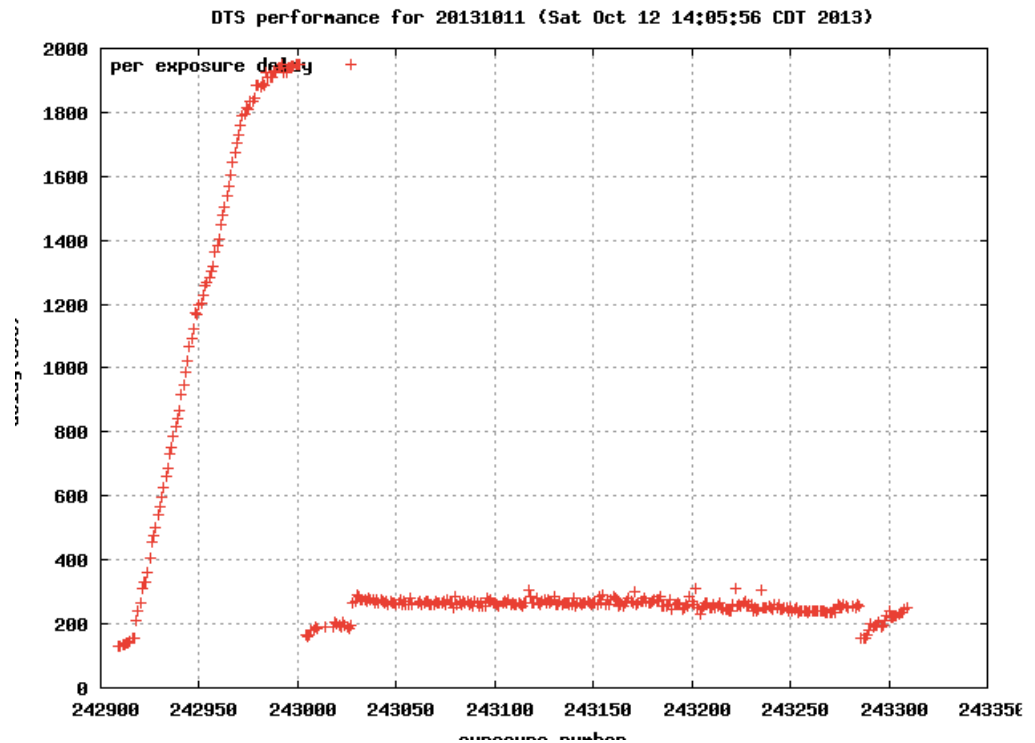


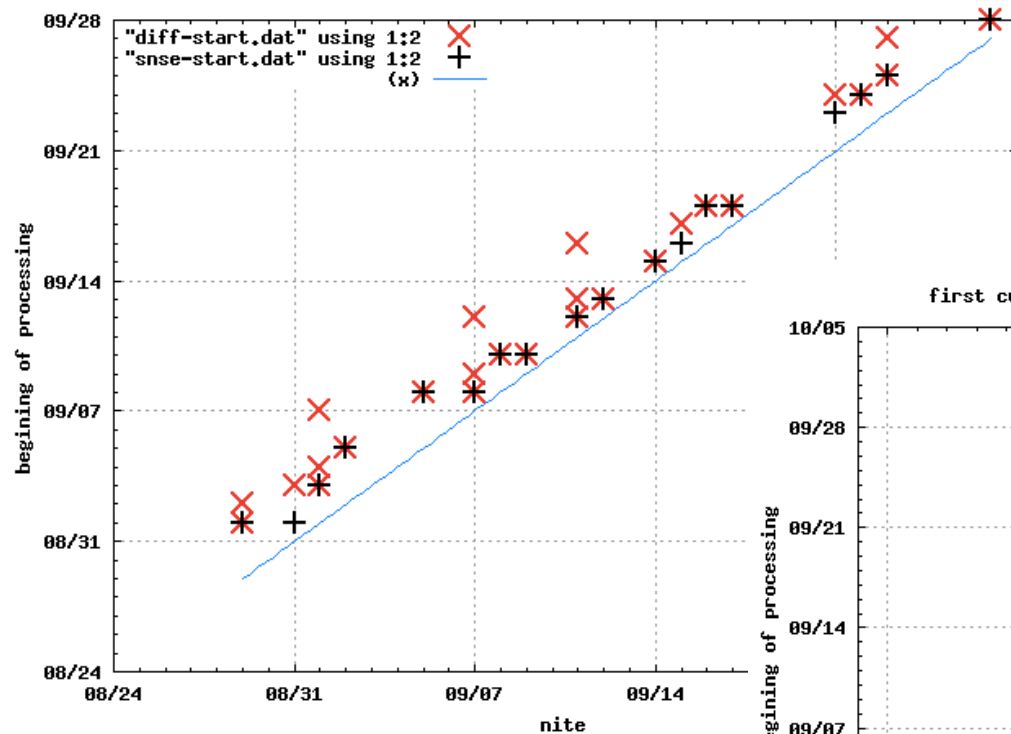
Figure -- Number of seconds between exposure and ingest at NCSA. The initial sequence are flats, bias and darks. Science exposures arrive and are ingested within 3 minutes.



