



CMB-S4

Eli Dart

LBNL & ESnet

Data Management L3 Lead for Data Movement

Credits

- This presentation contains the work of many people
- Most of the material is from others
- I will do my best to represent the project here, but science questions should be directed to the scientists (which I would be happy to do)
- Thank you for the opportunity to present CMB-S4 to you today!

What Is CMB-S4?

- The 4th generation ground-based CMB experiment.
- The 1st ground-based CMB *project*:
 - Designed to meet critical science thresholds, not to do the best we can under a particular budget cap.
 - Can't fail, not best effort.
- Making the full scope of CMB science available to the entire community:
 - Using the best technologies & techniques of all previous experiments.
 - Making the full scope of CMB science available to the whole community.
- Planned as a joint DOE (HEP) and NSF (Astronomy + Physics + Polar Programs) project:
 - Adding DOE capacities and capabilities to the longstanding NSF program.
 - Enabling unprecedented scaling (10x any previous experiment).

History

- 2013 - CMB community converges around CMB-S4 in Snowmass process.
- 2014 - P5 recommends CMB-S4 “under all budget scenarios”.
- 2015 - First CMB-S4 workshop held; biannually ever since.
- 2015 - NAS identifies CMB as one of 3 strategic Antarctic science priorities.
- 2016 - AAAC convenes the CMB-S4 Concept Definition Taskforce.
- 2017 - AAAC unanimously accepts the CDT report.
- 2018 - The CMB-S4 collaboration adopts its bylaws and is officially formed.
- 2019 - DOE takes CD-0 identifying the need for CMB-S4; NSF provides pre-project funding through U Chicago to develop the preliminary design.
- 2020 - CMB-S4 is TRACEd by Astro2020; DOE selects LBNL as the project lead lab.

Primary Science Goals

GOAL 1: Test models of inflation by measuring or putting upper limits on r , the ratio of tensor fluctuations to scalar fluctuations.

GOAL 2: Determine the role of light relic particles in fundamental physics, and in the structure and evolution of the Universe.

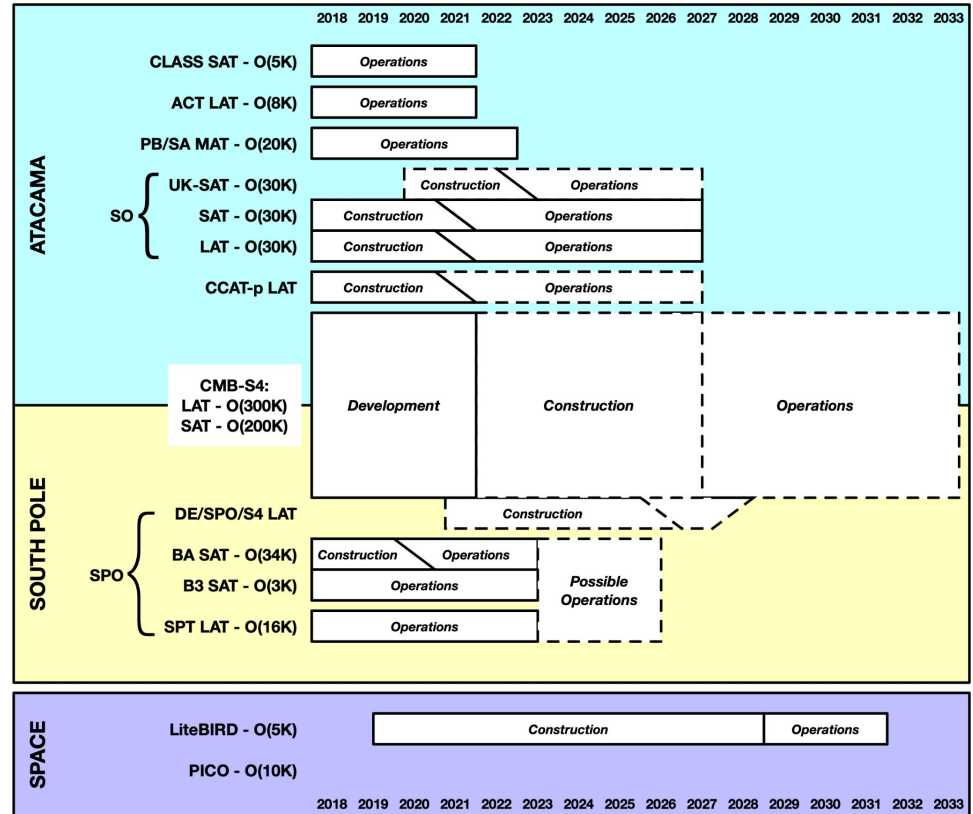
GOAL 3: Measure the emergence of galaxy clusters as we know them today. Quantify the formation and evolution of the $z \geq 2$ clusters and intracluster medium during this crucial period in galaxy formation.

