Boston University Information Services & Technology
111 Cummington Street
Boston, Massachusetts 02215

Boston University

Cyberinfrastructure Plan

I. BU Research Infrastructure

o The Massachusetts Green High-Performance Computing Center (MGHPCC)

BU recently partnered with government and regional institutions to establish the Massachusetts Green High Performance Computing Center (MGHPCC), operated by BU, Harvard, MIT, Northeastern, and the University of Massachusetts—five of the most research-intensive universities in Massachusetts—dedicated to serving our growing research computing needs more collaboratively efficiently, ecologically, and cost-effectively. BU high-performance computational resources are housed at the MGHPCC, taking advantage of an abundant source of clean renewable energy from the Holyoke Gas and Electric hydroelectric power plant on the Connecticut River. This project created a world-class high-performance computing center with an emphasis on green sustainable computing, offering unprecedented opportunities for collaboration among research, government, and industry in Massachusetts. MGHPCC is the first academic research computing center to achieve LEED Platinum certification, the highest level of green certification awarded by the Green Building Council.

MGHPCC comprises 33,000 SF of space optimized for high performance computing, a 19 MW power feed, and high-efficiency cooling to support up to 10 MW of load with provisions for expansion to 30 MW. The MGHPCC owns the 8.6 acre site, providing substantial room for expansion. Network infrastructure includes a fiber loop passing through Boston, the Northern Crossroads (NoX), and New York City, connecting BU to the MGHPCC at 40 Gb/s. BU is also actively engaged in other national and regional consortia where we provide technical, strategic, and governance leadership.

o Research Computing

Information Services & Technology (IS&T) Research Computing provides specialized hardware, software, training, and consulting for all areas of computational research at BU. Research Computing resources are typically used in computational science and engineering, simulation, modeling, bioinformatics, genomics, data analysis, and other disciplines that require high-performance computing, massive storage, or sophisticated visualization. Resources are managed in close consultation with BU's Research Computing Governance Committee, our Center for Computational Science, and the Rafik B. Hariri Institute for Computing and Computational Science & Engineering.

Our primary computational facility is a Linux-based Shared Computing Cluster (SCC) with approximately 8,000 CPU cores and nearly 3 PB file-system. Forty-six SCC nodes are equipped with NVIDIA GPU accelerator cards, for a total of 236 GPUs and 115,456 CUDA cores. The SCC comprises heterogeneous subsets of nodes interconnected by GigE, 10GigE, QDR and FDR InfiniBand. The SCC uses Globus Online to expedite bulk data transfers with unimpeded 10GigE access to Internet2 and other research networks.

SCC resources are available on both a fair-share, no-cost basis to all faculty-sponsored research groups and via buy-in funded by researchers for priority use. Excess buy-in capacity is returned to the pool for general use.

SCC resources are managed by Research Computing at no charge. A dedicated service model provides fee-based management for other research computing systems that don't align with shared or buy-in models. Finally, a co-location model offers stand-alone rack space, power, cooling, and network access to proprietary resources.

Research Computing supports approximately 2,200 individual researchers working on more than 450 different research projects supported by more than $340M in external funding across our Charles River and Medical Campuses.

Research Computing also provides both individualized consulting and group training. More than 30 training classes are held each semester covering research computing basics, programming, mathematics, data analysis, and scientific visualization. Several hundred researchers attend these sessions each term.

o Communications Infrastructure

The BU campus network provides high-speed access for over 100,000 devices to all institutional information, communication, and computational facilities, as well as the Internet, regionally aggregated resources, and advanced networks such as Internet2. Tens of thousands of ports and wireless access points are interconnected via optical fiber and a robust hierarchy of high-speed routers and switches. The University has three full Class B IPv4 address space assignments.

Our network was designed and deployed with flexibility, reliability and scalability as core tenets. We adhere to best-practice distribution standards, from secure stacked network closets, patch panels, and cabling, to pervasive IP spoof checking via Reverse-Path Forwarding validation on all segments and institutional IP address verification at borders. We also maintain a comprehensive database that maps the network to floor plans spanning hundreds of buildings.