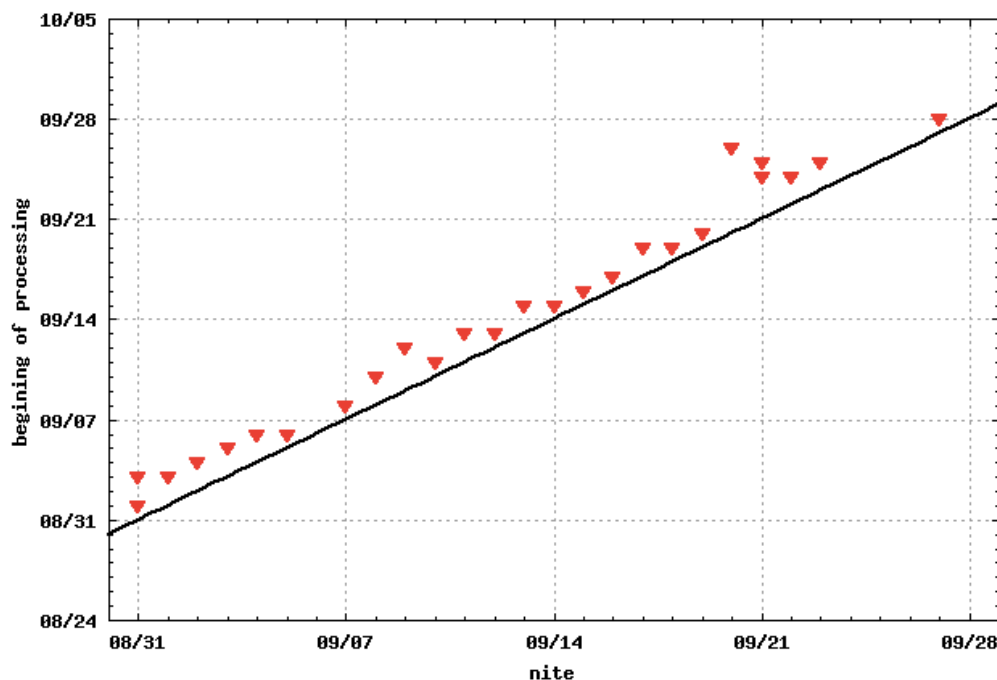
Quick Turn-around processing for SNE and image verification