GOAL 4: Explore the mm-wave transient sky and measure the rate of transients. Use the rate of mm-wave GRBs to constrain their mechanisms. Provide mm-wave variability and polarization measurements for stars and active galactic nuclei.

Meeting these goals will enable a wealth of other CMB/mm-wave science.

Context

US CMB landscape in the 2020s
as of 2018 (pre-COVID)

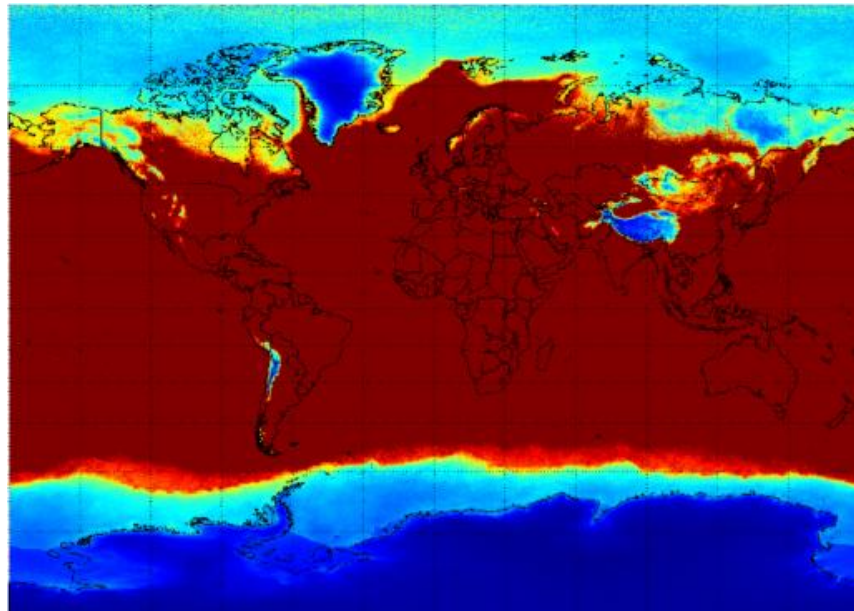


Experiment Design

Science Goal Design Parameter	Inflation	Light Relics	Galaxy Clusters	Transients
Map Depth (Detector-Years)	Ultra-Deep	Deep	Deep	Deep
Sky Area (Sites, Survey Strategy)	Small	Large	Large	Large
Angular Resolution (Mirror Size & Quality)	Low + Moderate	Moderate	High	High
Observing Cadence (Survey Strategy)	-	-	-	Daily
Frequency Coverage (Sites, Bandpasses)	Wide	Moderate	Moderate	Moderate

Sites

- Ground-based CMB observations are limited by the atmosphere: we need high, dry, sites.
- The South Pole and Chilean Atacama are the highest, driest sites.
- The US CMB community has a long history of working at both, and significant infrastructure is already in place for CMB-S4 precursors (South Pole Observatory; Simons Observatory & CCAT-prime)

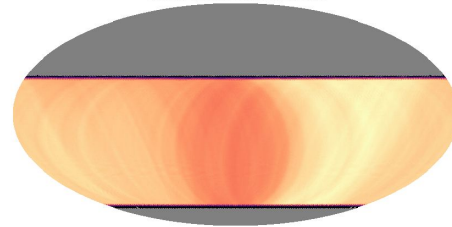


Mean precipitable water vapor across the globe. Candidate sites (dark blue) are the South Pole, Chilean and Argentinian Atacama Desert, Tibetan Plateau & Greenland.