Wired and encrypted 802.1x wireless network access is available throughout campus, including all residence halls and classrooms. The University recently invested several million dollars to ensure the entire wireless network conforms to the most current 802.11n standards. Redundant connections to the Internet and Internet2 operate at 10 Gb/s with automatic failover.

II. Planning

Cyberinfrastructure planning at BU is an integral part of comprehensive institutional IT strategy developed by stakeholders within an overarching governance framework. BU leadership established formal IT governance several years ago as described at bu.edu/tech/about/governance/.

Institution-wide IT strategy was formulated and published at bu.edu/tech/plan/ in 2014.

Note that, even prior to establishing formal governance and an IT strategic plan, the University had an extraordinarily rich and enduring history of commitment to shared infrastructure in support of research and education as highlighted in **Exhibit A** below.

The cyberinfrastructure goals highlighted below include and augment the 2015-20 BU Technology Plan Initiatives.

o Communications Infrastructure

- Ensure rapid and secure provisioning of new services by deploying industry standard Identity and Access Management (IAM) infrastructure, including federated identity.
- Deploy Eduroam and IPv6 transport for our clients and public facing Internet services. Toward that end, we have a /32 IPv6 allocation from ARIN which we now announce on edge routers to Internet2 while an internal technical working group formulates [1] IPv6 address allocation strategy, [2] a core protocol deployment plan to utilize IS-IS routing, [3] an assessment of network component IPv6 readiness, and [4] a plan for implementation of an IPv4 to IPv6 gateway. Eduroam was deployed in June 2014.
- Provide unified, high fidelity communication and collaboration services for one-to-one, one-to-many, and many-to-many voice and video communication that integrate messaging, scheduling, shared workspaces, and social networks while employing social networking to enhance online research communities.

o Research Computing

- Support data lifecycle, including the capability to securely store, protect, share, publish, archive, and sunset research data in accordance with policy and data management.
- Develop a dynamic, collaborative pool of human resources for the development and support of computational software for research.
- Ensure that shared research computing services fully support researchers on both the BU Charles River and Medical campuses, enabling both optimal use of resources and maximum collaboration.

o Regional Resources

- Build on the Massachusetts Green High Performance Computing Center (MGHPCC) to provide sustainable, shared research computing, and storage infrastructure to researchers.
- Join peer institutions and industry in developing safe and effective means to leverage cloud infrastructure for research, including high performance computing and big data.

Exhibit A: Campus and Regional Resources

Campus Resources

- ### Communications Infrastructure

BU developed one of the largest open-access computing and communication services in education in the early 70s, supporting thousands of subscribers. We were among the first institutions to establish an IP-based campus network in the early 80s and connect it to NSFNET in partnership with regional research universities.

BU has been a leader in regional networking for more than thirty years. We brought BITNET to New England in the early 80s and we subsequently founded NEARnet, the New England component of NSFnet, with Harvard and MIT. The three institutions later founded the Northern Crossroads in 1999 and BU was the first institution to connect to the vBNS, providing New England with shared access to the earliest vestige of Internet2.

- ### Research Computing

BU is a pioneer in centrally supported computational research and the deployment of leading edge shared scientific computing resources. We were one of the first universities to install IBM 3090-based vector processing, Thinking Machines Corporation CM2 and CM5 parallel computers, SGI Origin2000, IBM Power4 p690, and IBM Blue Gene/L. In 1989, the University created the *Center for Computational Science* for institutional scientific computing and visualization support, providing focus to drive computational science and advanced research computing. That group subsequently evolved into *Research Computing*.