DARK ENERGY

SESNE and diff processing v.s nite updated Sun Sep 29 01:06:32 CDT 2013



first cut processing v.s nite updated Sun Sep 29 01:02:29 CDT 2013



10/15/13

Dark Energy Survey -- CHEP13



Common Support Components

DARK ENERGY
SURVEY

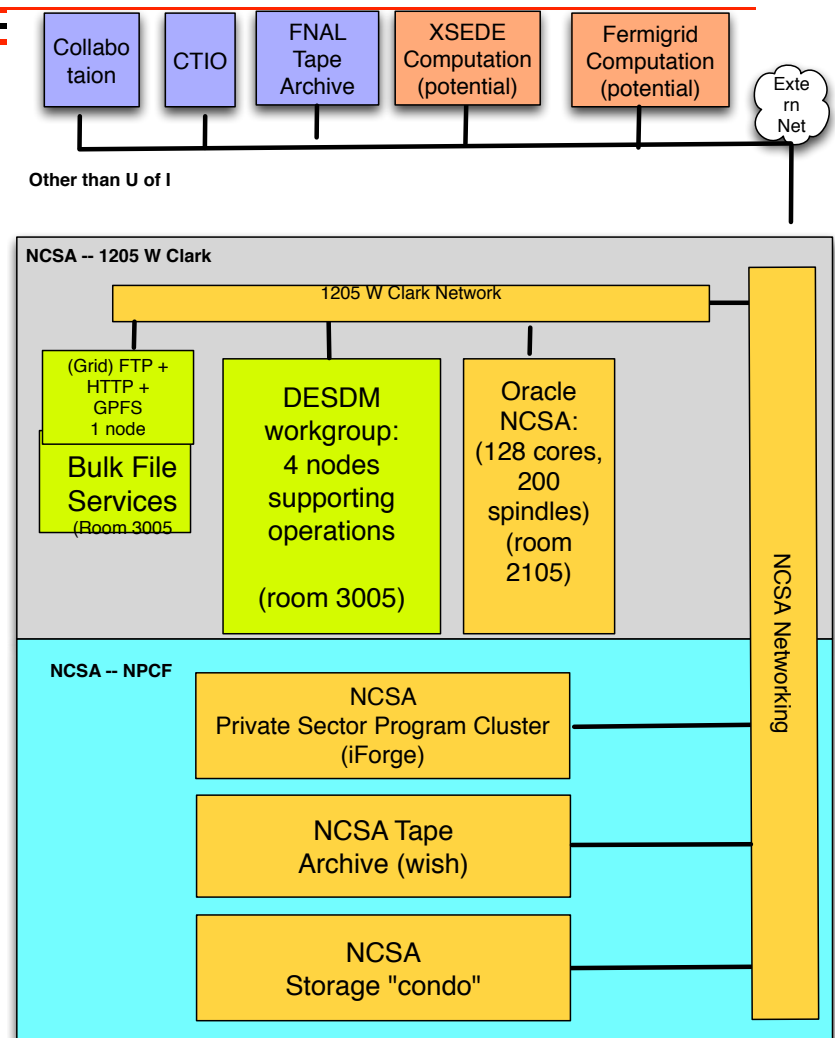
DESDM's direction is to work to use and help develop reusable services for its platforms wherever possible. Its vision is to make **NCSA excellent and experienced in hosting large collaborations.**

Successes

- Spinning archive is on the NCSA storage "Condo"
- Primary computing is 18 AMD nodes in the PSP cluster.
- 100TB Oracle Cluster is in a shared Oracle RAC

Desiderata

- Identify management support



10/15/13

Dark Energy Survey -- CHEP13



DARK ENERGY
SURVEY

Software integration

- Are implementing changes to our workflow system, adopting a philosophy that codes are “hostile”.
- DES makes substantial use of community codes, especially the AstrOmatic suite of codes.
 - Such codes are not bespoke for DES, and do not comply to a framework
 - Integration of specific modules are via FITS format files
 - an old, but highly accepted, file format standard
 - Standardized at the byte level, many independent IO packages.
- A good number of scientist know these codes.
 - Expect the workflow system to record the command line and configuration of these codes.
 - Not for example, invocation of a DES specific wrapper.

10/15/13

Dark Energy Survey -- CHEP13



Reduced Coupling Increased Coherence.

