

Georgia Tech Facilities, Equipment and Other Resources

1 The Partnership for Advanced Computing Environment (PACE)

In 2008, Georgia Tech created a campus strategy in support of high performance computing (HPC) resources for the campus research community. This work began as a series of monthly meetings with top campus researchers who used HPC capabilities via either small local clusters in their labs or the national facilities provided by NSF and government agencies. Out of more than a year of work, a technology model grew for central hosting of computing resources that would be capable of supporting multiple science disciplines with shared resources, dedicated resources, a group of expert support personnel, and physical data center facilities for high-density hosting of infrastructure.

As a result, Georgia Tech's senior academic, research, business, and IT leadership made significant initial investments and continuing matching and personnel funding. Recognizing this program must become a strong partnership focused on serving faculty needs, it was named the Partnership for an Advanced Computing Environment (PACE). PACE is funded through a mix of central funding (70%) and faculty funding (30%) that has proven sustainable since its inception in 2009 despite the rapid growth.

As of December 2023, the compute resources managed by PACE are comprised of ~2,000 CPU/GPU nodes totaling ~47,000 cores and ~500 GPUs, plus more than fifteen petabytes of total storage. Clusters managed by PACE include the general-purpose research cluster, Phoenix, the NSF MRI-funded cluster, Hive, a secure data enclave, Firebird, and an instructional cluster, ICE. These resources are used by over 3,300 active faculty, researchers, and students. PACE provides power, cooling, networking, and high-density racks, as well as a multi-tiered storage system including home directory, project space, and high-performance scratch space that provides a unified file system across the whole ecosystem. PACE recently adopted a new cost model based on resource consumption rather than specific equipment acquisitions, while also providing a limited free tier available for any GT faculty member and their research groups.

1.1 PACE Personnel

Currently, the PACE team consists of 23 full-time employees, plus a half-time employee for management and project support, as well as multiple students. This includes five management positions leading teams of Research Scientists, System Engineers, and Customer Support Professionals. Nine faculty-facing Research Scientists provide consultation services, training, and collaboration with the GT researchers and their external collaborators. Ten systems administration experts manage systems, storage, and cybersecurity. One Senior IT Support professional provides customer facing support and ticket triage.

GT is committed to supporting research as illustrated by the newly constructed Coda office tower and data center (see below). The Coda ecosystem is designed as a future-forward workplace enabling unique collaborations and partnerships between higher education and industry – both established and start-up. The overall theme is the nexus of research cyberinfrastructure, data sciences, and discovery focusing on HPC/HTC, high performance storage and analytics capabilities, as well as leadership class network services including research and education (R&E) backbone access.

1.2 Service Portfolio: AI/ML/DL, OSG, LIGO, Science Gateway and beyond

PACE's service portfolio constantly adapts to the rapid changes in computational needs of Georgia Tech researchers, including increasing demand for AI/ML/DL workloads, installation and optimization of scientific software, training workshops, and consulting engagements.

1.2.1 Phoenix cluster

In Fall 2020, PACE deployed the Phoenix cluster that is Georgia Tech's latest leading-edge supercomputer, which ranked 277 on the November 2020 Top500 list with 1.84 Linpack petaflops of computing power, and has continued to grow the system with additional hardware. The Phoenix accelerates the computational research efforts of the Institute's diverse community in support of data-driven work including but not limited to astrophysics, biology, health sciences, chemistry, materials and manufacturing, public policy, among other disciplines. In addition to the 34,816 cores from the 1,395 nodes, there are 57 GPU nodes with 2x Tesla V100, 37 GPU nodes with 4x RTX6000, and 4 GPU nodes with 8x NVIDIA H100 GPUs that are instrumental in accelerating AI/ML/DL workloads. The system interconnect is Mellanox HDR InfiniBand, with HDR100 100-gigabit links to each compute node, and top-of-rack switches connected via multiple HDR200 200-

gigabit links to a central core switch. Multiple petabytes of high performance storage is provided by a DDN SFA 18k storage device running Lustre. Supporting server infrastructure and storage is deployed in the Coda Enterprise hall, while compute nodes are deployed in the Coda research hall. Backups of research data are housed in the BCDC data center (see below).

