NSF Campus Cyberinfrastructure and Cybersecurity for Cyberinfrastructure PI Workshop

Recently the National Science Foundation (NSF) held it's NSF Campus Cyberinfrastructure and Cybersecurity for Cyberinfrastructure PI Workshop in Albuquerque, New Mexico on October 2-4, 2017. The workshop provided an opportunity for NSF award recipients to meet in-person, exchange project findings, interact with national cyberinfrastructure experts and collaborate across program areas, project areas, and project regions.

LEARN and its members contributed to the rich content of the workshop. Rice University's Klara Jelinkova provided a presentation on CC* Networking Infrastructure: Improving Network Infrastructure to Enable Large Scale Scientific Data Flows and Collaboration. Also, LEARN's President and CEO, Pankaj Shah, moderated the Cyberinfrastructure to Support Large Data Transfers in Genomics Research Panel. The panel was comprised of Deniz Gurkin of the University of Houston and Zac Blue from the Baylor College of Medicine.

The workshop was held in conjunction with the ESnet Site Coordinators Committee (ESCC) meeting and The Quilt Fall Member Meeting, providing more opportunities for networking, collaborative discussion, and establishing and strengthening relationships within the research community.

LEARN congratulates its members on their NSF awards and their contributions to the community.

You will find more on our member researchers' projects on the pages that follow.

ATTENDEE	TITLE	ABSTRACT
Zac Blue	**CC*DNI Networking Infrastructure: Enabling Frictionless Scientific Data Transfers in the Texas Medical Center**	The Texas Medical Center (TMC) in Houston, Texas, is home to a 15-year networking association called the Collaborative Healthcare Alliance for Technology (CHAT), comprised of 11 research and clinically focused organizations and over 70,000 researchers, educators, students and staff. CHAT is led and operated by the founding member, Baylor College of Medicine (BCM), in support of the community's explosive research-centric, network-intensive needs. This project replaces aging infrastructure with state-of-the-art equipment, while simplifying the design, and simultaneously enabling frictionless, high-speed data transfers by creating a first-ever Science DMZ, directly benefiting research labs such as: - Baylor's Human Genome Sequencing Center, led by Dr. Richard Gibbs, which has transferred hundreds of terabytes of genomic data in the past year, most notably as a part of the CHARGE project; - The Center for Genome Architecture producing and sharing some 50 Tbytes of new sequenced data each year; and - The Center for Metagenomics and Microbiome Research engaged in the production and collaborative processing of tens of terabytes of microbial genomic data. While updating the underlying infrastructure, the project also upgrades the connection between CHAT and its nearest upstream provider, the Southeast Texas GigaPoP (SETG), from 10 to 100 Gbps, matching the recently upgraded 100Gbps connection between SETG and LEARN, the Texas-wide research and education network backbone that ties the R&E community to Internet2. By streamlining research data transfers, the investigators will empower researchers at BCM and other CHAT partners to hasten the eradication of today's greatest health challenges.

ATTENDEE	TITLE	ABSTRACT
Priscilla Parsons	**CC*DNI Networking Infrastructure: Data Driven Network Infrastructure Upgrade for Lamar University Research**	Lamar University is introducing a Science DMZ designed to efficiently achieve higher levels of performance, reliability and predictability for critical STEM science research projects and address next generation big data research needs. The proposed cyber infrastructure provides connectivity to state and national high performance computing facilities, including those supported through the NSF/XSEDE program, from multiple research locations across campus. Lamar University faculty research includes, but is not limited to, biology, chemistry, environmental science, electronic systems, biomedical diagnostics, natural disasters, engineering, and high-performance computing and data analysis. The dedicated Science DMZ provides transformative capability for both research and educational programs by interconnecting research-intensive areas on campus to one another through a 10G fiber backbone while removing obstacles to efficient data flows between research laboratories and external collaborative computational and analytical facilities. The new design provides multiple 10Gbps-routed ports, DMZ switches for link consolidation and aggregation, a high-performance data transfer node, and a Perfsonar node for performance monitoring and testing. In addition, the optical fiber network is upgraded to single-mode fiber connections for research-intensive areas along with distribution and access layer switching to provide 10G capacity. The new network also provides broader impact benefits to graduate and undergraduate students by incorporating project design and operation into classroom lecture, student engagement in intensive computational and data driven research and independent student research efforts.