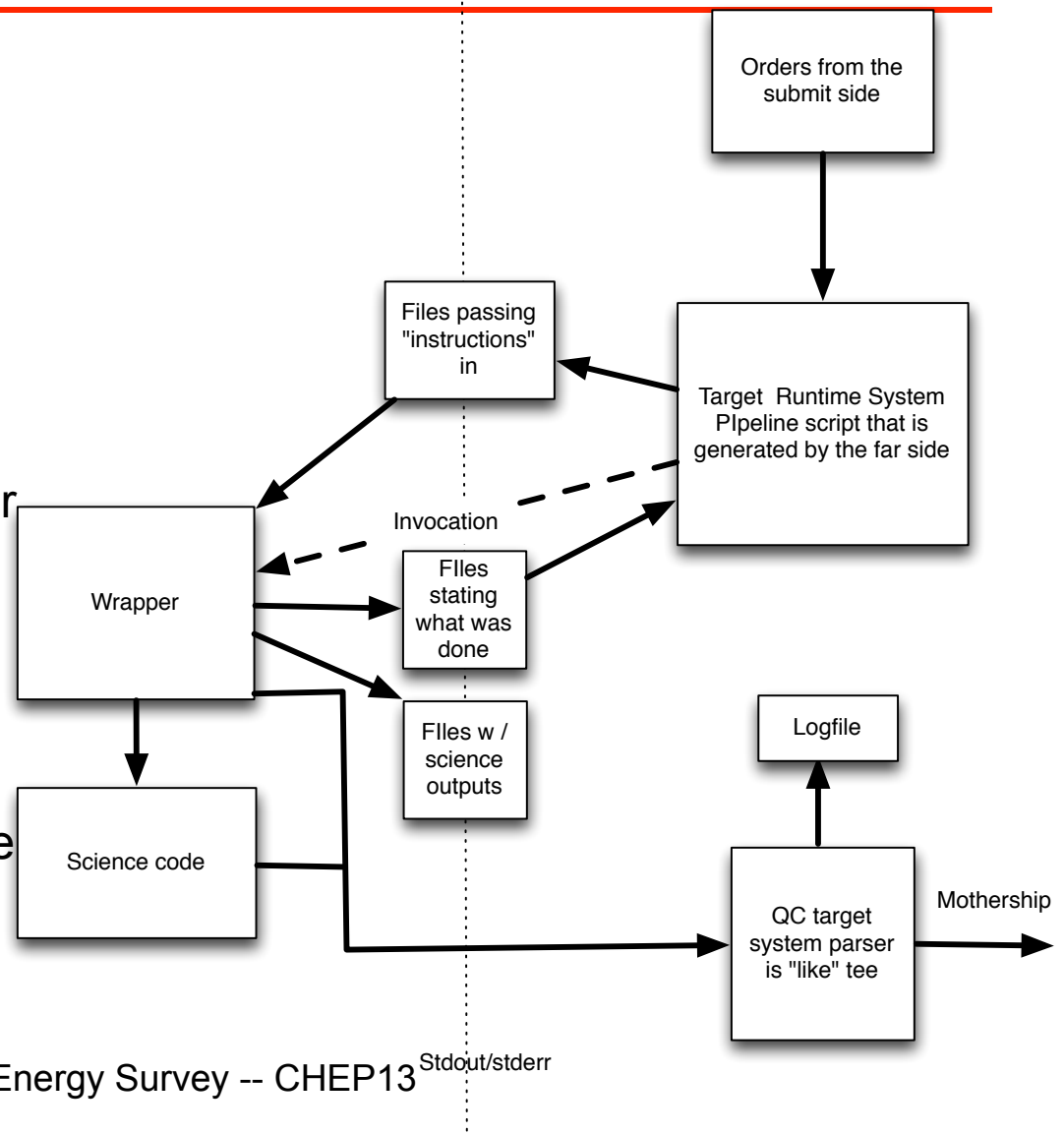
DARK ENERGY
SURVEY

Architectural approach divides the run time system into three main components:

- Science Code
- Wrapper
- Runtime framework

Aside from the benefits of factoring. The project will deliver benefits in:

- Operability,
- Unit running,
- Adequate provenance
- Elimination of “replica catalogs” and other immature internals
- Potential to Run on OSG.



10/15/13

Dark Energy Survey -- CHEP13



DARK ENERGY
SURVEY

Provenance Improvements.

