

SPT-3G Computing

Benedikt Riedel^{1,*}, *Lincoln Bryant*¹, *John Carlstrom*^{1,2}, *Thomas Crawford*², *Robert W. Gardner Jr.*¹, *Nicholas Harrington*³, *Nicholas Huang*³, *Alexandra Rahlin*⁴, *Judith Stephen*¹, and *Nathan Whitehorn*⁵

¹Enrico Fermi Institute, University of Chicago, 933 E. 56th Street, Chicago, IL 60637

²Kavli Institute for Cosmological Physics, University of Chicago, 5640 S Ellis Ave, Chicago, IL 60637

³Department of Physics, University of California, Berkeley, 366 LeConte Hall MC 7300, Berkeley, CA, 94720

⁴Fermi National Accelerator Laboratory, PO Box 500, Batavia IL 60510

⁵Department of Physics and Astronomy, University of California, Los Angeles, 475 Portola Plaza, Los Angeles, CA 90095

Abstract. SPT-3G, the third generation camera on the South Pole Telescope (SPT), was deployed in the 2016-2017 Austral summer season. The SPT is a 10-meter telescope located at the geographic South Pole and designed for observations in the millimeter-wave and submillimeter-wave regions of the electromagnetic spectrum. The SPT is primarily used to study the cosmic microwave background (CMB). The upgraded camera produces an order of magnitude more data than the previous generations of SPT cameras. The telescope is expected to collect a petabyte (PB) of data over course of five years, which is a significantly larger data volume than any other CMB telescope in operation. The increase in data rate required radical changes to the SPT computing model both at the South Pole and University of Chicago. This paper will describe the overall integration of distributed storage and compute resources into a common interface, deployment of on-site data reduction and storage infrastructure, and the usage of the Open Science Grid (OSG) by the SPT collaboration.

1 Introduction

The South Pole Telescope (SPT) [1, 2], see Figure 1, is used to observe the cosmic microwave background (CMB) to uncover important features of our Universe and the physics that govern it. The SPT is a 10-meter telescope located at the National Science Foundation (NSF) Amundsen-Scott South Pole station, the best site on Earth for microwave observations, and is optimized for sensitive, high-resolution measurements of the CMB. It is funded jointly by NSF and the Department of Energy (DOE).

Since the construction and first light of SPT in 2007, the SPT collaboration has completed two surveys of the CMB in South Polar sky:

- SPT-SZ: 2500-square-degree survey (2007-2011) [3]
- SPTpol: 500-square-degree survey (2012-2016) [4]

*e-mail: briedel@uchicago.edu



Figure 1. South Pole Telescope. Credit: NSF/SPT

Besides the survey area, the main difference between the surveys was the detector array; the field in general has been characterized by rapidly increasing channel count and additional information (polarization, in particular).

The SPT-SZ and SPTpol observations have led to groundbreaking results that have moved the field of CMB research forward in significant ways. These results include the first galaxy clusters discovered using the Sunyaev-Zel'dovich (SZ) effect [5] and the first detection of the elusive “B-mode” pattern in the polarization of the CMB [6].

The third-generation camera for SPT, SPT-3G, [7] was deployed during Austral summer 2016-17 (first light January 30, 2017) and delivers a large improvement in sensitivity over the already impressive SPT-SZ and SPTpol surveys. This increase in sensitivity comes from two technological advances:

- Improved wide-field optical design that allows more than twice as many optical elements in the focal plane, and
- Detector pixels sensitive to multiple observing bands in a single detector element.

The sensitivity of the SPT-3G receiver will lead to precise constraints on the sum of the neutrino masses and potentially deliver a detection of the primordial B-mode signal from a background of gravitational waves from the epoch of inflation.

2 Computing Requirements

The significant advances in sensitivity delivered by the SPT-3G receiver come primarily from increasing the number of detectors at the focal plane of the telescope by an order of magnitude. With this comes a concomitant increase in the requirements for data storage and computing needs. For a five-year observation time, an estimated 1.2 PB of storage and 150M CPU hours are required. This is a far greater storage and computing need than any other CMB observatory currently in operation. The SPT collaboration engaged with the Open Science Grid (OSG) [8] group at University of Chicago to maintain data analysis and storage infrastructure at both the South Pole and at the University of Chicago for the SPT-3G collaboration. The collaboration has decided to utilize a high throughput computing model, a first for a CMB experiment.

3 South Pole Computing

The main tasks for the South Pole computing infrastructure are:

- Perform first-pass data analysis and lossy compression for transfer over satellite, and
- Store full-fidelity data set for retrieval during the Austral summer.

