

Rutgers University: Office of Advanced Research Computing (OARC) Facilities, Equipment, and Other Resources

January 2020

OARC is a centrally funded support organization providing technical and scientific support and training to faculty, staff, and students in the use of its extensive computing, networking, storage, and software resources. OARC's computational scientists, known as Research and Education Facilitators (REFs) are experienced independent researchers with deep expertise in biochemistry, bioengineering, business administration, chemistry, data science, genomics, mathematics, mechanical engineering, microbiology, and physics. OARC's infrastructure administration team provides expertise in storage management, security, advanced networking, and automation. Together, these technical teams support and collaborate with investigators on demanding problems requiring advanced cyberinfrastructure to facilitate a path to solution.

Most of OARC's computing systems provide free, open-access to all students, staff, and faculty comprising the Rutgers research community as well as immediate, priority access for investigators who purchase compute nodes. Owners purchase the dedicated resources they need (compute nodes and storage) and OARC, through its university funding, supplies the matching infrastructure (e.g., power, cooling, network access, security, administration) along with both technical and scientific support. Jobs submitted to OARC's computing systems are managed in multiple ways depending on job type (owner or open-access), resources requested, current utilization of federated cluster members, and a variety of other factors. OARC's systems are also capable of enabling bursting of submitted jobs into commercial and academic cloud services when local resources are heavily utilized or other factors favor external resources.

Primary Computing Systems

The majority of OARC's resources are managed under a federated community cluster model named Amarel in memory of the prominent Rutgers artificial intelligence researcher, Dr. Saul Amarel. The Amarel community cluster is the umbrella for distributed resources that number approximately 600 nodes providing 15,500 compute cores, 124 NVIDIA GPUs (Tesla, Volta, and Titan models), Intel Arria FPGAs, and Intel N3000 NICs. Several memory configurations are supported with capacities from 128GB to 1.5TB and including Intel Optane persistent memory modules. All cluster nodes are connected with Mellanox InfiniBand FDR or EDR fabric. Major components of this distributed infrastructure are installed at the New Brunswick, Camden, and Newark campuses. Resources are built out through funding from university, state and federal agencies.

OARC also manages the Caliburn system which was funded by The State of New Jersey and provides computing and storage resources for qualifying proposals from researchers, commercial users, and public offices throughout the state. Caliburn comprises over 700 compute nodes and nearly 24,000 Intel Xeon cores with an Intel Omni-Path network fabric.

Data Storage and Movement

OARC's main storage facilities include four IBM Spectrum Scale appliances distributed across the University with an aggregate of 6 PB usable capacity. General purpose NAS appliances with an aggregate capacity of 1.6 PB provide backup and archive services for OARC's systems on the various campuses. To facilitate distributed data services, Spectrum Scale AFM (Active File Management) is used to provide cached copies of user account data across federation members. AFM utilizes OARC's 100 Gbps research network, CICnet, described in more detail in the next section. AFM utilizes dedicated data transfer nodes (DTNs) architected to optimally utilize the capabilities of CICnet and have been designed to move data across CICnet at speeds approaching 100 Gbps. The FIONA recipe for DTN has been also adopted by OARC to provide an affordable but flexible high-performance path to allow departments, labs, and other communities with demanding data movement requirements a simplified path to aggregate local data and move it into the OARC federated compute environment.

Data Center

Rutgers has seven data centers geographically distributed totaling about 23,000 sq ft and a combined electrical input of 1.1 MW. OARC's anchor facility is located at the Hill Research Computing Data Center which provides 200 KW of power and attendant cooling. Most power is UPS-supported with backup diesel generators.

Location	UPS Power (kilowatts)	Cooling (tons)
Piscataway (NB Campus, Hill Center)	600 (200 non-UPS)	225 (estimated)
Camden	35	3.5
Newark (Engelhard Hall)	52	30

Campus and Regional Networks:

Production

Rutgers University's production data network, RUNet, is a purpose-built network of nearly 100 routers and 2,500 switches servicing constituents of 8 campuses across three metropolitan areas and several dozen remote locations throughout New Jersey. The network interfaces with service providers at two of its main Points of Presence (POPs) in Philadelphia, PA and Newark, NJ. RUNet provides access to 30 Gbps of Internet transit, 10 Gbps of Internet2 IP transit, 10 Gbps of Internet2 Advanced Layer 2 Services, 20 Gbps of private Internet IP Exchange services and several private network interfaces to select content providers. Rutgers is also an anchor member of the NJEDge consortium, a collaboration network of NJ Higher Education.

Research

Funded in part by the NSF (Grant #OAC-1659232) and jointly supported by OARC and the telecom division of the Office Information Technology, CICnet is a 100 Gbps SDN network offering Rutgers investigators advanced compute and file access services while also serving as a testbed for new technologies and high-speed connections to local, regional, and national services in academic and commercial domains.

Regional Resource Collaboration

Rutgers has taken a leadership role in establishing the Eastern Research Network (ERN). Through a developing partnership of educational institutions, research facilities, regional network providers, and Internet2, the ERN is committed to providing layered and transparent access to shared data and computing facilities for research projects located at member sites. As the ERN develops, research communities across the region will have access to a broad range of services and resources that may not be available on any one campus alone. Academic institutions, research and educational networks and corporate participants are all working together to demonstrate the value of the ERN. Syracuse, UMaine, NJIT, OSHEAN, NJEdge.Net, Kinber, Google, Lenovo and SchedMD are among the expanding list of ERN member organizations.