- The provenance model conforms to the Open Provenance Model (OPM) with simplifications that remove components that DES is unlikely to use.
- The database schema supporting the model is based upon the schema proposed in *Storing, Reasoning, and Querying OPM-Compliant Scientific Workflow Provenance Using Relational Databases* (Chunhyeok Lim et. Al), again with simplifications that remove certain features.
- A prominent feature is the use of recursive queries.
 - This gives immediate utility in our relational database without export to a specialized provenance query engine.



Example Provenance Query

- What files were used as inputs when generating D00158794_Y_c23_r4p01_scampcat.fits

```
> select used.filename from used,wgb where wgb.filename='D00158794_Y_c23_r4p01_scampcat.fits'
```

FILENAME
sex.nnw
sex.param_scamp
D00158794_Y_c23_r4p01_detrend.fits
sex.conv
sexforscamp.config

- Get command line for execution that produced D00158794_Y_c23_r4p01_scampcat.fits

```
> select exec_name,cmdargs from wgb where filename='D00158794_Y_c23_r4p01_scampcat.fits'
```

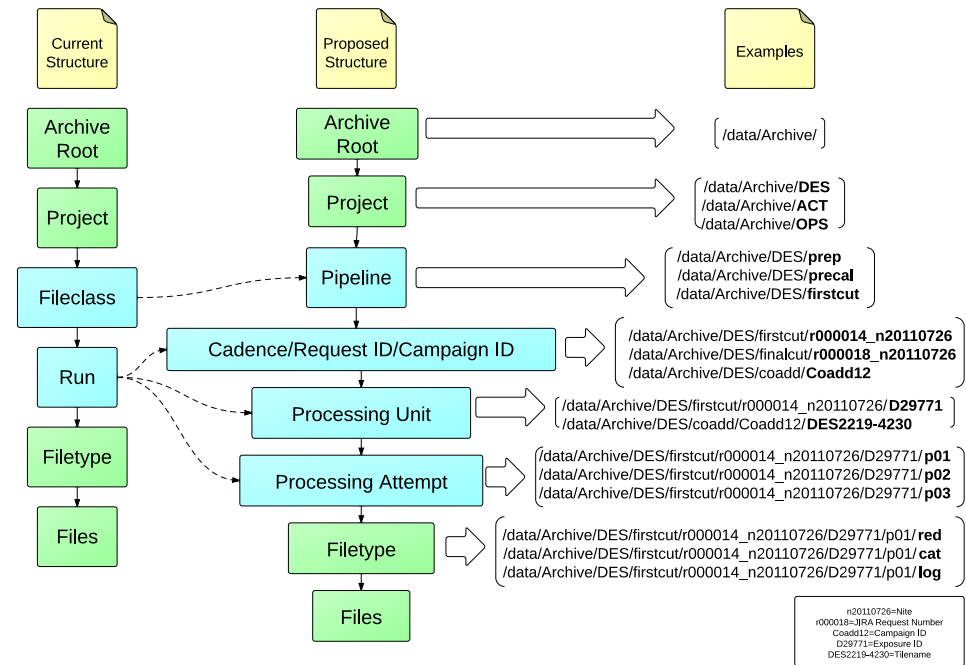
EXEC_NAME	CMDARGS
sex	'red/D00158794_Y_c23_r4p01_detrend.fits[0]' -CATALOG_NAME 'cat/D00158794_Y_c23_r4p01_scampcat.fits' -DETECT_THRESH '2.5000' -FILTER_NAME 'config/sex.conv' -FLAG_IMAGE 'red/D00158794_Y_c23_r4p01_scampcat.fits' -SATUR_LEVEL '47770.0' -STARNNW_NAME 'config/sex.nnw' -WEIGHT_IMAGE 'red/D00158794_Y_c23_r4p01_scampcat.fits' -WEIGHT_TYPE 'MAP_WEIGHT' -c 'config/sexforscamp.config'



Naming

DARK ENERGY
SURVEY

- Changes to the file tree will allow operation to quickly access file for debug.
- New Canonical Tree will scale at each directory level.
- Uniquely named files facilitate QA.
- Example of coupling: Current pathnames.