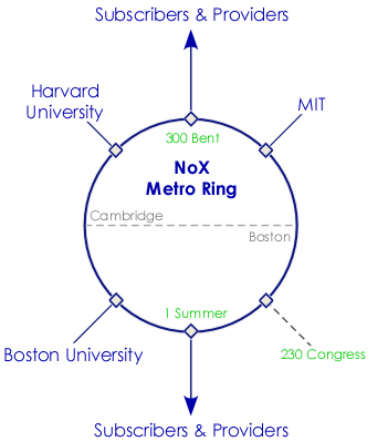
Regional Resources

Boston University has a long history of collaboration and partnership with national computing centers and bridging local resources nationally. BU was a partner in the Metacenter Regional Affiliates program and with NCSA in the Partnerships for Advanced Computational Infrastructure. Today, BU participates in the XSEDE Champions program, facilitating the use of NSF resources by regional researchers. We are also working through XSEDE to bring other cyberinfrastructure campus bridging technologies to local and regional systems using Globus Online in a pilot project. Additional discussions regarding a Genesis II pilot are ongoing.

In collaboration with Harvard, BU runs the New England ATLAS Tier 2 Center, one of five such centers in the US and a participant in the Open Science Grid. ATLAS is one of two general purpose experiments running on the Large Hadron Collider at CERN involving over 3,000 physicists world-wide. The NE ATLAS Tier 2 supports thousands of researchers with 4,500 compute cores and approximately 3.2 PB of storage.

In 1998, BU joined forces once again with Harvard and MIT to found the Northern Crossroads (NoX), a regional collaborative established to share human, material, and intellectual resources to foster the development and delivery of advanced network services in New England. The NoX fosters aggregation and collaboration among participating institutions and our respective academic communities and consolidation of regional representation to the national and international community. NoX participants meet regularly to consider issues at hand and to hear

from and talk with advanced network advocates about their commitment to R&E initiatives such as Internet2.



In 1999, the NoX established a robust service platform to support the goals of its participants, including sharing the cost of connections to Internet2 and provide a high-performance regional exchange for participants & commodity Internet service providers. In 2003, BU, Harvard, and MIT executed a twenty-year indefeasible right of use with Level3 for a 7.4 mile radius 144 optical fiber loop that runs throughout Boston and Cambridge providing access to major carrier facilities. Portions of this optical ring are allocated to each of the three partners and, in turn, to the NoX, greatly enhancing options for connectivity and providing essentially unlimited intercampus capacity. When subscribers connect to this exchange, local traffic remains within the exchange, improving performance and reducing demand on more congested and costly backbone networks.

In 2000, BU and the NoX joined forces with other R&E regional network service providers to form The Quilt where leaders throughout the regional networking community build on the intellectual capital and best practices of network service providers worldwide.

BU was an author of the 2000 proposal to extend Internet2 through the Sponsored Education Group Participant (SEGP) program. In 2005, BU was asked to participate in Internet2 strategic planning and has also served on various Internet2 governance and advisory councils, including the Network Planning and Policy Advisory Council (NPPAC) and Internet2's Governance and Nominations Committee (GNC).

BU is leading the Massachusetts Open Cloud (MOC) project, a collaboration between universities, industry and the Commonwealth to create a new public cloud, designed and implemented in Massachusetts. The MOC is the first cloud based on the "Open Cloud eXchange" (OCX)—a novel, transformative marketplace model for public cloud offerings. Partners in the collaboration include BU, Harvard, UMass Amherst, MIT, and Northeastern University, as well as the MGHPCC and Oak Ridge National Laboratory. The MOC is funded by a $3M Commonwealth of Massachusetts MassTech Collaborative Matching Grant Award and more than $16M of matching grants from industry participants.

The Commonwealth Computational Cloud for Data Driven Biology (C3DDB) is a computer system funded by the Massachusetts Life Sciences Center to support research that connects life science research with emerging innovative big data analytics. The C3DDB is located at the MGHPCC and operated jointly as a shared resource by MIT, BU, UMass, Northeastern, and Harvard.

BU is partnering with Harvard, MIT, BU, UMass, and Northeastern in the North East Storage Exchange (NESE) to design, build, manage, and deliver a multi-petabyte, multi-tier, shared storage exchange for science with a sustainable operating model. NESE leverages and extends existing regional infrastructure to support computationally and data intensive research, with an emphasis on lowering barriers to collaboration across institutions and disciplines.