ATTENDEE	TITLE	ABSTRACT
Deniz Gurkan	**CC*DNI Networking Infrastructure: Custom Science DMZ Per Research Lab with a Secure Invitation to Opt-In**	Science data transfers originating from various research labs to other collaborating labs and cloud resources have long been enabled by significant investments in research infrastructure at the university level. However, when a research lab transfers large scientific data sets, many hurdles still exist for research and Information Technology (IT) communities to work together in creating secure topologies for research collaboration. The University of Houston's vision is to enable the abstraction of circuit setup, traffic steering, and leveraging of network functionality for end-users to opt-in to connect their data with researcher labs in a customized fashion with relevant data sharing policies enforced on the topology. This project: (1) deploys network function instantiation (NFI) within main distribution frame (MDF) components at strategic buildings with research labs on university campuses; (2) implements an interdisciplinary data sharing isolated network customized for an air quality and healthcare research use case; and (3) deploys the NFI capability as a pilot on LEARN (Lonestar Education and Research Network) for future support of science data flows. This work has a broad impact across many dimensions: isolated networks per research lab creates an environment for accelerated innovation, free of operational constraints. The proposed prototype research topology supports researchers in eliminating the data sharing problems, with concepts that extend to all research collaborations.

ATTENDEE	TITLE	ABSTRACT
Andrea Fumagalli Gi Vania	**CC*DNI Integration: PROnet: A PRogrammable Optical Network Prototype Serving the Campus**	Without high-performance networking big data and high-performance computing have very limited value. It has long been recognized that the highest performance data transfers should directly run the on top of optical circuits without Ethernet switching or IP routing. This direct layering avoids switching-related packet loss and provides a cost-effective service. Optical equipment now exists that will permit the dynamic establishment of single application light paths. However, the time required to set up optical circuits is too long for short or small transfers, taking a second to set up a 100 Gigabit/sec link in order to transfer a small file does not make sense. This project is building an integrated campus network where small data transfers (megabytes or gigabytes) ride conventional networks and big data transfers (terabytes or petabytes) ride directly on optical circuits. The project is building PROnet, a campus network at the University of Texas Dallas, which provides the lowest possible network latency (signal propagation only), scalability to 100Gbps and beyond, and equipment having low power consumption and complexity. A user-friendly interface will be developed and deployed for the network inexperienced researcher to request the provisioning of as needed optical. A PROnet controller will be developed for the scheduling and automatic provisioning of optical circuits, which will shorten the circuit reservation procedure time significantly by removing human intervention from the entire process. For inter-domain operation, the open platform OSCARS will be extended to work with PROnet. With this OSCARS extension, provisioning of hybrid (optical and virtual) circuits spanning across multiple (optical and non-optical) domains will be a reality. The software will support secure use of the network via InCommon identities.

ATTENDEE	TITLE	ABSTRACT
Klara Jelinkova Joseph Ghobrial	**CC* Networking Infrastructure: Improving Network Infrastructure to Enable Large Scale Scientific Data Flows and Collaboration**	Campus networks are required to protect information and inspect data flows to safeguard security and privacy. However, researchers need open and unfettered access to large data flows and instruments across the globe to reduce time to discovery. The Rice University network is a shared resource that not only needs to support the administrative and teaching functions, but also enable scientists to use that network in new and innovative ways for research. Five key data-intensive application teams act as drivers of the new extension of network functionality and are providing feedback to the technical design staff. These application areas include earth and atmospheric sciences; urban data science; computational biosciences and neuroengineering; particle physics and distributed cluster computing. These applications build on long-term science investments aimed, amongst others, at understanding seismic events, the weather patterns in the Gulf regions and beyond, as well as urban trends in large, diverse cities such as Houston, TX. The basic model adopted by the project is "the science DMZ." A Science DMZ is "a portion of the network, built at or near the campus local network perimeter designed so that equipment, configuration, and security policies are optimized for high-performance scientific applications rather than for general-purpose business systems." This approach allows Rice to aggressively upgrade its network capacity for greatly enhanced science data access. This project supports 100 GB/s flows between the data transfer facilities at our off-campus data center and national and international R&E data repositories and takes advantage of SDN (Openflow) mechanisms.