Survey Strategies

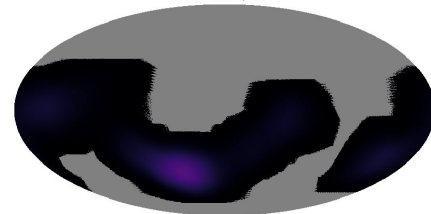
- CMB-S4 is unique in having *two* exceptional observing sites available.
- The biggest difference between the sites is in the types of sky surveys their latitudes can support.
 - Wide-area surveys can only be performed from the Atacama.
 - Compact ultra-deep surveys can only be performed from the South Pole.

Chile LAT modulated high cadence hits



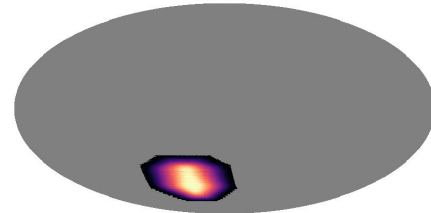
Chile wide survey hitmap

Chile SAT deep hits



Chile deep survey hitmap

Pole SAT deep hits



South Pole ultra-deep survey hitmap

Telescopes

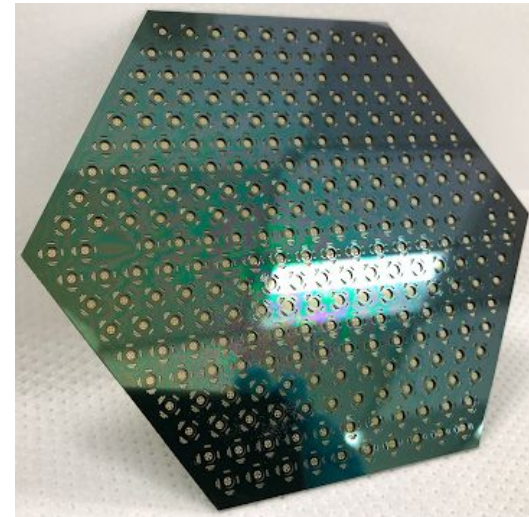
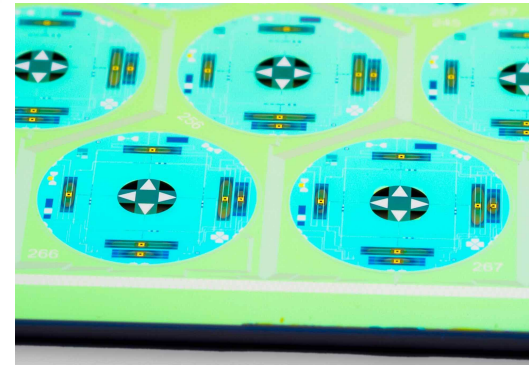
- Large Aperture Telescopes
 - 2 x 6m segmented mirror in Chile
 - 1 x 5m monolithic mirror at South Pole
- Small Aperture Telescopes
 - 6 x 3 x 0.5m at South Pole
 - Possible to relocate to Chile



Map Depths

- Detectors
 - 500,000 cryogenically-cooled superconducting transition edge sensors
 - 125,000 dual-polarization dichroic pixels
 - 500 wafers.
- Years
 - 7-year observation duration for all surveys

Survey	Detectors	Detector-Years
SAT	150K	1100K
South Pole LAT	115K	800K
Chile LATs	245K	1700K



Timeline

- Construction project: 2019-27 (Astro2020 & federal budget permitting)
- Staggered deployment across both sites: 2027-29
- Operations: 2029-36

