**Facilities, Equipment, and Other Resources**

The Yale Center for Research Computing (YCRC)[[1]](#footnote-1) supports and provides access to four high performance computing (HPC) clusters located in an HPC data center at Yale’s West Campus facility in West Haven, Connecticut. Divisions between clusters are based on the research applications used: non-biomedical research in the Faculty of Arts & Sciences, biomedical and psychology/HIPAA. Within these divisions, any researchers can access a portion of the clusters that is designated shared usage free of charge. Additional dedicated resources can and have been purchased by research groups.

The Grace HPC cluster is intended for non-biomedical research in natural science, social science, engineering, and other fields. It comprises over 835 nodes (servers) containing nearly 23,000 total CPU cores. This cluster uses Linux as its operating system and Slurm for scheduling. Each node on the cluster has 20-36 CPU cores with 4-9 GB of memory per core. Some nodes have Nvidia GPUs or additional memory per core. Grace's network infrastructure is high bandwidth, low latency InfiniBand. Grace has ~2.6 Petabytes of usable storage in its high-performance GPFS parallel file system. The cluster was acquired using a combination of internal university funds and individual faculty grants. Grace generally operates at over 80% of its theoretical capacity.

The Farnam and Ruddle HPC clusters are intended for use in biomedical research. Access to the Ruddle cluster is restricted to researchers engaged in genome sequencing at the Yale Center for Genome Analysis (YCGA). Together these clusters total nearly 400 nodes (servers) with 8,000 total CPU cores. These clusters use Linux as their operating system and Slurm for scheduling. Each node on the cluster has 20-36 CPU cores with 5-9 GB of memory per core. Some nodes on Farnam have Nvidia GPUs and/or additional memory per core. Farnam and Ruddle use Ethernet for networking. Together, Farnam and Ruddle provide over 4 Petabytes of usable storage deployed across several high-performance GPFS parallel file systems. These clusters were acquired using a combination of NIH grants, grants from private foundations, and internal university funds.

The Milgram HPC cluster is HiPAA-aligned and supports computational research involving regulated identifiable data. At the present time, use of the Milgram cluster is restricted to specific research groups in the Department of Psychology performing neuroscience research. It comprises 65 nodes (servers) with 1,764 total CPU cores. Each node on the cluster has 20-36 CPU cores with 5-9 GB of memory per core. Milgram has 1.2 usable Petabytes of high-performance storage in its GPFS parallel file system. During the Spring of 2020, YCRC plans to expand both the storage and computational capability of this cluster.

In addition to storage facilities associated with the HPC clusters or specific departmental or laboratory facilities, Yale operates two large research storage facilities connected to a high-speed 100 Gbps Science Network. For research data in active use, the Storage@Yale file storage system provides more than 2.5 PB of mirrored storage that is optimized to provide research labs, departments, schools, and individual researchers the ability to store and use large quantities of data on the HPC clusters or any Windows, Mac or Linux computer connected to the Yale network. For archive storage, Yale operates a mirrored tape library. Researchers can stage data to and from HPC storage through archive nodes attached to the Science Network, and the system automatically writes the data to the archive storage.

Throughout its campus, Yale operates a high-performance network infrastructure in support of research computing and HPC. The 100-Gbps Science Network connects the HPC datacenter and the main campus. It also provides a direct connection to the Internet2 through a non-firewalled Science DMZ, currently running at 10 Gbps, but scalable up to 100 Gbps should that be required in the future. Virtual LAN technology is used to segregate access and applications on the network, and 10-Gbps connections are provided via the VLANs to a number of individual laboratory and departmental storage and server facilities.

To facilitate high-speed data transfers, the YCRC holds a site license for the Globus file transfer and sharing software. The YCRC supports a number of Globus endpoints for the HPC clusters. Other Yale departments and laboratories can create endpoints supported by departmental or IT staff. All endpoints are connected to Yale’s Science Network and Science DMZ, enabling high-speed data transfers among Yale, national supercomputing facilities, and other universities with whom Yale researchers collaborate world-wide.

Yale has a separate main campus network based on a 10-Gbps backbone, with most buildings connected via 1-Gbps local networks. The main campus network is currently used for all university purposes other than high-speed data transfer through the Science Network VLANs. It also provides commodity Internet connectivity via multiple 1-Gbps connections from several commercial vendors. In addition to the direct connection from the Science DMZ, Yale provides one firewalled 10-Gbps connection to Internet2 from the main campus network. Over the next several years, Yale will upgrade its main campus network, incorporating both a faster backbone and software defined networking technology.

1. For additional information about the YCRC, visit <http://research.computing.yale.edu/>. [↑](#footnote-ref-1)