Montana State University
Campus Cyberinfrastructure Plan
January 2015

A. Executive Summary

Montana State University, the land-grant university for the State of Montana, is developing and deploying a data-centric cyberinfrastructure (CI) emphasizing digital scholarship in support of the institution's learning, research, and outreach missions. Montana State University is a Carnegie RU/VH classified institution attracting over $100M in sponsored research investment annually with primary federal funding from the National Science Foundation and National Institutes of Health.

Over the last four years, the campus has matured its investment strategy in cyberinfrastructure by recognizing the need for common institutional infrastructure investments to benefit the broader academic community. The 2012 University Strategic Plan codified these thoughts with explicit goals calling for increased investments in infrastructure to support broader accessibility. In the last two years, the academic and administrative components of campus have united to leverage faculty startup funds to create a community-computing cluster administered by full time information technology professionals.

A critical component of Montana State's CI strategy is a partnership between the Information Technology Center (ITC) and the Library to support the generation, curation, discoverability, and long-term archiving of research data. Working together, the two organizations leverage one another's expertise to avoid redundant resources and provide integrated platforms that tie back to academic publications and instruction.

In some campuses, CI is the infrastructure of the 1%; investments are made that advance scientific discovery but have little visible impact on the broader campus. Montana State believes that a focus on digital scholarship with specific programmatic efforts to integrate research data into the classroom can help assure that CI investments have broader impact across the land-grant instructional mission.

Montana State University has the following long-term CI goals:

1. Create an abundant and flexible CI driven by digital scholarship.
2. Develop a hybrid model that recognizes unique local needs for resources while balancing against the economies of scale that can be gained through leverage of additional public and private sector CI resources.
3. Use research investments for CI to create new opportunities for instruction and outreach driven by scientific discovery.

B. Data Intensive Institutional Cyberinfrastructure

B.1.1 Conceptual Overview-Cyberinfrastructure development at Montana State University is enhanced by a "digital scholarship" model that supports the research enterprise and teaching mission. The model creates a virtuous feedback loop resulting in powerful research developments that improve the health of research and learning communities.

Montana State defines digital scholarship as the creative and deliberate application of computing technologies to research. It is often multidisciplinary, requires a high degree of collaboration, and is influenced by and engaged in digital culture. Digital scholarship thrives in an environment where knowledge and data are openly shared. The successful application of this model requires a

foundation of Library and IT services. Our goal is to enhance the capacity of faculty and students in research, collaboration, modeling, and experimental design to cooperatively increase our institutional reputation, external funding, and scholarly output. This is the "digital scholarship" virtuous feedback loop, supported by and driving infrastructure and services to attract and retain top talent.

B.1.2. Data Management/Curation-The Library and ITC have implemented services to archive and preserve MSU's research data and other scholarship. ScholarWorks—MSU's Institutional Repository—is structured to link datasets to publications, thereby contextualizing and clarifying research results through the supporting data. Open access to scholarship enables and encourages collaboration between partners across state, national, and international institutions.

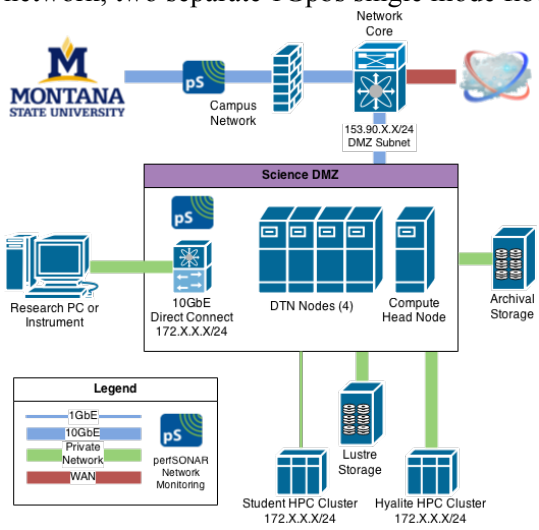
The Library provides consultations on metadata and data documentation for researchers, as well as acting in a liaison role to connect researchers to research cyberinfrastructure (see section B3). The Library pursues digital library development to support online collections of MSU scholarship, and it actively engineers discoverability in commercial search engines for digital archival collections and scholarly content archived in ScholarWorks.

B.1.3. Future Directions-The Library and ITC will continue to expand research data services, with a focus on curating and publishing MSU research data for use by students in the classroom. Montana State is also developing a set of services that will help university departments, centers, and institutes establish semantic entities in order to enhance MSU's online profile and institutional reputation. The goal for MSU's online discoverability includes creating machine-actionable components to tailor our content for search engines and semantic analysis

B.2. Montana State University's Networking Infrastructure

B.2.1. Montana State University Local Area Network (MSUNet): The foundation of the MSU network is two Cisco Nexus 7000 core routers. These cores are physically separated in the two data centers on the campus that are connected to one another with a 100 Gbps backplane. Each Nexus 7000 is dually connected to eight-fiber distribution network at 10Gbps.

