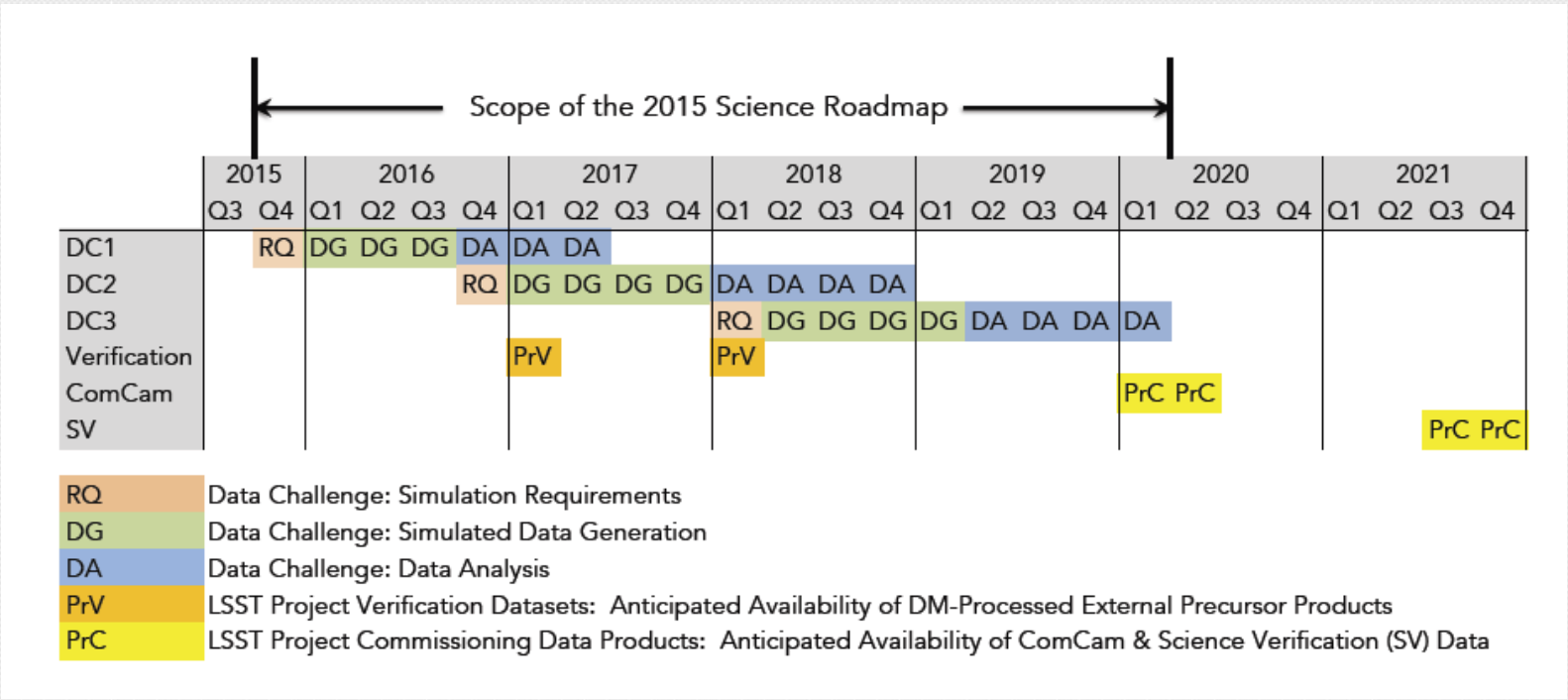
DESC++ Simulations

Andrew Connolly (UW), Julian Borrill, Salman Habib, Alina Kiessling, Rachel Mandelbaum, Peter Nugent, Michael Schneider

Driving development with data challenges



DESC will undertake a series of progressively more complex data challenges leading up to ComCam at the end of 2019

- Cosmology simulations are core to many of these challenges
- Data challenges don't need to be over arching or large but do need to be properly motivated for the science/analysis/systematic they will address

Data challenges

Table 3.1.1: DC datasets with image data by challenge and working group.