Single Epoch:

Camera ExposureID Band CCD RequestID Attempt Type
D29771_g_c04_r000014p01_detrend.fits

10/15/13

Dark Energy Survey -- CHEP13



DARK ENERGY
SURVEY

Data Distribution Mechanisms

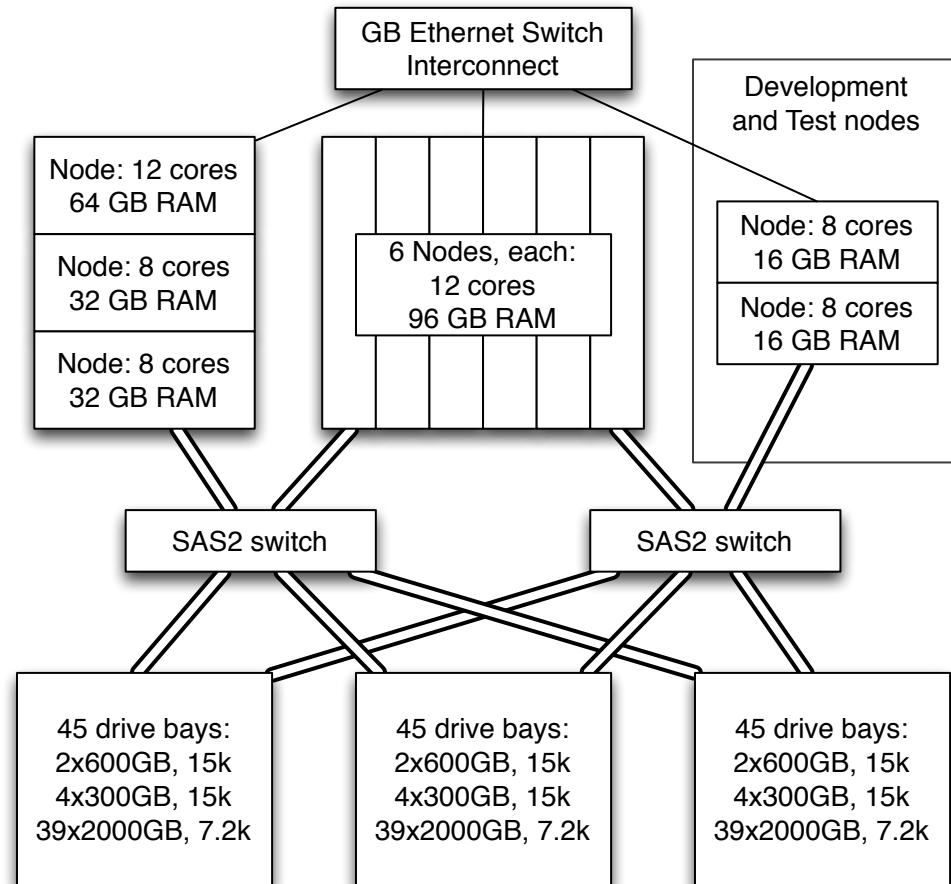
- Philosophy – simple tools anyone can use.
- DB –
 - User accounts in Oracle DB.
 - Each user has own schema, can see common schema.
 - Operations and Science Schemas.
 - Science schema are optimized for read access, queries are transparently parallelized.
- Files –
 - The main access is https
 - Have already been thanked for the slowdown due to the (s).
 - Have prototype code to redirect to http via temporary links.



DARK ENERGY
SURVEY

Database Cluster

- 11 Nodes
- 128 cores
- 736GB RAM
- 135 drives
- 241TB storage
- 6GB SAS2 storage fabric