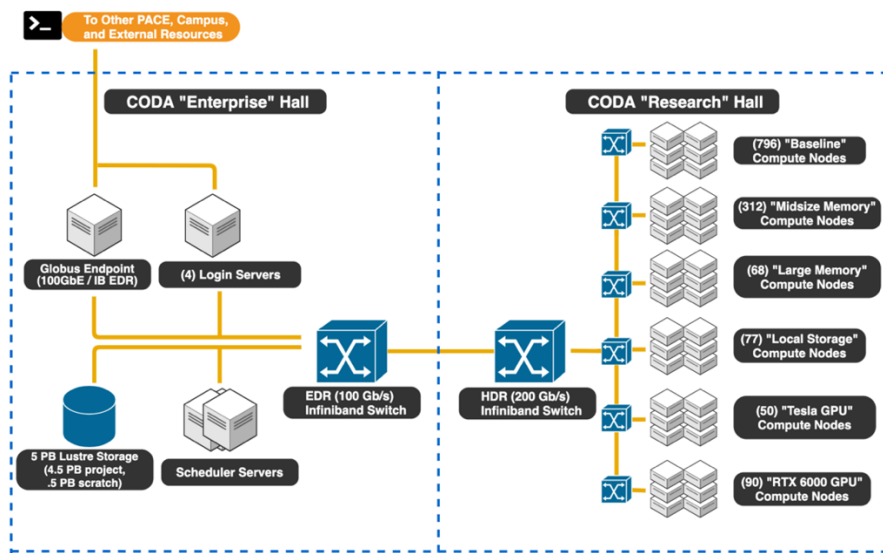


Figure 1 – Phoenix cluster architecture design.

1.2.2 NSF Hive cluster

In 2019, PACE completed the deployment of a large-scale cluster "Hive" through an NSF MRI award 1828187: "MRI: Acquisition of an HPC System for Data-Driven Discovery in Computational Astrophysics, Biology, Chemistry, and Materials Science." This cluster offers a rich variety of different capabilities from a base set of CPU nodes to large/extreme memory nodes, as well as fast local storage and GPUs (Figure 2). PACE has partnered with the Cyberinfrastructure Integration Research Center (CIRC) from Indiana University to provide a science gateway, facilitating intuitive access to the Hive cluster for use by GT researchers, XSEDE users, and researchers from partner Minority Serving Institutions in the region. Architecturally, Hive is similar to the Phoenix cluster. The system interconnect is Mellanox EDR InfiniBand, with a 100-gigabit link from each compute node to a central core switch. Multiple petabytes of high performance storage is provided by a DDN SFA 14k storage device running Spectrum Scale (formerly GPFS). Supporting server infrastructure and storage is deployed in the Coda Enterprise hall, while compute nodes are deployed in the Coda research hall. Backups of research data are housed in the BCDC data center (see below).

