Campus Cyberinfrastructure Plan
University of North Carolina at Charlotte

Introduction

The purpose of this Cyberinfrastructure Plan (CI Plan) is to document the current capabilities, strategy and plans for cyberinfrastructure at The University of North Carolina at Charlotte (UNC Charlotte). The CI Plan is a required document for responses to certain National Science Foundation solicitations. Moreover creating, implementing and updating cyberinfrastructure plans are critical to enabling research and advanced education in a forward thinking research university.

UNC Charlotte is committed to building on its strong research and teaching capabilities with a goal of joining the elite ranks of R1 research university as defined by the Carnegie Classification and has established an 'R1 Commission' to build a roadmap to achieve that goal. Robust cyberinfrastructure plans are critical to that goal. UNC Charlotte has also recently consolidated and streamlined the delivery of all information technology services including research and academic computing into one organization called OneIT. The document is reviewed and approved by the University Chief Information Officer and Vice Chancellor for OneIT.

For the purposes of the CI Plan, we will define cyberinfrastructure as the computing, communication and security infrastructure required to support advanced research and training in science, social science and engineering disciplines beyond the capabilities required for office productivity, remote and automated learning and administrative functions of the University. To be clear, for cyberinfrastructure to be productive and secure, a key requirement is to be closely integrated with and maximally leveraging the broader IT infrastructure and services of the University.

Further, this document primarily focuses on the cyberinfrastructure, which supports multiple colleges and departments within the University. Future versions of the document may incorporate key department level infrastructure.

Strategy

At UNC Charlotte, we provide core cyberinfrastructure capabilities through shared resources with dedicated staff, with planned upgrades and expansions to support a growing research footprint at the University. We leverage Campus IT infrastructure wherever possible to eliminate duplicative services and provide seamless cross service access, with well managed security to maximize researcher productivity and information assurance.

We gather requirements for the environment through a variety of methods including direct communication with researchers as well as the IT staff supporting the respective colleges and departments. We also have an advisory group for Research Computing with faculty and administrative representatives from the University research community. We consult with industry experts including peer groups at other institutions, vendors and community to track technology, identify and implement best practices.

We are also working to expand access to advanced cyberinfrastructure by encouraging the use of national resources through XSEDE and other computational resources. We have and will continue to maintain high speed external internet connectivity to facilitate use of external resources and to support our faculty's extensive collaboration with other institutions.

Core Campus IT Capabilities Leveraged for Research

Core IT Security

Security is a core requirement for any IT environment. UNC Charlotte utilizes a Defense-in-Depth strategy to implement a high availability Zero-Trust architecture. The university information security governance program utilizes ISO 27005 risk management strategies to mitigate unnecessary risk and reduce overall threat exposure. Through the use of administrative, physical, and technical security controls we have implemented a cybersecurity architecture that provides a high level of information assurance, reliability, and access control. University administrative controls align with ISO 27002 requirements, while supporting technical and physical controls are layered across the network and computing hosts. UNC Charlotte has multiple full-time cybersecurity personnel dedicated to the information security program providing services such as incident response, vulnerability management, endpoint security, compliance, and application security. The University is a member of the InCommon federation and leverages that for all SSO. The University is registered with the Research and Scholarship Entity Category and it also meets InCommon Baseline Expectations for Trust in Federation.

Core Campus Networking

The campus network infrastructure is centered around a 100-gig capable Multiprotocol Label Switching (MPLS) core, with all major distribution layer devices on campus enjoying 10-gig and 40-gig link connectivity to the core devices. The MPLS architecture provides the ability to create multiple isolated networks while utilizing a common network hardware infrastructure. Communication between the isolated networks is governed by a Firewall at the MPLS fusion point. The campus is connected to our internet provider MCNC by way of two redundant 20-gig etherchannels which utilize diverse physical paths for campus egress. Access to the Internet2 backbone, and round the clock Internet connectivity support are also provided by the MCNC service. In addition the campus has a tertiary link which can scale up to 10-gig in the event our two primary links go down. The primary wireless network for campus participates in Eduroam, and offers secure, world-wide, roaming access developed for the international research and education community.

The campus has obtained an /32 IPv6 block. There is currently a project underway to microsegment the campus, both inside and outside the datacenter. During this project IPv6 will be enabled as sections are micro-segmented. This will allow for the controlled rollout of IPv6 and the decommissioning of IPv4 over time.

Advanced Cyberinfrastructure

The core cyberinfrastructure supporting research at UNC Charlotte is organized under the University Research Computing (URC) organization within the OneIT organization.