DC	WGs	Name	Input	Catalog area	Cadence[a]	Image Area	Bands	Depth
DC1	CL/LSS	`HaloCat 1.0`	N-body		N/A	N/A	N/A	N/A
DC1	SL/SN	`Twinkles 1.0`	CATSIM (T)[b]	100 sq arcmin	DDF 10 yr	100 sq arcmin	ri	r=27.0
DC1	LSS/WL	`DC1 Phosim Deep`	CATSIM	80 sq deg	WFD 10 yr	80 sq deg	r	r=28.0
DC2	SL/SN	`Twinkles 2.0`	CATSIM (T)[b]	100 sq arcmin	DDF 10 yr	100 sq arcmin	ugrizy	i=24.0
DC2	SL/WL/LSS CWG/CL/PZ TJP	`DC2 Mock Lightcone`	CATSIM (N)[c]	300 sq deg	WFD 10 yr	100 sq deg	ugrizy	r=27.5
DC3	WL/LSS CL/PZ/TJP CWG	`DC3 Mock Lightcone`	CATSIM (NN)[c]	18000 sq deg[d]	WFD 10 yr	3000 sq deg	ugrizy	r=27.5
DC3	SL/WL/LSS	`DC3 Mock ComCam Survey`	CATSIM (NN)[c]	1000 sq deg	WFD 6 month	1000 sq deg	ugrizy	r=27.5

DC3: 3-5 Gpc^3 10^8-10^9 M_o (TBD)

Dataset Generation

- Data Challenge DC1 (2016) 5M CPU/hrs
  - **HaloCat**: 3 Gpc DM simulation (in hand)
  - **Twinkles 1.0**: 1 CCD 800 visits (ri bands)
  - **Phosim Deep**: several 80 sq degree regions 5x1300 visits (r band)
- Data Challenge DC2 (2017) 20M CPU/hrs
  - **Twinkles 2.0:** 1 CCD 2500 visits (ugrizy bands)
  - **DC2 Mock Light cone:** 300 sq degrees of mock catalogs, 100 sq degrees of imaging (ugrizy)
  - Used to define the tools for large analyses
- Data Challenge DC3 (2018) 200M CPU/hrs
  - 18,000 sq degrees of mock catalogs, 3000 sq degrees of imaging (ugrizy)
  - 100x increase in volume and will be a major challenge
  - TBD

Generation of Cosmological Simulations

- Mock catalogs (CS1, CS2, CS3)
  - **RQ: Which parameters need simulating and how (shape, size, …). What mocks should we be generating (volumes, mass resolution, …)**
  - SW: Ability to integrate “any” simulation into a catalog framework
  - VA: Define how we validate a mock catalog (against parameterization or existing data sets) and apply this to current mocks
- Validation
  - **How do we validate a mock catalog (against existing data sets or precursor knowledge). Critical component to the simulations.**
  - Prediction tools (in collaboration with TJP)
  - Tools to predict mass function, power spectrum, correlation function etc (e.g. emulators).
- New simulations (hydro and the inclusion of systematics)
  - Hydro simulations, systematics simulations and non LCDM models

Tasks for simulations

LSST Dark Energy Science Collaboration /... / Cosmology Simulations

Key Tasks for SRM

Edit Watch Share ...

Created by Karl Krughoff, last modified by Katrin Heitmann on Sep 07, 2016

Key Projects/Task	Due date	Description	People working on this	Input needed from other WGs	Existing / completed work	Status
CS1		**Standardize project wide definitions for galaxy observables**	@Adrian Pope, @Nelson Padilla, @Eve Kovacs, @Joanne Cohn, @Risa H. Wechsler, @Danila Korytov		Github CS1	
CS1.1.1	12/15	Schedule and conduct a first meeting of interested parties to identify the specific quantities that need to be standardized.		all		DONE
CS1.1.2	12/15	Assemble all standard methods for calculating identified properties.	@Adrian Pope			ONGOING
CS1.1.3	12/15	Deliver first version of document outlining galaxy properties of interest (Confluence?).				ONGOING
CS1.2.1	06/16	Research measurement methods for the galaxy properties delivered in CS1.1 to determine the optimal measurement technique for each one.				WAITING
CS1.2.2	09/16	Produce a library to execute the necessary measurements so that there is a reference implementation for the project.				WAITING
CS1.2.3	09/16	Produce a formal document hosted in a version control system (github) that describes the measurement techniques and outlines the relevant decision making points.				WAITING

http://ls.st/e53

Common simulations…

- Computational resources (Habib et al 2016)
  - Current projections of available HPC resources (2015)
    - 100-500M core hours for individual projects
    - Primarily DOE and NSF resources
  - Current projections of available HPC resources (2020)
    - 1B core hours hours for individual projects
    - Primarily DOE resources
    - HEP resources available at 100B to 1T core hours at LCFs and NERSC