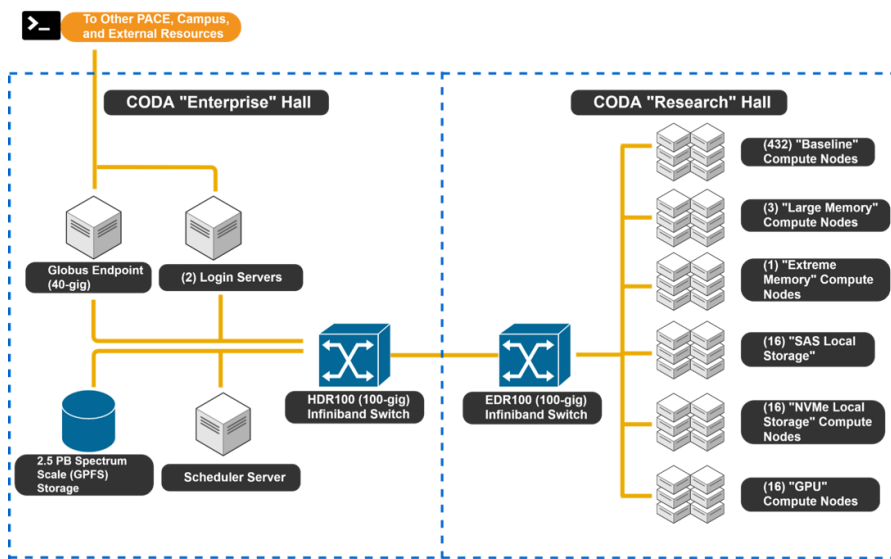


Figure 2 – Hive cluster architecture design.

1.2.3 Instructional cluster (PACE ICE)

To support growing demand for coursework requiring research cyberinfrastructure, PACE operates the “Instructional Cluster Environment (ICE)”. These resources offer an educational environment that’s identical to production research resources, providing thousands of students at the graduate and undergraduate level with ample opportunities to gain first-hand scientific computing experience including HPC and GPU programming each year. Furthermore, the entire PACE scientific software repository is made accessible to all ICE students, providing an education environment that mirrors production research clusters in every aspect. In addition to credit-bearing courses, ICE enables PACE team to develop and host hands-on tutorials and training workshops to help GT researchers and students improve their computational skills.

In partnership with the College of Computing and College of Engineering, ICE resources include 60 Cascade Lake nodes, 10 RTX 6000 GPU nodes, 25 NVIDIA V100 nodes and 20 8x NVIDIA H100 nodes optimized for single precision workloads. The 20 8x NVIDIA H100 nodes represent a significant collaboration with the College of Engineering, which will support a wide base of AI education at Georgia Tech, and function as an “AI Makerspace” in the tradition of other physical Makerspaces at GT. 18x additional H200 nodes have been shipped and will be installed in late 2024 to expand the AI Makerspace as part of the ongoing effort.

1.2.4 LIGO and Open Science Grid

PACE custom built a cluster for the LIGO project, consisting of 25 compute nodes, each equipped with 24-core Cascade Lake CPUs with 196GB memory. This cluster is fully integrated into Open Science Grid (OSG) to run jobs submitted locally and remotely using a HTCondor gateway integrated with OSG's central glidein factory service. PACE manages all of the infrastructure and services that are specific to LIGO runs. These services include scheduling, gateway, GridFTP, resource monitoring and accounting, proxy service (Squid) and several monitoring tools and health checks.

PACE has extended OSG services to other projects, namely LIGO, IceCube and CTA/VERITAS, which also rely on the OSG ecosystem. As a part of this effort, PACE received an NSF CC* award 1925541: “Integrating Georgia Tech into the Open Science Grid for Multi-Messenger Astrophysics”. With this award, PACE added CPU/GPU/Storage to the existing OSG capacity, as well as the first regional StashCache service that benefits all OSG institutions in the Southeast region, not just Georgia Tech.

1.2.5 CUI cluster

PACE manages a Controlled Unclassified Information (CUI) cluster for sensitive research that requires compliance with NIST 800-171 such as ITAR and various export control regimes. The CUI cluster is comprised of 49 compute nodes, each with 24-core Cascade Lake CPUs and up to 768GB of memory. In addition, this cluster features a node with 4 single-precision GPUs.

All the CUI systems are physically isolated from the main clusters, with their own network, HDR InfiniBand, storage, and supporting IT infrastructure and compliance mechanisms.

1.2.6 ScienceDMZ

PACE has deployed a ScienceDMZ, which utilizes 40 and 100 Gb/s connectivity to the Georgia Tech research network and Globus for GridFTP file transfers. Data transmitted through the ScienceDMZ is replicated in real time to high-performance storage available within PACE clusters.

1.2.7 Archive Storage

The PACE archival storage service is designed to house those important datasets that need to be retained but are not in active use. A design optimized for low cost and reliability makes PACE archival storage an attractive component of data management plans. Widely familiar to the national research community, Globus is used to transfer data to and from the archive. Data within the archive is internally replicated multiple times to ensure availability, and is connected via a 10-gigabit path.

