**Facilities, Equipment, and Other Resources**

The Yale Center for Research Computing (YCRC)[[1]](#footnote-1) operates five primary high performance computing (HPC) clusters located in an HPC data center at Yale’s West Campus facility in West Haven, approximately 7 miles from the main Yale campus.

Faculty of Arts & Sciences Research: For non-biomedical research in science, engineering, and other fields, the YCRC currently runs the Omega and Grace HPC clusters, totaling over 1,200 nodes and nearly 13,000 cores. These two clusters are traditional Linux clusters, containing dual-processor nodes that have 8-28 cores per node and 4-9 GB of memory per core. A small number of nodes have Nvidia GPUs and/or additional memory per core. The clusters use InfiniBand networks for computational work. Omega and Grace share ~2.6 Petabytes of usable storage deployed in a high-performance GPFS parallel file system. Yale acquired the Omega and Grace clusters using internal funds primarily, except for a modest number of cluster nodes that were funded by individual faculty grants or gifts. Limited access to these two clusters is available free of charge to all faculty, students, and research staff in the sciences, social sciences, engineering, and other fields for non-disease-specific scientific research. Portions of the clusters have been dedicated to specific research groups in astrophysics, climate science, high-energy physics, energy sciences, engineering, life sciences, and other fields. In general, these clusters operate at between 80% or more of their theoretical capacities.

Biomedical Research: For biomedical research, the YCRC currently runs the Farnam and Ruddle HPC clusters, totaling nearly 500 nodes and approximately 9,000 cores. These clusters primarily support NIH-funded biomedical research, including disease-specific research. Most of the nodes on both clusters contain dual 10-core processors and up to 128 GB of memory, and they are connected via 10 Gbps Ethernet to a 40Gbps network backbone. A few nodes on both clusters have 32 cores and 1.5-2.0 Terabytes of memory, and a small number of nodes on Farnam contain Nvidia K80, 1080Ti, P100, or Titan V GPUs. Together, Farnam and Ruddle provide over 5 Petabytes of usable disk storage. These clusters were acquired using a combination of NIH grants, grants from private foundations, and internal university funds. Access to Farnam is provided free of charge for biomedical and life science researchers at Yale, primarily in the Yale School of Medicine. Several specific research groups have secured special priority access by contributing funds to the cluster purchases. The Ruddle cluster is restricted to research related to genome sequencing performed at the Yale Center for Genome Analysis (YCGA). Over the past several years, the compute and storage resources on these two clusters (and two predecessor clusters) have been nearly fully subscribed by existing users, and the clusters typically operate at over 80% of their theoretical maximum capacities.

Psychology Research: The YCRC operates the Milgram HPC cluster to support computational research in neuroscience in the Department of Psychology. This cluster, which is HiPAA-aligned to permit use for projects involving regulated identifiable data, comprises 72 nodes totaling 1,920 cores. Most nodes have 256 GB of memory. Data storage is provided on a 1.2 Petabyte high-performance GPFS parallel file system. Use of the Milgram cluster is restricted to specific research groups in the Department of Psychology.

Storage Facilities: 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 described below.

*Active Storage:* The active research file storage system provides more than 2.5 PB of mirrored storage that is optimized to provide groups (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.

*Archive Storage:* Yale operates a mirrored tape library designed to provide long-term archival data storage. Users stage data through archive nodes attached to the Science Network, and the system automatically writes the data in a tape library at West Campus and replicates it to a second tape library located in downtown New Haven.

Network Infrastructure: Yale operates a high-performance network infrastructure in support of research computing and HPC. A 100-Gbps Science Network physically connects the HPC datacenter and the two research storage facilities. 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.

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. Yale connects to the Internet2 via the Connecticut Education Network (CEN), which connects, in turn, to the Northern Crossroads GigaPop (NOX). In addition to the direct connection from the Science DMZ, Yale provides one firewalled 10-Gbps connection from the main campus network.

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

1. Additional information about the YCRC is available at http://research.computing.yale.edu/. [↑](#footnote-ref-1)