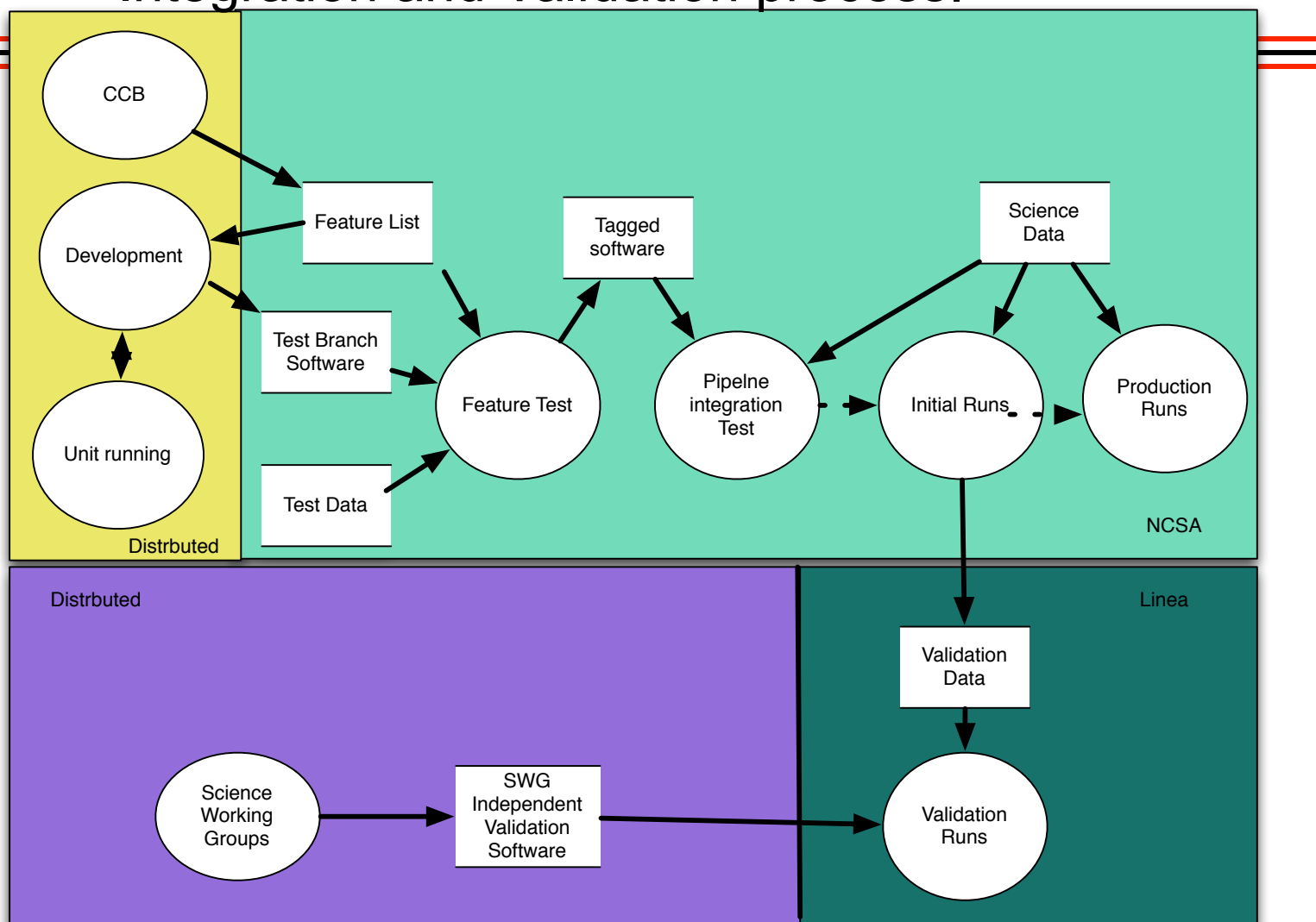
10/15/13

Dark Energy Survey -- CHEP13



DARK ENERGY
SURVEY

New Structure supports this Development Integration and Validation process.

