1.09 - Data Management

Julian Borrill (he/him)
for the DM team

CMB-S4 Collaboration Meeting, May 9-13 2022

Scope

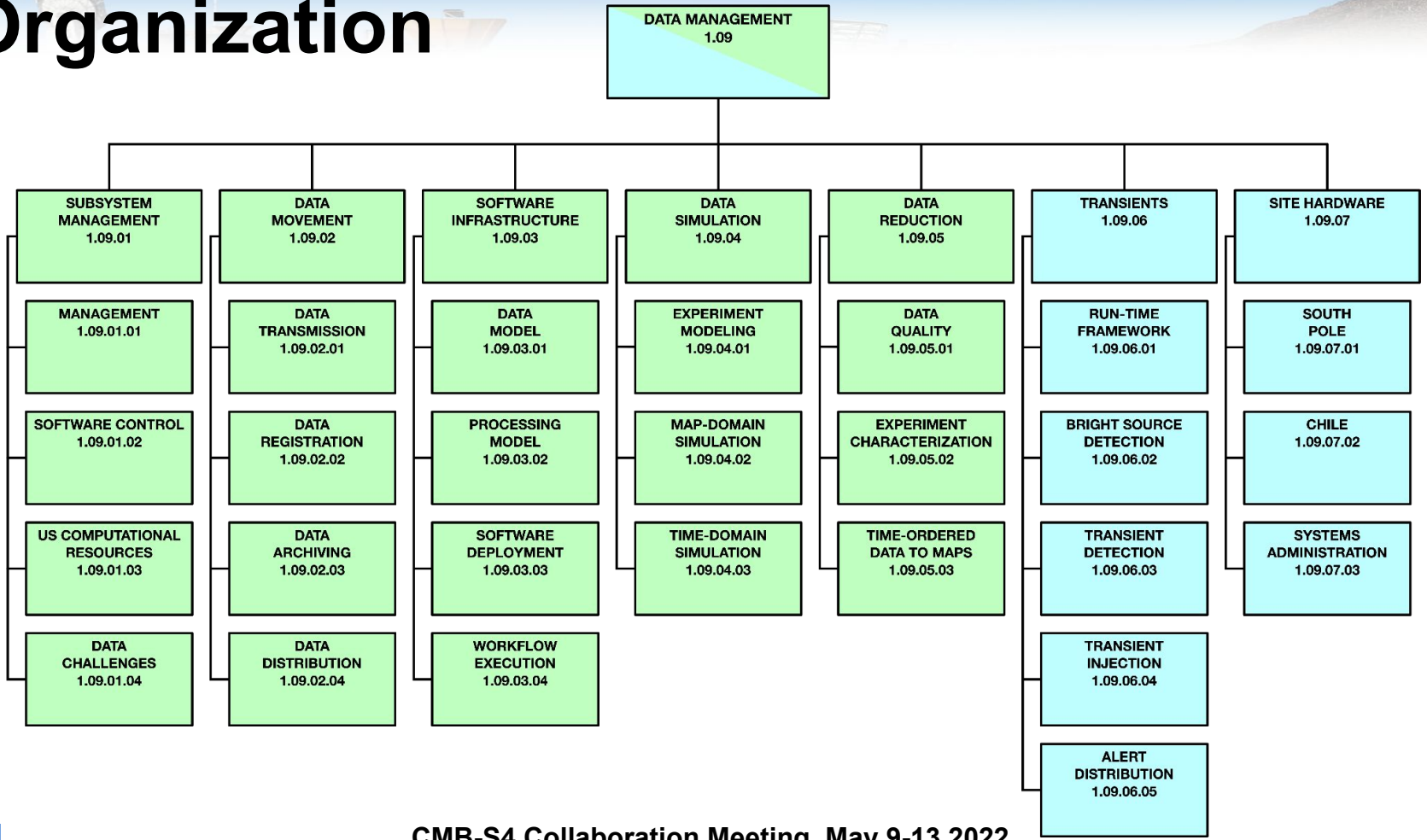
In operations Data Management will:

- Receive raw telescope data from Data Acquisition and secure it
- Conduct per-observation data quality checks and issue timely alerts
- Detect per-observation transients and issue timely alerts
- Deliver science-grade single-frequency maps (and ancillary data products) to the collaboration AWGs for science analyses
- Distribute all data products and the software used to produce them to the community

In construction Data Management does:

- Develop and deploy all systems needed for operations prior to the start of commissioning
- Use these systems to support experiment design validation and verification

Organization



Leadership Team

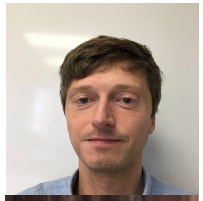
Intentionally distributed & representative team - Pole/Chile, lab/university, cosmological/computational, …

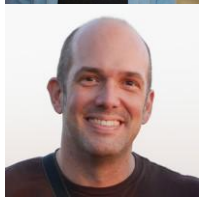
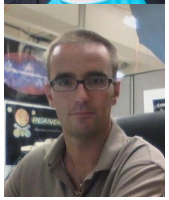
1.09 DATA MANAGEMENT

Julian Borrill (LBNL)
*Kevin Long** (LBNL)


1.09.01 DM Management	1.09.02 Data Movement	1.09.03 Software Infrastructure	1.09.04 Data Simulation	1.09.05 Data Reduction	1.09.06 Transients	1.09.07 Site Hardware
*Kevin Long** (LBNL)	Eli Dart (ESNet) tbd	Debbie Bard (NERSC) Ted Kisner (LBNL)	Sara Simon (FNAL) Andrea Zonca (UCSD)	Colin Bischoff (Cincinnati U.) Reijo Keskitalo (LBNL)	Don Petravick (UIUC) Nathan Whitehorn (MSU)	Tom Crawford (U. Chicago) Sasha Rahlin (FNAL)