Distribution switches are connected via fiber to 48 building head end switches. Of these switches, 47 are currently 1Gbps while 1 is at 10 Gbps. The original fiber plant build out includes 8 single mode and 12 multimode fibers to each connected building via 8 intermediate distribution frames. MSUNet serves 8600 wired hosts from 344 networking devices with an average utilization of 2.5Gbps of the Wide Area Network. In addition to the interior campus network, two separate 1Gpbs single mode fibers provide connectivity to the affiliated MSU Research Park and the Museum of the Rockies.



B.2.2. MSU's Science DMZ: In order to meet the data needs of MSU researchers, a physical portion of the campus network has been reserved specifically for the transmission of research data.

The Science DMZ, deployed in compliance with the DOE standard, is explicitly designed as a dedicated portion of the MSU network that is distinct from the enterprise network. The architecture of the system includes both publicly addressable systems within the MSU network and, for increased security, private subnets where

necessary: for example, high performance compute nodes, etc.

The DMZ backbone includes capacity for the local resources, but is also intended to connect research labs and instruments directly to the DMZ if required. Internally, the DMZ uses 10GbE networking to connect the Data Transfer Nodes (DTNs) to the WAN and to research cyberinfrastructure. High performance storage and research archival storage systems are connected through the DMZ to the DTNs, enabling high-speed data ingress to any connected lab or instrument. Conversely, this also allows high speed data egress from either storage system (or connected labs and instruments) nationally and internationally. The recommended software tool for data transfer is Globus, using a provided public endpoint to the DTNs.

It is of note that the Science DMZ was entirely self-funded by MSU. This component of research cyberinfrastructure was considered a high priority by the institution to help support EPSCOR related research from the Institute on Ecosystems.

B.2.2.1 Network Performance Monitoring for the DMZ: perfSONAR nodes measure network performance from both within the Science DMZ and within the enterprise campus network. This allows technicians to monitor and diagnose data transfer issues between remote locations, the campus network, and the Science DMZ.

B.2.3. Transition to IPv6-With recent upgrades to the core and data center network MSU is now more fully able to support IPV6. In 2015, MSU will undertake a readiness audit to determine the current utilization of IPV4 space on campus and best opportunities for a phased deployment of IPV6. Dual stack deployments will likely be the foundation for this transition. The audit will focus first on assessing the utility and effectiveness of using IPV6 within the MSU Science DMZ with a focus on the community cluster, research data repository, data transfer and network performance monitoring tools.

B.2.4. Anti-Spoofing BCP38-MSU's CISCO routers are all BCP38 compliant and the campus is actively using these capabilities in both our routers and firewalls. In addition, the campus is planning a pilot of enhanced network security software.

B.2.5. MSU's Wide Area Network-MSU's Wide Area Network (WAN) provides 30 Gbps of broadband networking connectivity for research and commodity traffic. The physical WAN is a collaborative effort through the Northern Tier which is used by MSU to reach the Pacific Northwest GigaPop (PNWGP). PNWGP provides research-networking connectivity to MSU via egress to Internet2 and also commodity Internet services to campus. The MSU WAN includes connections to the MSU sister campuses, the State of Montana network, and the University of Montana through SummitNet, a private MPLS cloud.

B.2.5.1. Regional Networking Collaboration with the Northern Tier and Pacific Northwest GigaPop- MSU's IRUs through the Northern Tier provide two westbound10Gbps DWDM waves and two eastbound traveling through Minneapolis with connectivity to StarLight in Chicago. MSU presently has an active10Gbps connection west to Seattle and east to Billings. An additional 20Gbs will be activated by March 2015. Campus utilization of the WAN at peak times currently reaches approximately 3Gpbs.

B.2.5.2. Internet2-Through our relationship with the PNWGP, MSU has access to Internet2 Advanced Layer 2 Services as well as CENIC and other significant national and international research hubs across the world. MSU is an active member of Internet 2 and has leveraged a number of Net+ service offerings including Eduroam, Office 365 Education, and the InCommon

certificate service. Strategically, MSU will continue to use these community efforts to further deploy core services based on need and available fiscal resources.

B.2.5.3. Bozeman Fiber-In addition to active involvement in national and regional networks, MSU also spearheaded efforts to enhance local broadband. In 2014, the Bozeman City Council voted to move forward with the deployment of an open access public fiber network. A dedicated connection from the community fiber network to MSU presents opportunities for peering relationships with public entities, a redundant physical path for campus to network services from PNWG, and a new way to deliver local campus CI investments to the community. MSU envisions using the new community fiber to enhance STEM collaborations with local educators.