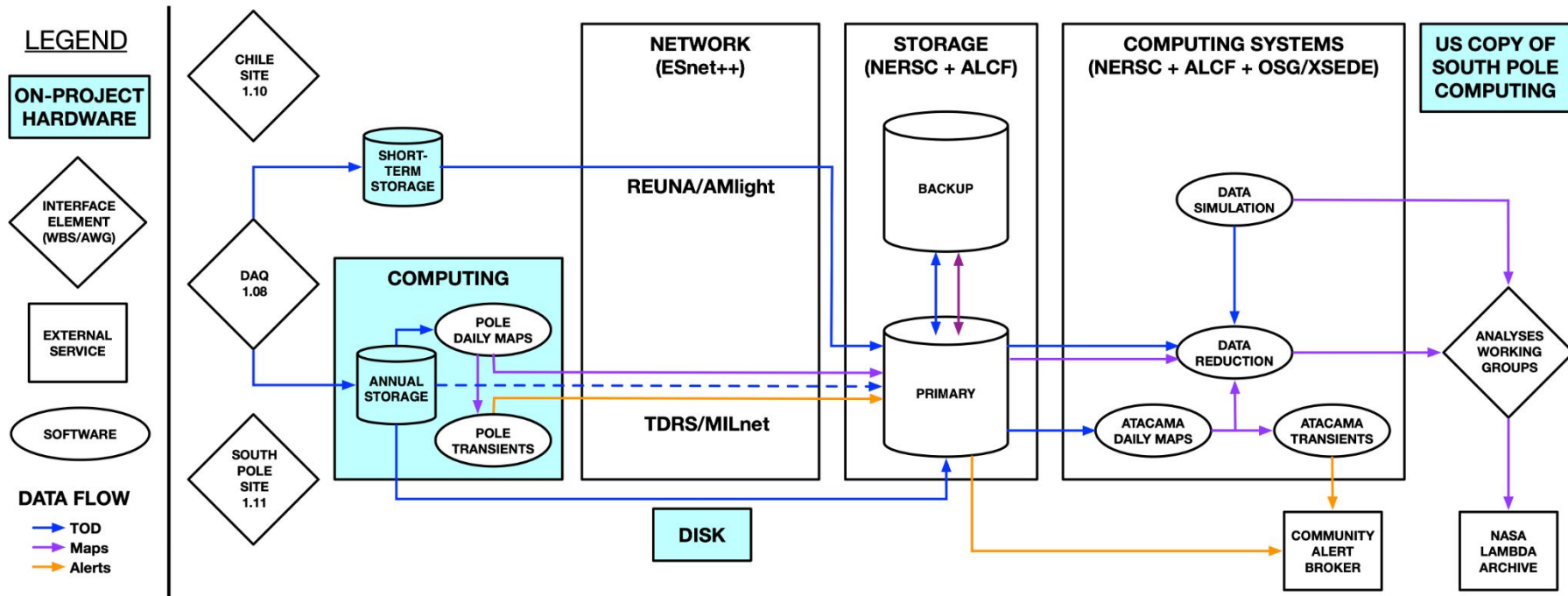
Natural progression:

- Late 10s: 4 x single-site, single-aperture (ACT, BICEP, POLARBEAR, SPT)
- Early 20s: 2 x single-site, dual-aperture (SO, SPO)
- Late 20s: 1 x dual-site, dual-aperture (CMB-S4)

COVID-19 Pandemic has caused some delays

- Working with funding agencies to determine new schedules
- Project is continuing to move forward as best we can

Data Management Schematic



Note: named resources are anticipated, not confirmed.