2 Cost model

In January 2021, Georgia Tech formalized the adoption of a new method of funding research cyberinfrastructure. The new model aims to be more sustainable and flexible and has the full support of Georgia Tech's executive leadership. The new model is based on actual consumption of resources – similar to that of commercial cloud offerings from Amazon (AWS), Google (GCP), and Microsoft (Azure). In this new model, PIs are charged only for what they actually use rather than for some fixed capacity which may remain idle. Other advantages of this model include:

- Researchers have more flexibility to leverage new hardware releases instead of being restricted to hardware purchased at a specific point in time. Researchers can tailor computing resources to fit their scientific workflows rather than being restricted to the specific quantity and configuration of compute nodes purchased.
- Rapid provisioning without the requirement to wait for a lengthy procurement period to complete.
- Insulation from failure of compute nodes. Compute nodes in need of repair can be taken out of service and repaired, allowing jobs to proceed using other compute nodes in the pool rather than decreasing the capacity available to a particular user.
- A free tier that provides any PI the equivalent of 10,000 CPU-hours (per month) on a 192GB compute node and one TB of project storage at no cost.
- PACE staff will monitor the time jobs wait in the queue and procure additional equipment as needed to keep wait times reasonably low.
- Note that a similar consumption model has been used successfully at other institutions such as Univ. Washington and UCSD, and this approach is also being developed by key sponsors (e.g. NSF's cloudbank.org).

PACE has twice secured the necessary campus approvals to waive the F&A overhead on PACE services, as well as those from commercial cloud providers, for proposals submitted to sponsors between January 1, 2021 and June 30, 2025. While not formally committed to at this point, this waiver is anticipated to continue.

Complete details on the cost model and the rates are published at this site: <https://pace.gatech.edu/update-gts-research-cyberinfrastructure-cost-model>.

3 Distributed and Federated Academic Cloud (VAPOR)

In 2014, Georgia Tech created a campus strategy in support of cloud computing resources for the campus research community. The Virtual Applications and Platforms for Operations and Research (VAPOR) facility is a joint project between the Colleges of Computing, Engineering, and Science, PACE, and the Office of Information Technology (OIT). The facility is comprised of several Redhat RDO and Redhat Enterprise OpenStack environments. These individual cloud environments are federated together using the Redhat Cloudforms open hybrid cloud-management framework. All of the VAPOR cloud services utilize the campus standard identity management and authentication mechanisms to facilitate federation and sharing of resources between campus units.

4 Physical Facilities

Georgia Tech facilities are designed to meet the demanding requirements of modern HPC, data, and network systems, maximizing their availability with power, cooling, and storage redundancy measures. Research cyberinfrastructure operated by PACE resides in two data centers - Coda, the newest and primary facility, and the Business Continuity Data Center (BCDC).

4.1 Coda

Georgia Tech and PACE are anchor tenants in Coda, a multi-block 21 story technology focused building offering 645,000 sq. ft. office space, 80,000 sq. ft. data center space, as well as street-level retail and collaborations spaces. Approximately half of the office space is occupied by various Georgia Tech academic schools and research centers, OIT, and interdisciplinary research neighborhoods. The other half is occupied by industry partners.

Coda supports the economic development of Atlanta and the State of Georgia through job creation, new tax revenues, and a technology cluster. It drives anticipatory innovations in research cyberinfrastructure by serving a diverse research community by converging industry, research, and educational leadership in a dynamic, world-class environment. It also provides a HPC and data center space to commercial companies to become the 'de facto center for excellence' for HPC in Atlanta. It is expected to create a new ecosystem based around a unique facility modeling high-end computational/network/data-intensive hosting defining the future in trans-disciplinary research, eco-friendly practices, and public/private partnerships.

Power capacity allocated for Georgia Tech use in the Coda data center is comprised of two separate operating environments—500 kW for enterprise workloads and 1,500 kW for research cyberinfrastructure. While the enterprise space is designed with a traditional 2N configuration for both power and cooling, the research space is configured with an Uninterruptible Power Supply (UPS) coverage and no generator backup power. The data center is designed for expansion to both generation and delivery up to a total of 10 MW, installed in phases as demand increases.

