Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2020 Proceedings | Advances in Information Systems Research

Aug 10th, 12:00 AM

Accelerating FIU's science research and education towards discovery and innovation by leveraging FIU's Science DMZ

Vasilka Chergarova
Florida International University, vc574@mynsu.nova.edu

Julio Ibarra
Florida International University, julio@fiu.edu

Cassian D'Cunha
Florida International University, cdcunha@fiu.edu

Mike Kirgan
Florida International University, michaelk@fiu.edu

Rolier Cabrera
Florida International University, rolier@fiu.edu

See next page for additional authors

Presenter Information

Vasilka Chergarova, Julio Ibarra, Cassian D'Cunha, Mike Kirgan, Rolier Cabrera, David Rotella, Jeronimo Bezerra, Rene Rotella, and Heidi Morgan

Accelerating FIU's science research and education towards discovery and innovation by leveraging FIU's Science DMZ

Completed Research

Vasilka Chergarova
Florida International University
vchergar@fiu.edu

Julio Ibarra
Florida International University
julio@fiu.edu

Cassian D'Cunha
Florida International University
cdcunha@fiu.edu

Mike Kirgan
Florida International University
michaelk@fiu.edu

Rolier Cabrera
Florida International University
rolier@fiu.edu

David Rotella
Florida International University
david@fiu.edu

Jeronimo Bezerra
Florida International University
jbezerra@fiu.edu

Rene Rotella
Florida International University
rener@fiu.edu

Heidi Morgan
University of Southern California
hlmorgan@isi.edu

Abstract

Research faculty and their students are spending too much time on data management issues related to the transfer of data between networks. As the campus cyberinfrastructure increases data production, the transport capacity of the network must increase proportionally to deliver the data to the High-Performance Computing centers for analysis. This work presents the experience of a research project in implementing Science DMZ at Florida International University, which added six researchers and their laboratories to the Science Network and Science DMZ. The study applied qualitative approach to assess the researcher's science workflows in order to create a Science DMZ implementation plan and followed the Energy Sciences Network implementation guide.

Keywords

Science DMZ, Campus Network, Cyberinfrastructure

Introduction

Research increasingly requires collaboration across multiple institutions and the capabilities to access and share resources across multiple disciplines. Cyberinfrastructure (CI) underpins this collaborative science and engineering research. According to the Energy Sciences Network (ESNet) implementation guide and Dart et al. (2013), Science DMZ (demilitarized zone) paradigm consists of network design pattern that collectively use a cyberinfrastructure for science and education more efficiently and creates an optimized network environment. Science DMZ can be viewed as a friction free buffer between private and public networks. Some of the early adopters of Science DMZ approach include collaborative research projects at the University of Colorado in physics, Pennsylvania State University's College of Engineering and Virginia

Tech Transportation Institute collocation computing and storage resources project, and National Energy Research Scientific Computing Center' projects with scientists from National Oceanic and Atmospheric Administration Earth System Research Lab in Boulder, Lawrence Livermore National Laboratory, Iowa State University, SLAC Linear Accelerator National Lab, and Berkeley Lab (Dart et al. 2013).

This paper includes the architecture scenario, equipment, and security considerations while implementing Science DMZ (SC-DMZ) at the Florida International University (FIU) Campus. The use-case included science instruments such as the Wall Of Wind (WOW)(Aly et al. 2011), FIU's High-Performance Computing (HPC) Cluster, Magnetic Resonance Imaging (MRI) instrument at the Center for Imaging Science (CIS) (Fattal 2018), and Geographical Information System (GIS) (Krefft 2009) data. Other CI on campus, such as XNAT server (Marcus et al. 2007) that collects MRI imaging data, FIONA Data Transfer Node (DTN) server (Ramsey 2014) to transfer the MRI data to University of California in San Diego (UCSD), WOW file server that collects all the WOW experimental data and media files, and the perfSONAR (Hanemann et al. 2005) nodes, were also connected to the Science Network. Those scientific instruments perform a key role in supporting scientific collaboration through high-performance computing, network, and data, and are distributed across campuses, regional, national and international organizations, and span scientific communities. A brief description of the science application and its CI is presented in the next section.

The CIS at FIU is a multidisciplinary research center that was designed to support an integrated community of investigators at the forefront of imaging science, with an emphasis on functional neuroimaging research. The CIS includes a research-dedicated magnetic resonance imaging (MRI) facility that supports a 3T Siemens MAGNETOM Prisma. This MRI scanner is equipped to run the Human Connectome Protocol (Van Essen et al. 2012), which entails state-of-the art multiband data acquisition, advanced motion correction (PROMO), EPI distortion correction (EPIC), ultra-fast data acquisition facilitating HARDI, and high-resolution sub-second TR EPI acquisition. The GIS center at FIU supports the teaching and research activities of the FIU community in the areas of geospatial web, geospatial data/metadata creation, data management and dissemination, visualization, image processing of remotely sensed data, and vegetation classification and modeling. The collaborators at GIS obtain their data sets from various federal, state, and local county agencies. The data is obtained from satellite images and remote sensors on-site data collection. The WOW facility is affiliated with the International Hurricane Research Center (IHRC) and the College of Engineering and Computing (CEC) at FIU (Leatherman et al. 2007). The facility enables testing of entire building structures at full-scale, leading to performance-based design for hurricanes through direct correlation of wind speed with performance and damage levels. The WOW facility can test potential failure of full-sized structures, such as site-built or manufactured housing and small commercial structures. The WOW and the MRI are big data generating CI instruments at FIU. The data they generate is of much interest to research communities in the U.S. and internationally. The research experience for faculty and students working with these CI instruments is greatly enhanced from having high-throughput low-latency access to deliver and analyze very large data sets on the FIU HPC and at other remote compute centers. Use of the Science Network and SC-DMZ for at-scale experimentation increases capacity towards discovery.

