Introduction:

The Faculty of Arts and Science (FAS) Research Computing (FAS RC) team at Harvard University (https://www.rc.fas.harvard.edu/) provides state-of-the-art research computing resources to all its affiliated faculty, researchers, and students. These resources are provided to the community through a centralized high-performance computing (HPC) cluster, Cannon, that FASRC manages in partnership with the Massachusetts Green High Performance Computing Center (MGHPCC). Through FASRC, the community has access to HPC computing, secure enclave (FASSE), storage, scientific software, high-level technical and scientific consultations, training, documentation, and guidance on how to leverage hardware and software for their scientific applications and workflows on the cluster.

Compute Clusters

Cannon Cluster: As of May 2024, the Cannon cluster provides access to:

- 99,900+ CPU cores: Intel Xeon (Sapphire Rapids, Cascade Lake, Skylake, Icelake), and AMD (Genoa, Zen, Zen2).

- 1,000+ GPUs: NVIDIA Volta (V100), Ampere (A100, A40), and Hopper (H100).
- Network over 4*100 Gbps Ethernet (GbE) for high-speed data transfer, providing several petabytes of high-performance parallel storage.
- Compute nodes are interconnected over NDR (Next Data Rate) and HDR (High Data Rate) high speed Infiniband based on the generation of hardware.

FASSE Cluster (Secure Enclave): The FASSE cluster supports data hosting and analysis of up to Harvard data security level 3 data. The cluster is made up of:

- 48 nodes totaling up to 2,304 cores of water cooled Intel Cascade Lake cores.
- 4 nodes each with 4 Nvidia A100 GPU's.
- 18 nodes set aside for high memory work up to 500GB of RAM and 1 node with up to 2TB.
- Nodes are interconnected with high speed HDR Infiniband.

Access and Authentication

The FASRC environment provides end-to-end account and authentication services for the clusters, storage, and attached services. This provides each lab with discrete accounting as well as the ability to sponsor lab members and external collaborators.

Access to the cluster is available through standard means such as SSH with two-factor authentication, as well as through the browser-based Open OnDemand graphical interface for Cannon and FASSE jobs when connected to the FASRC VPN.

Data Storage

FASRC provides large-scale networked attached access to research data from a variety of endpoints: compute cluster, instrument workstations, and personal computers. Various storage types are designed for specific needs that stem from raw data collection, data processing and analysis, data transfer and sharing, data preservation, and archiving. These storage types have different data retention and replication policies and are offered under Storage Service Center.

	Tier 0	Tier 1	Tier 2	Tier 3
Description:	High Performance Lustre	Enterprise Isilon	NFS Storage	Tape
Performance:	High I/O	Tiered	Moderate	Low
Snapshot:	No	Yes	No	No
Disaster Recovery:	No	Yes	Yes	No
Mounted to cluster:	Yes	Yes	Yes	No
Samba/SMB mount:	No	Yes	Yes	No
Encrypted at rest:	No	No	No	No
Transfer Mechanism:	Globus	Globus	Globus	Other
Data Management:	Starfish	Data IQ	Starfish	N/A
Object Store:	No	No	No	No
Location:	Holyoke and Boston	Holyoke and Boston	Holyoke and Boston	Holyoke

Scientific Software and Applications

Research or Scientific software is a collection of tools, codes, and libraries available to researchers on the cluster that allows them to generate simulations, analyze results, and provide meaning to their data. FASRC maintains and provides free access to most commonly used research software across multiple disciplines, such as Matlab, R, Python, to its users. To effectively facilitate research, FASRC provides guidance and help to its users for installing any domain-specific tools and software, needed for the workflow, on the cluster. This is achieved by providing guidance through consultations, documentation, and training on installing software using package managers (such as Spack and Mamba), utilizing Containers, and/or installing from the source. The domain-specific software spans a wide range of areas ranging from AI/ML, Bioinformatics, Computational Chemistry, Applied Sciences, Astronomy, and Physics to name a few. FAS has site or volume licenses for several packages for cluster and/or desktop use, which are provided at no cost to faculty and students. Licenses are usually available upon request or are specific to schools or departments. However, desktop/laptop installation of research software is self-service and one can reach out to their local HUIT Desktop support for installation help.

Infrastructure

Besides providing compute and storage services, FASRC also supports instrument labs, e.g. sequencers, working with instrument vendors to help install and maintain controlling computers and the storage necessary for the proper operation of the instrument.

In addition, FASRC provides limited virtual machine (VM) provisioning and support, as well as hosting services in the datacenters we manage. We provide all necessary networking support for these services and can work with PIs, lab managers, and faculty

members to recommend IT architectures and solutions for any research computing needs.

Please either email [here](#) or enter a ticket [here](#) to request assistance or ask for additional details.

Security

FASRC Security works with a wide range of individuals and groups - researchers, School PrivSec Officers, HUIT, FASRC staff, data providers, and more - to help reduce risk across FASRC systems and services, meet the needs of researchers using sensitive data, and address compliance obligations.

While much of the work goes into transparent security controls for our infrastructure and the services you use, security engineering, or HUIT facing work such as investigating suspicious activity or compliance work, you might consult us if:

- You have questions like "Does *this-research-workflow* seem like a secure and reasonable approach?", "How does FASRC handle *this-security-process*?", or "Can I use *this-FASRC-service* to store or process Harvard Data Security Level 3 data?"
- You, your School PrivSec Officer, or a data provider for a DUA you are considering have questions about FASRCs security capabilities.
- You have questions about your security oriented responsibilities, or how to fulfill them. For example, [PI responsibilities at FASRC](#)

Contact security@rc.fas.harvard.edu for more information.

Research Data Management

[FASRC Research Data Management](#) offers tools, resources, and guidance to help individuals and groups manage research data throughout the entirety of a research project. Examples of supported topics include the development of data management plans, selection of data storage and transfer tools, data organization techniques, data ownership and security policies, and data sharing recommendations.

Contact the FASRC Research Data Manager for more information: sarah_marchese@fas.harvard.edu

Support and Training

FASRC provides support to all its users through [documentation](#), semester-based [training/workshop](#), and [consultation](#) sessions. We provide live online training sessions for new and advanced users that are paired with our online documentation. Videos of most sessions are made available for all users to view at their own pace on [Training Materials – FASRC DOCS](#) and [FASRC - YouTube](#).

FASRC drop-in office hours are held Wednesdays 12-3PM virtually and are open to all FASRC users with questions or issues related to computing, storage, data management, workflow execution, and so on. For help with technical issues, please come prepared with jobs and/or errors, and any other relevant material for a faster diagnosis of the problem. For more details, please visit [Training | FAS Research Computing](#)

FASRC also offers a help ticket system that is accessible to all our account holders and external users. Please either email [here](#) or enter a ticket [here](#) to request assistance.