Experimental Data	2013	2020	2030+
Storage	1PB	6PB	100-1500PB
Cores	10^3	70K	300+K
CPU hours	$3x10^6$ hrs	2×10^8 hrs	$\sim 10^9$ hrs
Simulations	2013	2020	2030+
Storage	1-10 PB	10-100PB	> 100PB - 1EB
Cores	0.1-1M	10-100M	> 1G
CPU hours	200M	>20G	> 100G

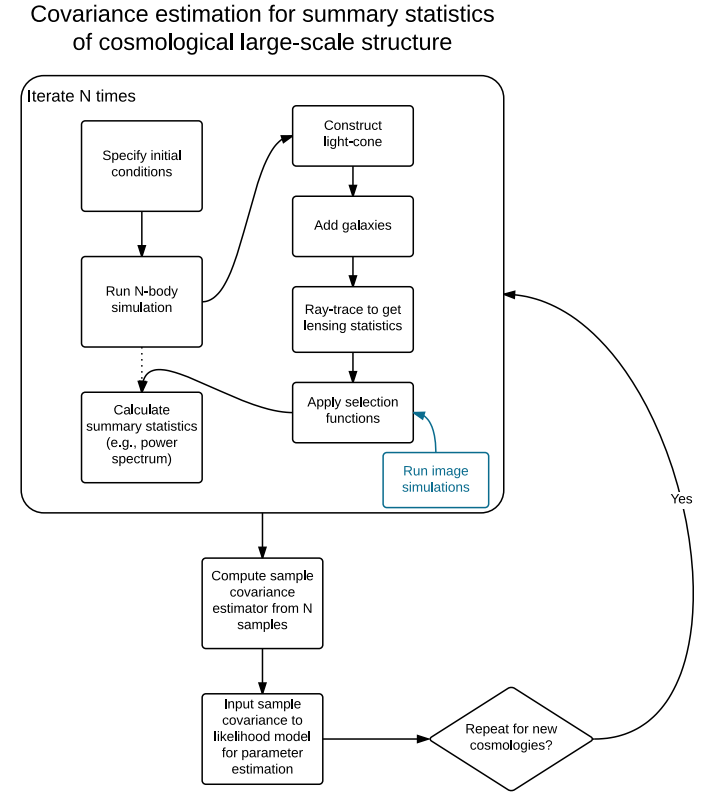
Year	Requested Allocation (M CPU-hrs)
2017	20
2018	150
2019	150
2020	300
2021	500

Common simulations needs

- Computational needs
  - Resources available for the scale of simulations required (with possible exception of Hydrodynamical simulations)
- **Infrastructure needs**
  - Sharing simulations (e.g. Millennium DB) and *mock catalogs* has greatly expanded the use of simulations in a broad range of applications
  - At 100 TB to 1 PB scales this infrastructure doesn't exist
  - Transfer of data will be a challenge (10 days for 100 TB at 1 Gb)
- Data exchange formats
  - Standard in observational datasets is FITS. No standard for cosmology (as groups adopt their own standards).
  - HDF5 is a possibility (really just a container class)
- **Mock catalogs**
  - Great progress in the development of mocks from SHAMs, SAMs, hydro
  - Still no next generation mock catalogs (or agreement on attributes or what they mean)

Primary drivers

- Covariances
  - Covariance matrices generated from series of mock galaxy catalogs (and associated numerical simulations). Brute force approach feasible in about 2025 (given computational projections) but significant advance due to *new statistical approaches* (the advantage of slacking)
- Baryonic physics
  - Astrophysical feedback mechanisms due to baryons not well understood.



Not just cosmological simulations