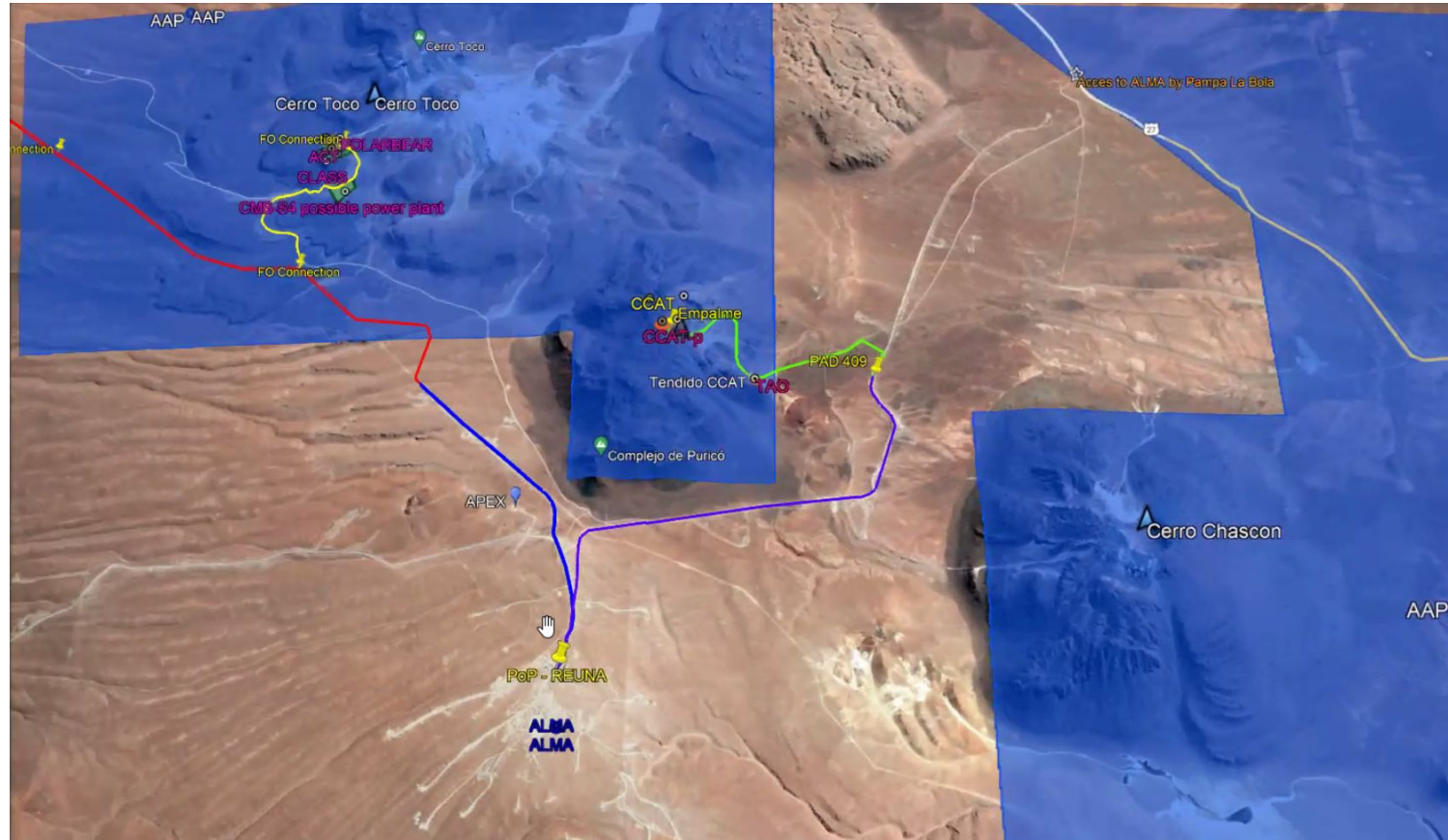
Atacama Site: Networking

- Compressed data rate ~1.2 Gbps
- Real-time data transfer to US data center (NERSC)
 - Transient alert analysis may be performed in transit on FABRIC nodes
- Scoping up to 1 month of on-site storage ~400TB
 - With 10 Gbps available, 4 days to clear a month-long backlog.
- Working closely with Simons Observatory to coordinate site networking
 - Eli Dart of ESnet as CMB-S4 Atacama data movement lead.

