

CompF4 Edge Services Summary

Topic Leads:

Ofer Rind <rind@bnl.gov>

Benjamin Weaver <benjamin.weaver@noirlab.edu>

Scope/Definition of the Topic

Edge Services operate at the interface between a data center and the wide area network, separated from the data center's core services. This includes middleware that facilitates user access between the data center and external systems (e.g. storage, databases, workflow managers). These services may be managed externally in partnership with the data center and federated across multiple data centers.

Scope/Definition of the Topic

- We expect significant overlap with other topics in Computational Frontier.
- Edge Services can be used “interior” to the data center as well as outside.
- Services may be user-supplied.
- Definitions are evolving rapidly.

Community Input/Relevant Whitepapers

- This list is not intended to be comprehensive.
- Alex Sim, *et al.*, Deploying in-network caches in support of distributed scientific data sharing, [arXiv:2203.06843](https://arxiv.org/abs/2203.06843).
- Dave Casper, *et al.*, Software and Computing for Small HEP experiments, [arXiv:2203.07645](https://arxiv.org/abs/2203.07645).
- Doug Benjamin, *et al.*, Analysis Facilities for HL-LHC, [arXiv:2203.08010](https://arxiv.org/abs/2203.08010).
- Tom Lehman, *et al.*, Data Transfer and Network Services Management for Domain Science Workflows, [arXiv:2203.08280](https://arxiv.org/abs/2203.08280).
- K. Herner, *et al.*, Towards an HPC Complementary Computing Facility, [arXiv:2203.08861](https://arxiv.org/abs/2203.08861).
- Maria Acosta Flechas, *et al.*, Collaborative Computing Support for Analysis Facilities Exploiting Software as Infrastructure Techniques, [arXiv:2203.10161](https://arxiv.org/abs/2203.10161).

Findings I - Examples

- [Spin](#) is a container-based platform at NERSC that allows users to deploy individualized application stacks, including databases, data access tools and analytics.
- Services such as [My NERSC](#) provide individual and collaboration-scale access to metadata, allowing for long-term storage planning, or tools to manage access permissions.
- Data transfer services such as [Globus](#) are routinely implemented in edge environments such as a dedicated DMZ or data transfer node.
- [NSF's SLATE](#) is intended to provide “DevOps”-like service deployment federated across many possible data centers.
- [ServiceX](#) combines both data extraction and delivery for high-energy physics data sets.

Findings II - Themes

Focusing on some common features of edge services

- Containers
 - Containers are most commonly orchestrated with [Kubernetes](#).
- Configuration with declarative language
- DevOps
 - Specifically with version-controlled repositories
 - “System as software”
 - “Infrastructure-as-a-Service”
- Federation
 - Federated identity
 - Federated services

Findings III - Concerns

- Future maintainers should not be subject to highly specialized software & systems.
- Kubernetes is powerful, but difficult to learn, configure and maintain.
- Security needs to be considered at every step.
- Third-party and federated identity: what happens when these services are unavailable?
 - Hypothetically lock out scientists because ORCID is down for maintenance?

Recommendations/Research Directions

- Edge services have the potential to play a key role in supporting increasingly complex and distributed computing models by:
 - allowing users more flexibility and control over their services;
 - offloading some of the customization and support burden from the facilities.
- Standardization and automation of edge services could particularly benefit smaller experiments with more limited development resources.
- Provisioning for edge services needs to rise to the highest level of data center planning.
 - Facilities need to support users in these efforts by providing easy “push button” solutions to common tasks.
 - Facilities need to take advantage of federated services, scaling capabilities beyond individual centers
- Edge services exist in many fields beyond HEP.
 - Share!

Findings I - Examples (Backup)

- [XCache](#) is a data caching service based on [XRootD](#). It is used by the LHC ATLAS collaboration for data distribution.
- [Open Science Grid \(OSG\) Data Federation](#) enables users and institutions to make datasets available to high-throughput computing environments.
- [Frontier Squid](#) is an OSG implementation of the [Squid](#) cache proxy software.
- [CERN OpenStack Private Cloud](#) is an Infrastructure-as-a-Service integrated with CERN computing. Similar to Spin, but VM-based.
- [Chameleon](#) is a cloud testbed service that Internet2 [OESS](#) orchestration.
- [FABRIC](#) combines testbeds (e.g. Chameleon) with CI services.