CMBS4-doc-931
DOI 10.5281/zenodo.8248049

CMB-S4 Experience with FABRIC for Data Challenge Zero

G. Daues, F Menanteau, D. Petravick
Center for Astrophysical Surveys
National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign
August, 2023

Abstract:
This technical note describes CMB-S4 experience with research infrastructure provided by the NSF Funded FABRIC project (NSF awards 1935966, 2029260, https://fabric-testbed.net/). We found we were able to validate this platform's architecture type for potential use in CMB-S4 for prompt processing use cases.

Motivation:
While the bulk of the computing for the CMB-S4 experiment can be satisfied by large batch facilities, such as NERSC. There is a need for prompt processing to support daily quality checks and to produce prompt transient directions of millimeter wave astrophysical phenomena for the community. The prompt processing system should be available each day observations are taken, be well connected to networking, and ideally supported as a part of an agency's computing facility to gain economies related to multiple users of the facility.

FABRIC as a model for infrastructure
To prototype such a Facility CMB-S4 has evaluated the NSF-Supported FABRIC research infrastructure as a prompt processing platform in the context of its Data Challenge Zero.

In brief, FABRIC provides a distributed mesh of compute and storage resources, and the ability to interconnect these resources with secure internal wide area virtual network connections. FABRIC supplies an infrastructure-as-software provisioning model which provides the capability to adjust provisioning to work around any local failures, and to scale with load.

Prototyping activities.
Two of us (Daues, Petravick) obtained credentials and established the "CMB-S4 Phase One" experiment within FABRIC. While FABRIC supports experiments using their Jupyter Hub, we elected to download the FABRIC Python developer software and control our experiment from Linux Hosts, since this would model how the experiment would operate in production. After we

configured the software we explored the FABRIC feature set and settled on these features for our prototyping. We would…

1) use a variety of FABRIC sites to represent the distributed elements of a final system.
2) Use multiple sites for computations to assess moving the computations to mitigate downtimes of any FABRIC site.
3) connect all these sites using FABRIC Level 3 networking, and chose to use IPV4 protocol because we understood it.
4) write a driver program that allocated resources based on declarations.
5) Initially nodes moving data to the glocal research networks were contended to a few fabric sites which supported IPv4 connections to the greater IP network. We would have to essentially use the control plane for initial data movement.

As prototyping progressed, used two additional new features, as the were deployed by the FABRIC project.

1) Persistent storage
2) Program controlled, full access to the internet for each node

System level provisioning of these features is managed using *planner.py* a tool we wrote. Functions provided by the tool are depicted in the figure, below.

```
    plan                show plan, but do not call FABRIC APIs
    apply               instantiate the plan by calling FABRIC APIs
    delete              delete a slice, if it exists
    print               Print information about nodes and networks in a slice
    json                emit json describing slice
    mass_execute        Execute command(s) on all nodes in the slice. if the
                        command begins with @ treat the argument as a file of
                        commands.
    execute             execute a command(s) on a specfic node
    debug               call a slice into memory and start the debugger
    slices              print the name of all my slices
    resources           experimental -- take a peek at resources
    health              test topology health by having each node ping all the
                        others
    aliases             Print out alias for ssh commands to each node. Note that
                        FABRIC nodes with IPV6 addresses cannot be reached from
                        sites not supporting IPV6. "-u prints correponding unalias
                        commands.
    dns                 Setup /etc/hosts so all node can addess all others by name.
                        These names can be made up and don't leak out of FABRIC.
    renew               renew a slice so it sticks around
    format              format fabric-provided storage at a site.
```

At the pipeline-infrasturcture level we are able to integrate a conventional production system using Apptainer containers, HTCondor, and file transfer methods supported by RUCiO, which is under evaluation by CMB-S4 as a component of its data management Framework. We were successfully able to run science pipeline payloads using this infrastructure.

Our work shows that FABRIC, while research infrastructure, is capable of supporting middleware typical of production systems used in experimental science. Moreover,This type of infrastructure seems capable of handling the prompt processing needs for CMB-S 4's Chilean Data, and for its South pole data, should the required network bandwidth to the South pole be available to the experiment.

Observations

Fabric provides these features which support reliable, prompt production:

Private, closed high bandwidth experiment-specific, distributed private networks. Mitigating the need for all distributed protocols to be hardened to the level required on the open internet, Implying greater choice at the application level.

Information security services (at least) on par- with current on-prem security services. Unlike on-prem security services, the meta-facility would have the distributed scope of the experiment as its prime concern. FABRIC currently provides bastion services, and hardened access to its Management network. Future infrastructure could provide scanning, log retention, log analysis, vulnerability scanning and other services currently available only in the scope of on-prem, or for a (substantial) fee for commercial cloud-based deployment.

Supporting software-as-infrastructure provisioning model, among distributed sites. This infrastructure-as-code model supports reliability by the ability to re-deploy the private network and computational resources to new sites. This allows long-running experiments to maintain operations with minimal disruption during hardware upgrades and allows all experiments to have higher availability than if locked into a particular site.

Satisfying common operations needs. A common infrastructure can support common modules to support the experiment's operations, and be a basis for a level of experiment support from the facility. As an example of this, FABRIC has a plug-in system level monitoring capability, allowing system level monitoring of all nodes in a deployment. Such a system also allows joint monitoring between facility and experiment. An additional example is that FABRIC maintains a base of trusted Operating system images.

For large project, the ability to built a distributed proxy for sites while the actual physical locations where the experiment will run are not yet constructed. Using FABRIC we've been able to have The southernmost node, at Florida international University, play the role of "observatory", etc…

Support add-on scope to a project. Often as an experiment evolves there are proposals for extra scope. For observational experiments, (such as CMB-S4) the science in the new scope can depend on getting prompt or real-time access to data. As an example, CMB-S4 will run for many years, and as the project is being designed, the transient sky in for its observational window is not well understood. If the project supports "taps" on its prompt data streams, add-on projects (which might have their own funds) that can be “grafted” onto CMB-S4, the computing for the add-on project is provisioned within the same common infrastructure.

Summary:

We have found no blockers for the use of infrastructure like FABRIC. There is potential for significant common services if something like FABRIC is available as common infrastructure for distributed experiments such as CMB-S4.