Cerro Toco Fiber Paths

Fiber optic connection being built from REUNA at ALMA to Simons Obs. and CMB-S4

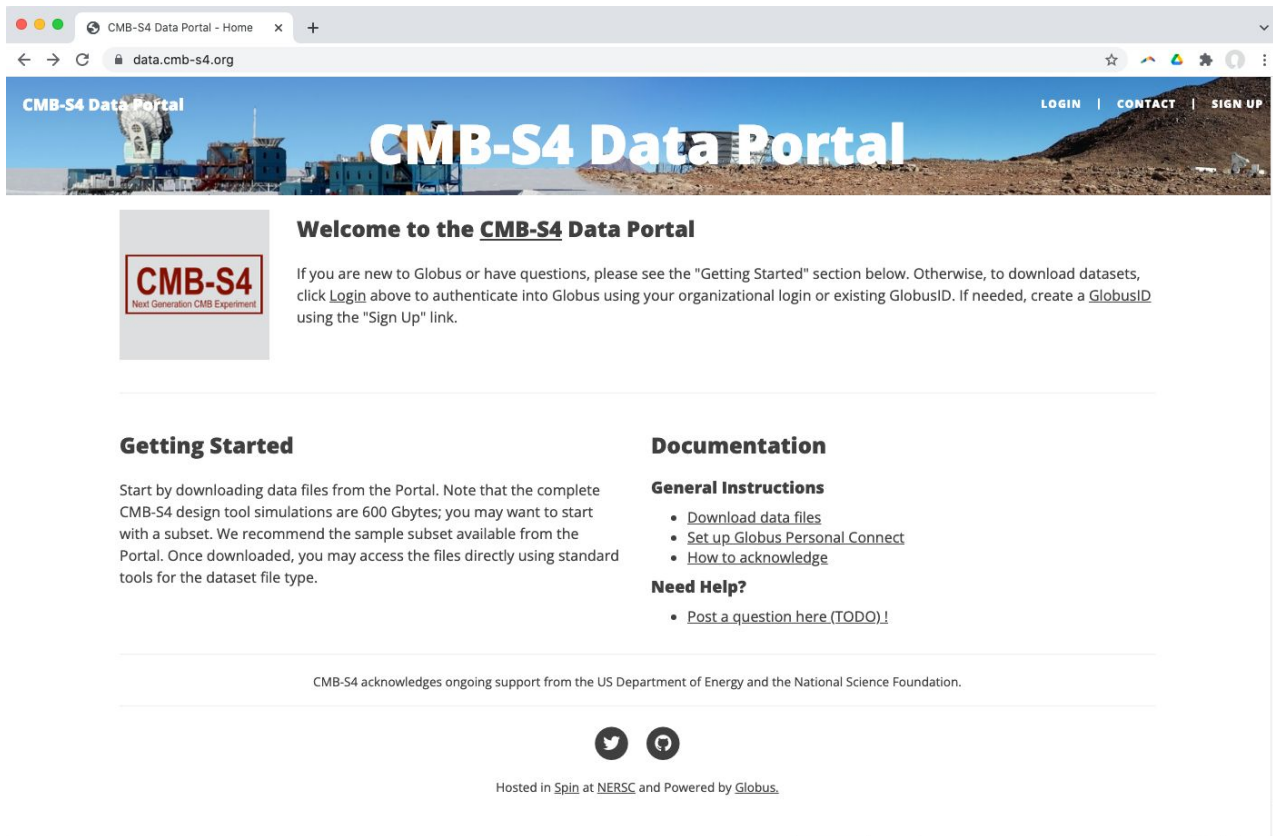
- Currently being sited in collaboration with Simons
- Collaboration with REUNA on services design



Prototyping: Data Distribution

- NERSC's Modern Research Data Portal:
 - Tool based on the Globus platform
 - Well documented, simple customization/initialization
 - Current interest from other collaborations (e.g. LZ)
 - Used by the LSST-DESC collaboration, among others
- Prototype CMB-S4 data portal
 - <https://data.cmb-s4.org/>
 - Preliminary data distribution service for Data Challenge 1

Prototyping: Data Distribution



The screenshot shows a web browser window displaying the CMB-S4 Data Portal. The browser's address bar shows the URL `data.cmb-s4.org`. The website has a blue header with the text "CMB-S4 Data Portal" and navigation links for "LOGIN", "CONTACT", and "SIGN UP". Below the header is a large banner image of a desert landscape with a radio telescope. The main content area features a "Welcome to the CMB-S4 Data Portal" message, a "Getting Started" section with instructions on downloading data, a "Documentation" section with links to "General Instructions", "Download data files", "Set up Globus Personal Connect", and "How to acknowledge", and a "Need Help?" section with a link to "Post a question here (TODO)!". At the bottom, there is a footer with social media icons for Twitter and GitHub, and a statement: "Hosted in Spin at NERSC and Powered by Globus."

CMB-S4 Data Portal - Home x +

data.cmb-s4.org

CMB-S4 Data Portal

LOGIN | CONTACT | SIGN UP

CMB-S4 Data Portal

Welcome to the CMB-S4 Data Portal

If you are new to Globus or have questions, please see the "Getting Started" section below. Otherwise, to download datasets, click [Login](#) above to authenticate into Globus using your organizational login or existing GlobusID. If needed, create a [GlobusID](#) using the "Sign Up" link.

Getting Started

Start by downloading data files from the Portal. Note that the complete CMB-S4 design tool simulations are 600 Gbytes; you may want to start with a subset. We recommend the sample subset available from the Portal. Once downloaded, you may access the files directly using standard tools for the dataset file type.