Each floor of the Coda data center is served by a dedicated chilled water loop that utilizes heat exchangers, heat recovery units, and both water and air side economization to deliver cooling at the most efficient PUE. Air is distributed within the suites either by 55 ton CRAHs for densities under 15 kW per rack, or rear door heat exchangers that tie into the chilled water loop via close coupled cooling for densities up to 50 kW per rack. The data center environmental conditions conform to the newer ASHRAE standards with water temperature for the rear doors is typically provided at 70 degrees F and an inlet air temperature averaging 76 degrees F.

Georgia Power is delivering electrical power via two geographically diverse switchable feeds at 20 kV and 4160 V via dedicated, underground, concrete encased duct banks that terminate in two, below grade transformer vaults, each containing up to five (5) x 3,000 kVA transformers. From there, power is stepped down to 4,160 V for distribution to secondary transformer vaults located proximate to each of the data center floor suites above, and then transmitted to research data suite PDUs at 415 V providing 240 V to each computer or PDUs at 208 V in the other spaces.

A research partnership with Georgia Power provides a self-contained power system nearby. This Microgrid system provides up to 1.5 MW of power to the data center using a combination of alternative technologies such as fuel cells, battery storage, and micro-turbines. The intent for the Microgrid is to provide protection and up to full operation during peak midtown grid utilization. The power feed will switch diverse sources automatically, and this happens within the minimum flywheel UPS operation time of 30 seconds. The Microgrid will auto start should both diverse source feeds fail or upon a Georgia Power directive. The facility is adaptable to incorporate future power generation technologies.

Dedicated security is located in the Coda building lobby and manned 24/7. Entry from outside requires a security badge or visual confirmation from the security network operation center. Once inside the lobby, all visitors must pass through a minimum of three security checkpoints before accessing computer racks. All ingress/egress points, hallways, loading areas and outside pathways are monitored by closed circuit video devices.

Coda is straddled by West Peachtree Street and Spring Street, which are two of Atlanta’s primary fiber pathways providing direct access to Level 3, Zayo, Fiberlight, AT&T and Comcast fiber.



Figure 3 - Coda building, depicting the office towers to the right, data center at the rear left, and the historical Crum & Forester building at the front left.

4.2 Business Continuity Data Center (BCDC)

The BCDC provides 270 kW of capacity, supported by two UPS systems with 2N redundancy backed by a generator sized to provide power for all of the systems in the facility. BCDC also utilizes a N+1 redundant 200 ton chilled-water cooling system in a traditional raised floor design to allow full distribution of cooling to all racks. BCDC is monitored by a 24x7 network operations team for cooling and power issues. Network connectivity between Coda and BCDC is comprised of multiple 100-gigabit ethernet connections as well as additional dark fiber for future expansion needs.

4.3 Physical Security

The Georgia Tech police department (GTPD), a division of the Atlanta Police Department, provides the general security on the campus. GTPD performs campus patrols, mobile camera monitoring and includes a SWAT response team for emergency preparedness and crime prevention. Georgia Tech data centers have badge level access and camera coverage including the building vicinity. Motion sensor alarms are configured to alert GTPD. All systems are monitored 24/7 by an operations team, located in Coda, which responds to emergencies and potential hazards, such as rising temperatures or chilled water leaks. Both data centers link through GTPD to the Atlanta Fire Department, which is located approximately 2.5 miles from campus.