The lossy compression to transfer science data over satellite is necessary, as there are only 125 GB per day allocated to SPT for data transfer on the South Pole Tracking and Data Relay Satellite System Relay (SPTR) [9]. The experiment produces roughly 400 GB per day of data. The remaining data is stored in full-fidelity at the pole, such that it can be retrieved by collaboration members during the austral summer.

At the South Pole, OSG staff deployed new computing infrastructure during the Austral summer 2016-17. The new hardware consists of six servers and two storage chassis.

A single storage controller node for the two storage chassis. One of the servers acts as a hypervisor that hosts central services such as Domain Name System, a Network File System (NFS) server, a login node, and a Puppet [10] server. The three remaining servers are utilized as worker nodes for an HTCondor [11] cluster. All machines of the same model have the same hardware configuration, such that they can act as a hot spare in case one of the machines catastrophically fails. Similarly, the HTCondor pool servers have fairly large local storage that can be used as replacement disks in the storage chassis

The storage chassis are used to provide a large storage pool for data analysis at the South Pole and as a data store for data that cannot be transferred via satellite. The large storage pool is managed using ZFS [12] and exported to the servers using NFS. The other storage chassis is configured as a "just a bunch of disks" (JBOD). This is done to allow for data retrieval at the end of every Austral summer, i.e. disks filled with data are replaced with new disks for the upcoming season of data-taking. The current plan is to swap out the required hard drives every season and replace them with previously used hard drives from the previous observation year.

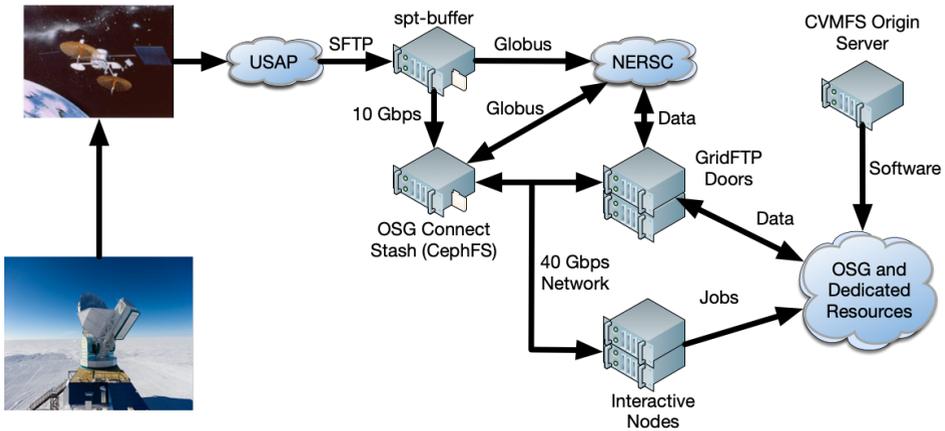


Figure 2. Overview of data and job flow for SPT. Credit: NSF/SPT/NASA

4 Northern Hemisphere Computing

The computing infrastructure in the northern hemisphere consists of two pieces: data management and data analysis. The data transfer from the South Pole is handled through the infrastructure provided by the United States Antarctic Program (USAP). The daily SPT data transfer is retrieved from USAP’s servers in Denver, CO to a dedicated server at the University of Chicago. From this server, the data is added to OSG Stash, a Ceph [13]-based filesystem located at the University of Chicago and available to users of the OSG, and replicated to the High Performance Storage System (HPSS) tape archival system at DOE’s National Energy Research Scientific Computing Center (NERSC) at the Lawrence Berkeley National Laboratory [14] using Globus Online [15] file transfer service, see Figure 2.

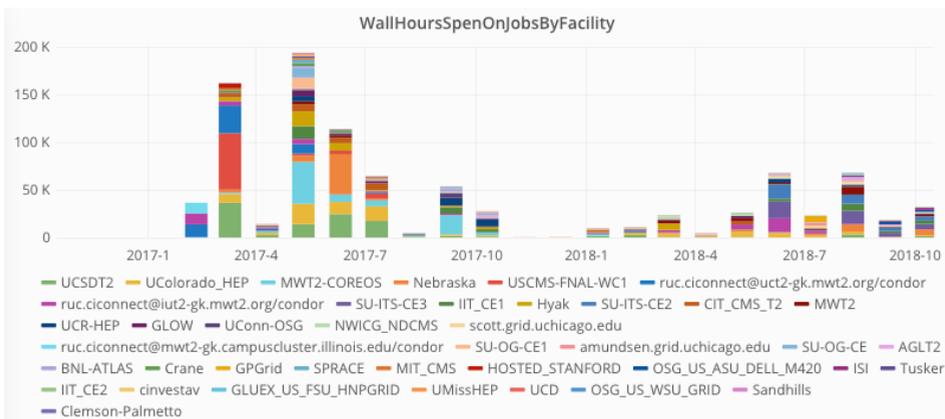


Figure 3. Summary of monthly Wall Hours used by SPT across the OSG.