Documentation

General Instructions

- [Download data files](#)
- [Set up Globus Personal Connect](#)
- [How to acknowledge](#)

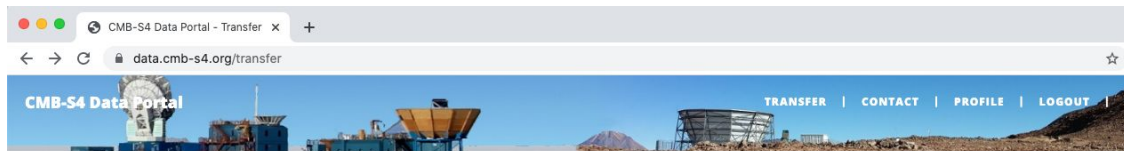
Need Help?

- [Post a question here \(TODO\)!](#)

CMB-S4 acknowledges ongoing support from the US Department of Energy and the National Science Foundation.

Hosted in Spin at NERSC and Powered by Globus.

Prototyping: Data Distribution












Datasets

Choose the dataset(s) you wish to download, then click the **Transfer** button to begin transferring the full dataset(s) from NERSC. **Or**, you can click on a catalog dataset to browse and choose individual files for transfer (not supported for image datasets). See [here](#) for detailed instructions.

To learn more about a dataset, click the "doc" link next to the dataset name.

Dataset filter: ☒ Design tool simulations 1 (DT1) ☒ Data Challenge 1 (DC1) ☐ Deprecated

Dataset Name	Size	Select
 DT1: Lensed CMB scalar only (doc)	30 GB	<input type="checkbox"/>
 DT1: CMB tensor only $r=3e-3$ (doc)	30 GB	<input type="checkbox"/>
 DT1: Foregrounds (doc)	30 GB	<input type="checkbox"/>
 DT1: Noise only no splits (doc)	31 GB	<input type="checkbox"/>
 DT1: Noise + Atmosphere, full and 7 splits, LAT Chile (doc)	172 GB	<input type="checkbox"/>
 DT1: Noise + Atmosphere, full and 7 splits, LAT Pole (doc)	200 GB	<input type="checkbox"/>
 DT1: Noise + Atmosphere, full and 7 splits, SAT Pole (doc)	121 GB	<input type="checkbox"/>
 S4 only DT1: Noise only no splits (doc)	31 GB	<input type="checkbox"/>
 DM only DT1: Atmosphere (doc)	31 GB	<input type="checkbox"/>

Transfer

Data Transfer From NERSC To ALCF

- Data transfer from NERSC to ALCF is important to CMB-S4 (backup data center)
- NERSC and ALCF high-speed data transfer clusters work well (8.2GB/sec or 65Gbps, almost 30TB/hour)
- data.cmb-s4.org portal uses different NERSC endpoint - runs at 1.1GB/sec, 8.8Gbps, or about 4TB/hour
 - Performance is adequate for current needs
 - Globus automation and other features make this easy from a people-effort perspective

ACTIVITY LIST ✓ CMB-S4 public data to alcf#dtn_theta
transfer completed

Overview | Event Log

Task Label	CMB-S4 public data to alcf#dtn_theta
Source	► CMB-S4 public data
Destination	► alcf#dtn_theta
Task ID	4cab966-0b84-11ec-bf0e-edb00af5aa74
Owner	Eli Dart (dart@globusid.org)
Condition	SUCCEEDED
Requested	9/1/2021, 05:25 PM
Completed	9/1/2021, 05:35 PM
Duration	9 minutes 34 seconds
Transfer Settings	<ul style="list-style-type: none">• transfer is not encrypted• overwriting all files on destination

358 Files

7 Directories

654.42 GB Bytes Transferred

1.14 GB/s Effective Speed

0 Skipped files on sync

0 Skipped files on error

[View debug data](#)

FILE MANAGER ✓ NERSC DTN to alcf#dtn_theta
transfer completed

Overview | Event Log

Task Label	NERSC DTN to alcf#dtn_theta
Source	NERSC DTN ⓘ
Destination	alcf#dtn_theta ⓘ
Task ID	a18670b4-fa0c-11eb-832c-f56dd2959cb8
Owner	Eli Dart (dart@globusid.org)
Condition	SUCCEEDED
Requested	2021-08-10 11:56 am
Completed	2021-08-10 12:05 pm
Duration	9 minutes
Transfer Settings	<ul style="list-style-type: none">• verify file integrity after transfer• transfer is not encrypted• overwriting all files on destination

19260 Files

211 Directories

4.44 TB Bytes Transferred

8.24 GB/s Effective Speed

0 Skipped files on sync

0 Skipped files on error

[View debug data](#)