5 Network and Connectivity

The Georgia Tech campus network consists of over **200,000** network ports across **285** buildings linked by 1,800 miles of fiber optic cabling. There are roughly **2,688** network switches from a diverse set of vendors. The campus provides a centrally managed WiFi service with over **7,400** access points. **Georgia Tech has 2x – 400 gig link connections to SoX for Internet connectivity in addition to various 10 and 100gig Internet bound connections. There are also 400 gigs of bandwidth to each data center for PACE and another 400 gigabits of bandwidth for enterprise services.** The campus network infrastructure is managed by the OIT Network Engineering team. This team is charged with providing 24x7 support of the network infrastructure, including WiFi, “from the wall jack to the Internet,”; responding to problems reported by the campus users, provisioning new equipment and connections, and troubleshooting service and performance issues. The Network Engineering team works closely with Georgia Tech CyberSecurity to define and implement policies and procedures for securing the campus IT infrastructure. The facilities provide 1 Gb/s, 10 Gb/s, 40 Gb/s and 100 Gb/s ethernet connections to all servers and HPC systems, as appropriate. The Coda data center uses EDR (100 Gb/s) and HDR (200 Gb/s) InfiniBand, with all servers and compute nodes connected to the ethernet network at least 10 Gb/s.

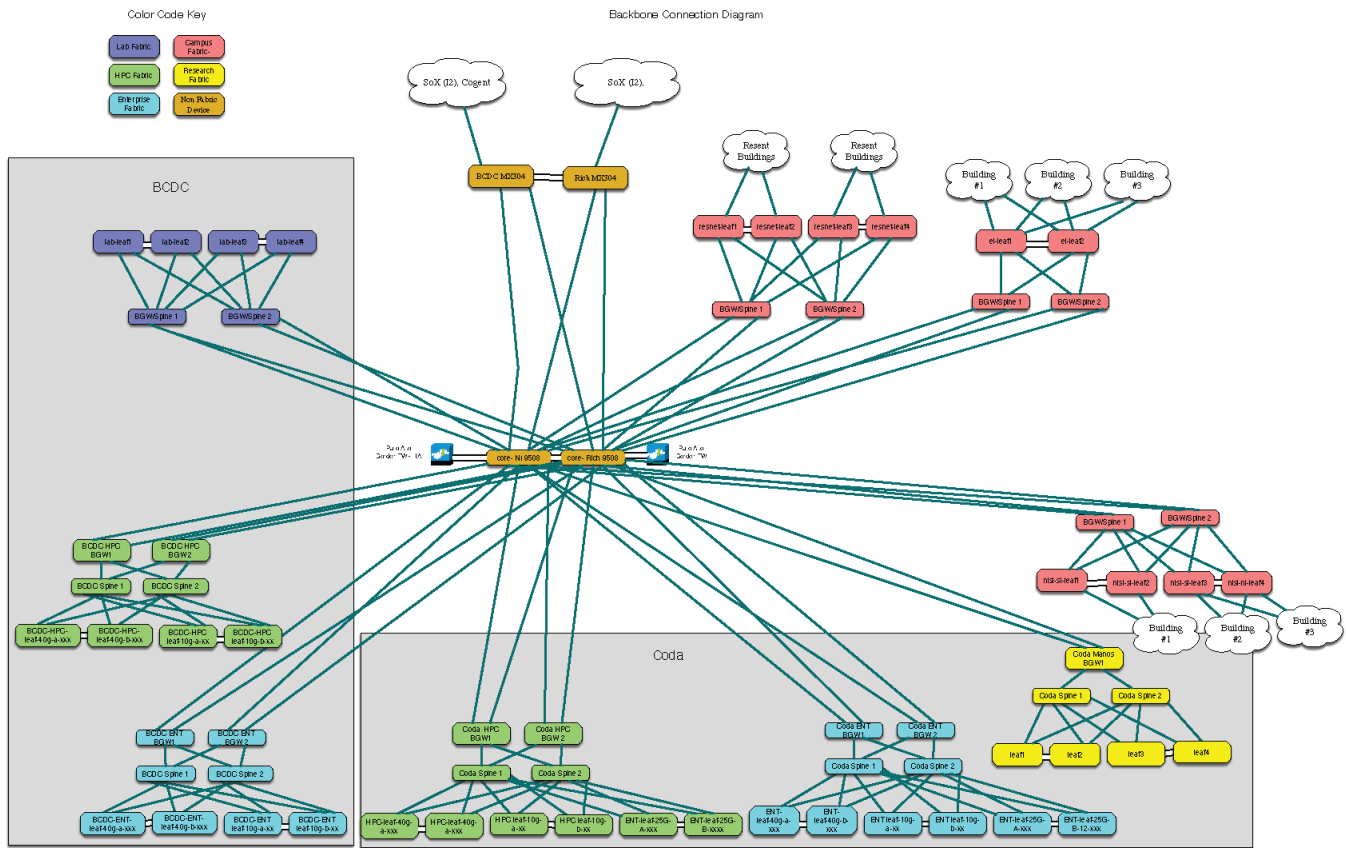


Figure 4 – Connections of core and border routers between BCDC, CODA, and SoX

5.1 Research Network Connectivity

The Coda data center has a fully redundant ethernet network topology, with separate fabrics for Research and for Enterprise, cross connected by four 100Gb/s links. Coda connects to the GT campus and BCDC with four 100Gb/s links. Figure 4 shows the connections of core routers within campus as well as connections to the Southern Crossroads Network (SoX)

5.2 External Network Connectivity

Georgia Tech led creation of the Southern Crossroads (SoX) fiber ring that connects various universities and select research institutions in the Southeast, and intends to locate network access points for that ring in the Coda. This combination of commercial and proprietary fiber gives the Coda a unique onramp onto the local and global fiber networks. SoX is a collaboration of 13 universities in the South joining forces to connect to the Internet2 network (vBNS). This collaboration, hosted at Georgia Tech and serving many