URC provides a network of High Performance Computing cluster partitions including storage services to support the research mission of the university. There are currently eight clusters or partitions that include approximately 5,000 computing cores, 57 TBs RAM, 58 Titan/Tesla/GeForce GPUs and 3,600 TBs (usable) disk storage. In addition, UNC Charlotte provides over 50 TBs of networked data storage specifically for faculty research for use outside the cluster environment. The resources are split between clusters focused on HPC and analytics. These resources have been upgraded to current technology on a regular basis, including significant additions in 2020 and early 2021. The plan is to regularly upgrade and expand this infrastructure to keep up with the growing demand of the university research community.

The primary HPC environment is built on the latest generation of Intel and AMD based compute servers, connected via an EDR and HDR generation Infiniband fabric from Mellanox/NVIDIA. A subset of the servers are equipped with NVIDIA GPU's. Each node in the cluster has high bandwidth access to a parallel file system, based on Lustre. The environment supports a variety of programming models including MPI, OpenMP and a broad spectrum of scientific applications. The resources are scheduled via SLURM to maximize throughput and ensure fair sharing of the capability. The analytics environment is Intel based servers configured as a Hadoop cluster and connected to the same Infiniband fabric as the HPC cluster. This would enable hybrid HPC/Analytics applications in the future. All nodes in the clusters are also configured as NFS clients with access to an NFS based storage environment described below. These environments are allocated to faculty and research staff members by request. The systems are used for a wide range of applications including computational fluid dynamics, molecular biology, DNA Sequence Analysis, Geographic Information Systems, Social Science and many other areas of research. To date through employment of fair share scheduling techniques we have not had to apply a peer review process for allocations, but may need to as demand increases.

The storage environment consists of both NFS and Lustre distributed file systems to meet the different IO profiles required by researchers. The storage is structured and allocated with distinct spaces for user home directories, project and scratch space governed through use of quotas.

Supplementing the primary HPC and Analytics clusters are separate smaller Educational Clusters for use by faculty and students for instruction in University courses in computational science, computer science, data science and other disciplines. These clusters are used by over 350 students per semester. The strategy for provisioning the Educational Cluster resources is based on 'trickle down' from the primary environments following major upgrade cycles enabling a high educational impact with a relatively small capital investment.

The environment also contains a Data Transfer Node (DTN) with a 10Gb connection to the internet. The DTN and its placement in the campus network was patterned from the Science DMZ templates and best practices. This dedicated system provides high-performance data movers running optimized bulk data transfer. Tools include GlobusOnline/GridFTP and a performance measurement/test node running perfSONAR

User accounts on the URC systems are aligned with the campus identity system and leverage the campus two factor authentication system (DUO). Security monitoring of the URC systems is also supported by the University security team under the direction of the University Chief Information Security Officer.

All of the above infrastructure is housed in secure data centers and is supported by a professional staff of four full time employees, who plan and run the infrastructure and directly support the university research community. Additional resources are available within the broader OneIT organization to support networking, storage, and security integration and alignment with other campus IT systems including technical consulting as well as governance. Close collaboration with and leverage of the IT staff supporting individual colleges is also an ongoing activity.

Ongoing and Planned Cyberinfrastructure Initiatives

During 2021 a number of initiatives are under way to improve and expand the University's cyberinfrastructure.

- The resources outlined above include a new AMD based cluster, an improved data protection scheme for the Lustre file environment. A powerful system to support advanced AI and other analytics work through a single node with 8 NVLINK connected NVIDIA A100 GPU's is also being installed in the HPC cluster. These new systems are being installed in an additional newly available data center connected to the primary data center via long distance Infiniband (NVIDIA/Mellanox MetroX2).

- The Research computing group is also supporting new initiatives to explore the use of HPC suitable containers using Singularity. This will enable a broader set of applications and the ability to rapidly deploy fully tested container based application environments. One of the projects dependent on Singularity also will support long lived certificate based access to enable unattended work flows.

- User and University stakeholder communication is also being expanded with redesign of the internal web site, the creation of science driven content highlighting the scientific research enabled on the environment.

- The group has also begun an initiative to promote the use of National Science Foundation supercomputing resources through participation in the XSEDE program's Campus Champions program. This effort is starting with pilot projects with faculty, which will be documented to enable other campus researchers to leverage these resources.

- A campus wide collaboration is in process to alleviate 'last mile' bottlenecks between departmental level data intensive instrumentation and other data sources and the HPC/Analytics environment. The effort includes a proposal to the NSF's CC* program.

- Self assessment is an ongoing process for UNC Charlotte's cyberinfrastructure including outreach and comparison of practices between UNC Charlotte and other similar institutions.

Summary

UNC Charlotte has invested in robust, professionally managed, cyberinfrastructure seamlessly integrated with campus IT and has processes in place to sustain the environment. Innovation initiatives are in place to continuously find ways to improve the research experience with the infrastructure. The University has a strong feedback mechanism to insure that the needs of the campus research community are being met.