B.2.5.4. Future Directions-Montana State University will continue to mature the networking foundation of the university with deployments of IPv6 and active network defense on the local area network. In addition, MSU will search for opportunities to leverage additional regional networking efforts via PNWG and Internet2 to meet future broadband needs of researchers on campus. Active monitoring of current network utilization, engagement with researchers, and the Data Census will drive these investments.

B.3. Community Cluster and Research Storage Repository-To meet the research, computational, and storage needs of researchers, MSU has established local resources for high performance computing, scalable high speed storage, and durable archival storage for data archiving and publication through ScholarWorks (the library's institutional repository for data discoverability). There are two major strategic goals of this institutional infrastructure:

1. Provides access to resources that are typically reserved only for a select few who operate these personal computational infrastructures in isolation and
2. To build an infrastructure that supports the "digital scholarship" virtuous feedback loop. It begins in the lab, proceeds through publication, is discoverable via the library, and available as a reusable component of curricula.

The Hyalite high performance computational cluster was installed in late 2014 and put into production in the spring of 2015. Researchers, the Provost, the Vice President for Research, and the Chief Information Officer collaboratively funded and have expanded the resource. The MSU IT Center's Research Cyberinfrastructure (RCI) division manages the resource centrally. The cluster currently consists of computational nodes connected to high-speed lustre storage. The cluster utilizes 10GbE RDMA networking to provide fast and efficient connectivity to the Science DMZ at a manageable cost. Hyalite is composed of 36 Compute Nodes for a total of 576 SandyBridge Xeon cores and a total of 2.3TB of RAM; connected to 320TB of lustre storage and 57TB of archival storage.

Hyalite is available to all MSU researchers, with access to priority queues for those that contribute funds for cluster expansion. This represents a unique opportunity for many faculty who otherwise would not have access to instruments of this caliber. Hyalite also includes a student-facing cluster composed of hardware that is no longer needed but is still viable for computational workloads. The cluster is available to all students at MSU and provides an HPC sandbox space for experimentation and classroom projects.

In addition to local resources, MSU has built strong partnerships with national research entities. A local XSEDE campus champion is available to assist researchers in gaining access to 383,000 SUs across six national supercomputing resources. Integration with InCommon grants researchers easy and efficient access to a majority of this infrastructure. MSU maintains a strong relationship with our peers: the university is a founding member of the Rocky Mountain Advanced

Computing Consortium (RMACC)1, and works closely with other regional high performance computing centers (RMHPC, Wyoming SCC).

A critical component of MSU's long term computational and data archival strategy is to determine the appropriate role for local campus resources while investigating the economies of scale that can be gained through partnership with commercial resource providers. MSU is actively engaged with both Amazon and Microsoft in defining "burstable" resources that could augment production runs that exceed campus capacity or that may more cost-effectively be brought to bear for long-term archival storage.

B.3.1. Future Directions-Montana State will continue to invest to increase the capacity of the community computing cluster and research storage while using research projects undertaken in coordination with the Library to identify the balance between local CI needs and the opportunities to leverage both national public resources (XSEDE) and private sector partners.

B.4. Identity Management-MSU LDAP and Active Directory identity sources are provisioned by custom feeds from our ERP system (Banner). Single sign-on authentication and authorization services are provided by the Central Authentication Service (CAS) supporting intra-campus access to web resources, and the Shibboleth federated identity service. InCommon membership grants MSU trusted and convenient single sign-on access to national and international online educational and research service providers.

C. Education and Outreach

C.1. Formal MSU Curriculum-Digital Scholarship is the foundation for MSU CI, and acts as the bridge to a variety of structured outreach and education efforts. A multi-departmental effort is underway in Business, Computer Science, and Extended University to create a new data science certificate. Institutional CI investments will be made available in off-peak demand cycles to students and instructors in order to have the widest impact possible.

C.2. Informal Education-The Library has also partnered with the Museum of the Rockies (a Smithsonian Museum Affiliate) to build public collections like the Indian Peoples of the Great Plains (http://arc.lib.montana.edu/indian-great-plains/) and provide archival and storage services for these collections. In addition to this campus-based focus, ITC and the Library are committed to working with the NSF Science Montana STEM program developed under the institution's EPSCOR grant to develop future modules that would leverage research data stored, curated, and archived at MSU. The Data Management Librarian conducts outreach with labs, departments, and faculty members in order to facilitate good data management practice and to encourage data archiving and sharing, especially in ScholarWorks.

D. Sustainability

D.1. Sustainability and Responsibility for Maintaining Federal Investments-Montana State University's CI strategy relies on leveraging federal research opportunities to enhance the infrastructure and services available to the local research community. In recognition of the stewardship responsibilities of these investments, Montana State commits to use central investment and personnel available in the Information Technology Center to support bringing and maintain these assets in the campus production environment for the use of the academic community.

[1] https://www.rmacc.org