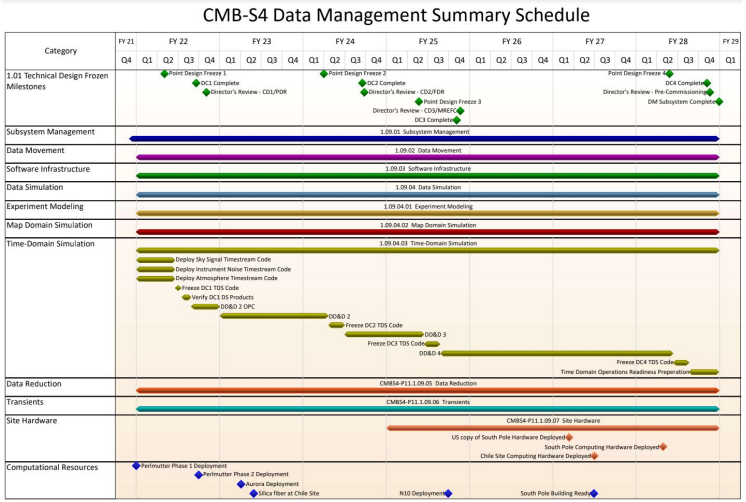





Technical Highlights

- Project Systems
- Data Challenge 1
- Analysis of Alternatives
- Transients & FABRIC

Project Systems

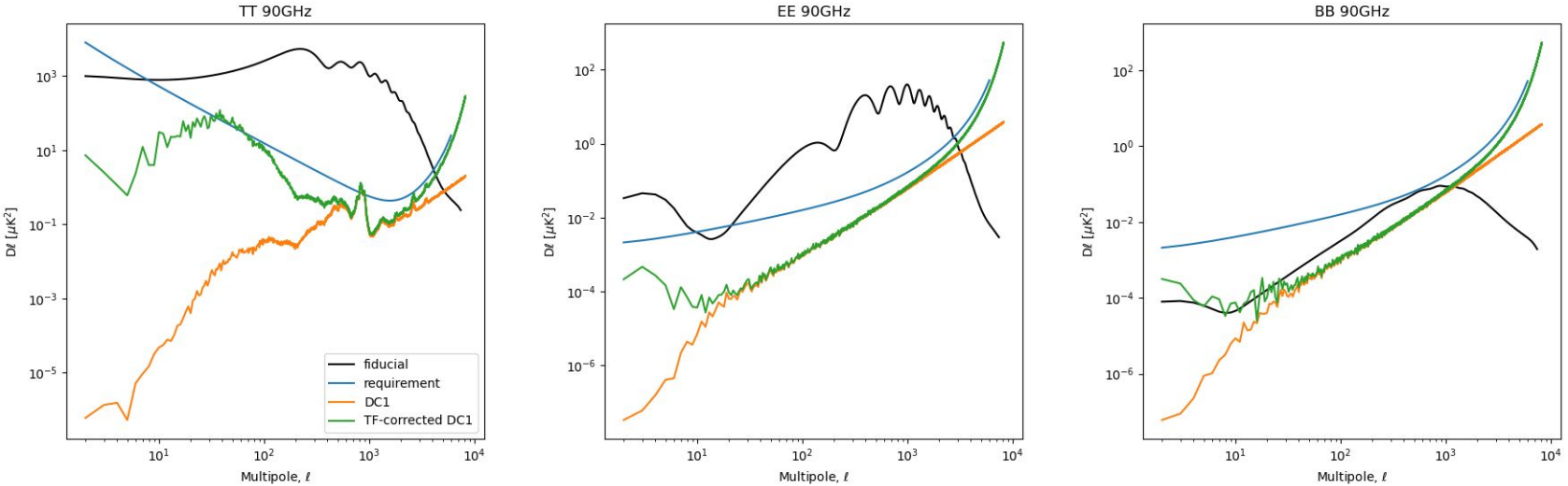
- Dash360: rebuilt the DM cost/schedule (as recommended by the fall Director's Review) to integrate the Data Challenges into the overall workflow
  - Interleaved 6-month Data Challenge and extended Design/Develop/Deploy phases.
  - DD&D to be populated at each phase - see DC2 session on Thursday.



- JAMA: redefined the DM project requirements to be more explicitly tied to the science requirements and to support flowdown through the WBS.
- JIRA: reset the DM project risks/opportunities to address South Pole bandwidth consistently and incorporate mitigations for all risks.

Data Challenge 1

- See DC1 session on Thursday.
- Ongoing activity - Chilean LAT data completed and available.
- Provides validation that the PBD meets the measurement requirements.



Analysis of Alternatives

- See Chile Alternatives & South Pole Alternatives sessions on Wednesday.
- Internal DM issues:
  - Site hardware in Chile to meet requirements of any additional telescopes (trivial)
  - Site hardware at South Pole to fit within power constraints (very complicated)
  - US computational resources (mostly trivial)
- DM supporting roles
  - Chile survey strategies
  - Chile observing efficiencies

Transients & FABRIC

- SPT transient detection pipeline being redeployed as a starting point for CMB-S4.
  - Transients to be injected into DC1 maps and then detected; tests for false positives and negatives
  - Obvious dependency on any revision/extension of transients science requirements.
- Working to deploy transient detection pipeline on FABRIC hardware as a test-case for us both
  - FABRIC deploys in-network computing resources that could facilitate timely alerts by performing detection while the raw data are in transit from Chile.

Plans For FY22

- Complete DC1
- Continue to support Analysis of Alternatives
- Inject/detect transients into DC1 data, including using FABRIC hardware
- Continue planning next DD&D phase, towards DC2
- Provide training