The University of Chicago OSG group set up two dedicated servers to allow the collaboration to perform interactive data analysis and submit large data reduction pipelines to the opportunistic computing pool on OSG, see Figure 3 for OSG usage. The collaboration

has yet to fully meet its computing requirements as the experiment has been having start-up difficulties. In addition to the servers, we deployed a copy of the SPT-3G software dependencies across the OSG and on the two dedicated nodes using the CERN-VM Filesystem (CVMFS) [16]. To allow for interactive data analysis, we have deployed a JupyterHub [17] instance on each server. The users can access the SPT data on OSG workers nodes using GridFTP [18] from OSG Stash.

To supplement the opportunistic computing pool provided by OSG, the University of Chicago staff utilized Virtual Clusters for Community Computation [19] project to add dedicated resources from campus clusters to the SPT computing pool. In production, we have connected the Hoffman2 [20] cluster at University of California, Los Angeles to the SPT resource pool. We are currently testing connecting the University of Chicago's Midway cluster [21] and the NERSC's Cori [22] to the resource pool.

5 Conclusion

We have described the computing workflow for the SPT 3G telescope both on site and at the University of Chicago. SPT using OSG as its predominant computing resource is the first instance in which a CMB telescope has used either OSG or a distributed computing model; the success of this approach paves the way for further use in upcoming challenges for CMB data analysis, in particular the upcoming CMB-S4 experiment.

References

- [1] J. E. Ruhl, et. al., Proc. SPIE Int. Soc. Opt. Eng. **5498**, 11-29 (2004), astro-ph/0411122
- [2] J. E. Carlstrom, et. al., Publ. Astron. Soc. Pac **123**, 568 (2011), astro-ph/0907.4445
- [3] C. L. Chang, et. al., AIP Conf. Proc **1185**, 475-477 (2009)
- [4] J. E. Austermann, et. al., Proc. SPIE Int. Soc. Opt. Eng. **8452**, 84521E (2012), astro-ph/1210.4970
- [5] Z. Staniszewski, et. al., Astrophys. J. **701**, 32-41 (2009), astro-ph/0810.1578
- [6] D. Hanson, et. al., Phys. Rev. Lett. **111**, 141301 (2013), astro-ph/1307.5830
- [7] B. A. Benson, et. al., Proc. SPIE Int. Soc. Opt. Eng. **9153**, 91531P (2014), astro-ph/1407.2973
- [8] R. Pordes, et. al., J. Phys.: Conf. Ser. **78**, 012057 (2007)
- [9] <https://www.usap.gov/technology/contentHandler.cfm?id=1971> [accessed 2019-01-28]
- [10] <https://puppet.com/> [accessed 2019-01-28]
- [11] D. Thain, T. Tannenbaum, and M. Livny, Distributed computing in practice: the Condor experience, Concurrency and Computation: Practice and Experience **17**, 323-356 (2005)
- [12] http://open-zfs.org/wiki/Main_Page [accessed 2019-01-28]
- [13] <https://ceph.com/> [accessed 2019-01-28]
- [14] <http://www.nersc.gov/users/storage-and-file-systems/hpss/> [accessed 2019-01-28]
- [15] I. Foster, IEEE Internet Computing **15**, 70-73 (2011)
- [16] P. Buncic, et. al., J. Phys.: Conf. Ser. **219**, 042003 (2010)
- [17] T. Kluyver, et. al., Positioning and Power in Academic Publishing: Players, Agents and Agendas, 87 - 90, (2016)
- [18] J. Bresnahan, lobus GridFTP: What's New in 2007, GridNets (2007)
- [19] L. Bryant, et. al., PEARC: Practice and Experience in Advanced Research Computing, (2018)
- [20] <https://idre.ucla.edu/hoffman2> [accessed 2019-01-28]

- [21] <https://rcc.uchicago.edu/resources/high-performance-computing> [accessed 2019-01-28]
- [22] <http://www.nersc.gov/users/computational-systems/cori/> [accessed 2019-01-28]