of the top universities and state networks in the Southeast, has provided quality high performance network access to many scientific resources around the world. Over the past 15 years, SoX has upgraded capabilities to the current 10 Gbps network service (Internet2, NLR, Google, ESNET, and other), and has completed an upgrade to 100 Gbps connectivity to the Internet2 AL2S service. OIT manages and operates the Southern Crossroads (SoX). This strategic position of Georgia Tech allows for high bandwidth connections to leading universities and national labs, with a 10 Gb/s link to Oak Ridge National Lab (ORNL) in particular. In addition, Georgia Tech is the regional GigaPOP for I2, and Southern Light Rail (SLR), the regional aggregation.

5.3 IPv6

IPv6 has been available for services on campus since 2008, with the DNS and other necessary IPv6 infrastructure in place for this period of time. Additionally, firewalls are available for all campus IPv6 networks including a self-service capability to rapidly reconfigure access controls as needed, IPv6 networks are treated as first-order entities, on par with support for IPv4 networks.

5.4 PerfSONAR

Georgia Tech uses perfSONAR to instrument its network for latency and bandwidth. The current infrastructure is a mesh, which includes 8 nodes at 10G and 1 node at 40G across the Georgia Tech network infrastructure. These nodes are distributed at key monitoring points including the main data centers and HPC areas across the campus. Also, the Georgia Tech perfSONAR mesh participates in a disjoint mesh with perfSONAR nodes in our regional network, the Southern Crossroads (SoX).

5.5 Detailed Network Facilities

5.5.1 At SoX

- Colo facilities in Atlanta: 55 Marietta, 56 Marietta
- Colo facilities in Nashville: 460 Metroplex
- 100 Gb/s fiber to 56 Marietta St
- Juniper MX960 (56 Marietta), DC power redundant, 2x modules, 2x 100Gb, 8x 10G, 3x modules 10x 1G; Route processors redundant, Switch control boards redundant
- Dell R610 server (55 Marietta), 10G NIC Intel
- Dell R610 server (Nashville), 10G NIC Intel
- Juniper MX240 (Nashville), DC power redundant, 2x 100Gb, 16x 10Gb, 20x 1G, Route processors redundant, Switch control boards

5.5.2 At Georgia Tech

- Multiple redundant fibers between Coda and BCDC data centers
- Juniper MX304s (Rich and BCDC), with redundant route processors and two 100/400 Gigabit Ethernet line cards. (<https://www.juniper.net/us/en/products/routers/mx-series/mx304-universal-routing-platform-datasheet.html>)