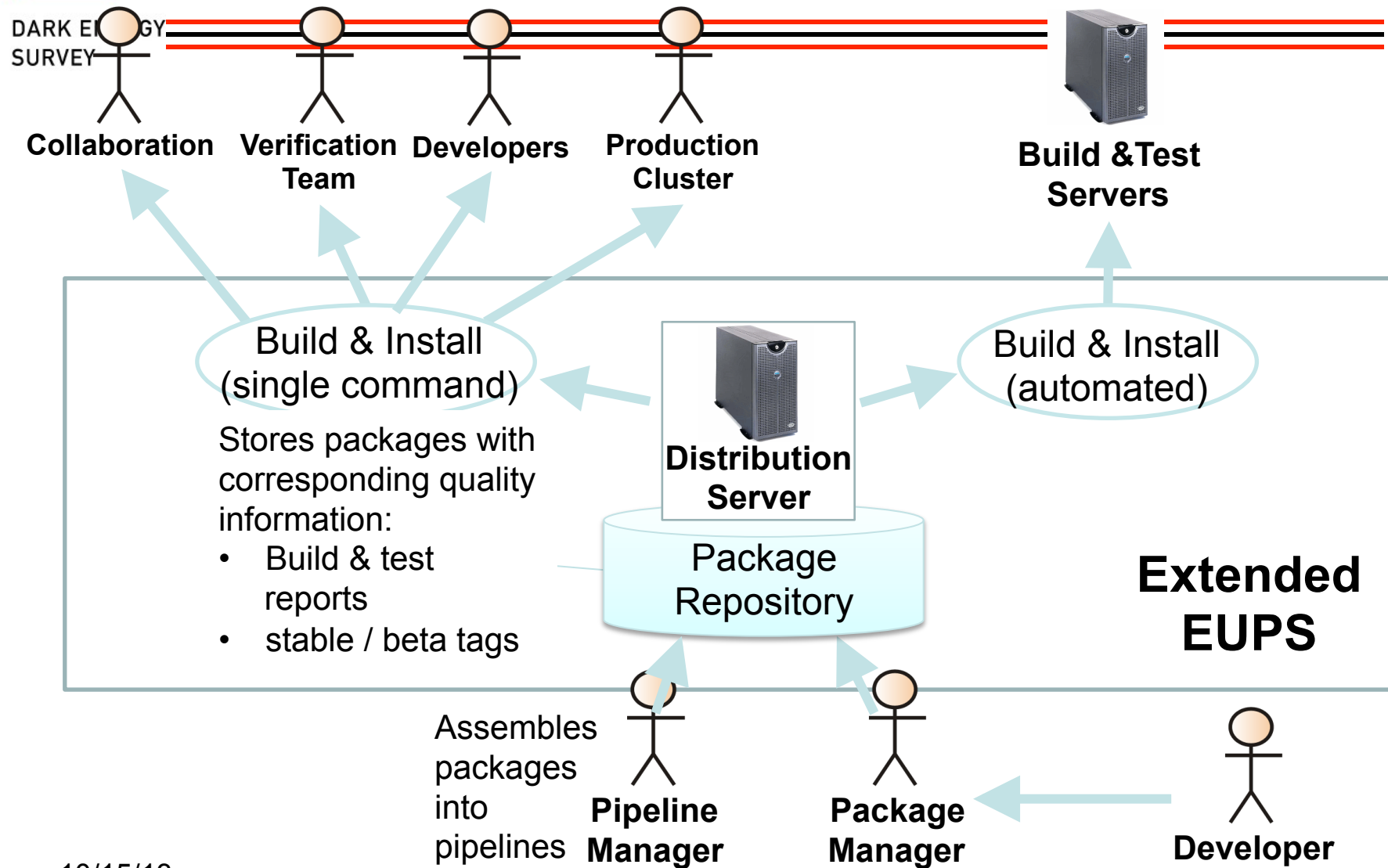


10/15/13

Dark Energy Survey -- CHEP13



Build / Distribute / Install



10/15/13

Dark Energy Survey -- CHEP13



DARK ENERGY
SURVEY

Summary

- DES data taking has begun!
- Data processing system
 - Produced a release based on an initial science verification year.
 - Processes data daily for SNE and L0 data verification.
 - Defined code and data QA processes.
- A set of upgrades are in progress to help the system scale for 5 years of operation. Upgrades include
 - Provenance system based on the OPM and literature.
 - Workflow based on integrating community codes communicating via files.
 - Scalable file structure.