Research Goals

The goal of this project was to enhance the research experiences for faculty and students by removing constraints that can result in less time being given to the research process. This project aims to provide FIU's premier research faculty and their students with unconstrained access to CI required for their research and educational activities, both on and off campus. Unconstrained access to on- and off-campus resources, such as the Texas Advanced Computing Center - TACC (Koziol 2014), the Open Science Grid-OSG (Pordes et al. 2007), the Sunshine State Education and Research Computing Alliance SSERCA resources, and the Extreme Science and Engineering Discovery Environment -XSEDE (Towns et al. 2011), is provided by FIU's Science Network and SC-DMZ. The SC-DMZ is a strategic CI that enhances research capabilities and creates an environment tailored to the needs of high-performance science applications. The use case described in this paper can bring valuable insides for entities (e.g., universities, research centers, labs) with similar science instruments that are considering the Science DMZ approach for friction-free data transfer over the network.

Research Method

A qualitative approach was applied to gather insights about the science flows, the cyberinfrastructure involved, and the collaborating sites per science domain. The data was collected during two workshops, an online survey, and several additional meetings with researchers from August to October 2018. The goal was to increase the understanding of how the data flows from source to destination and its interaction with dedicated and shared cyberinfrastructure. A project evaluator participated in the regular weekly meetings with the project team. In addition, there were several additional meetings with the project PI and the Campus CI Engineer, in order to gather more details about the project, the technical team involved in the implementation, and the researcher labs that needed to connect to the SC-DMZ network.

Two workshop meetings were conducted with the researchers from the three research groups: CIS, GIS, and WOW. The SC-DMZ team also invited participants from the university's IT department and other technology teams, which resulted in 19 workshop participants. After a SC-DMZ (Dart et al. 2013) concept presentation, the participants were asked questions about implementation and how the project could improve their research and collaboration.

An open-end questions survey was used to gather more details about the science workflows and the cyberinfrastructure involved, specific to the research domain. The initial draft was presented and discussed with the PI and Campus CI Engineer, which resulted in several rounds of editing similar to the Delphi method (Dalkey and Helmer 1963). The survey included a total of 21 questions separated into four categories: 1) Introduction to the project and identification of the research group; 2) Cyberinfrastructure section to gather information about computing systems, data storage systems, advanced instruments, data repositories, visualization environments, and people, linked by high-speed networks used for the research (Stewart et al. 2010); 3) Workflow section to gather information on the researcher's science workflows, the use of cyberinfrastructure, and how the workflows will use the SC-DMZ network; and 4) Data movement section to gather information about the size of the transferred data, collaborating sites, and additional comments. The survey was emailed using Qualtrics to a sample consisting of 13 researchers from three research groups: nine from CIS, one from GIS, and three from WOW. A follow-up email was sent two weeks later. The final survey version can be found at https://fiu.qualtrics.com/jfe/form/SV_e9RHV3tY4yOj89f. We received in total seven usable responses (four CIS, two WOW, and one GIS), which represent 53.84%. Demographic data was not collected in this study. The collected data was exported to an Excel spreadsheet and analyzed using Grounded theory (Charmaz and Belgrave 2007) approach with science workflow as unit of analysis. Multiple meetings with the PI and the Campus CI Engineer took place to clarify some of the researchers' answers due to their limited input. The Campus CI Engineer served as a liaison between the researchers, technical team, and the project evaluator.

Results

The sections below describe the interpretation of the researchers' answers to the questionnaire per science flow: CIS, GIS, and WOW. The interpretation includes science flow description, data size, collaborative sites where data is transferred to/from, cyberinfrastructure components identified in the science flow, additional researchers using the lab/science instrument, and future plans. *Verbatim researcher responses are shown in italics*. Our interpretation is shown in normal font.

The Neuroinformatic and Brain Connectivity (NBC) Laboratory

Science flow description: NBC team conducts its functional neuroimaging research at the CIS at FIU, using a research-dedicated MRI machine. A total number of 20 researchers, scientists, and graduate students are actively using the lab. According to the answer of question 12 from the questionnaire (Q 12: On average, how much time does a workflow(s) consume?), the workflow can take between minutes to days, depending on the specific process.

> *"Depending on the specific workflow, some workflows can take between minutes, hours, or days."*

These workflows are preprocessing, data quality control, and group-level analyses. The cyberinfrastructure involved in the execution of these workflows includes XNAT server, HPC cluster and a local computer.

"Preprocessing of MRI data from a single subject currently takes about 1 day to perform and data quality control takes about 3 hours."

Preprocessing of MRI data from a single subject currently takes about 1 day to perform and involves the following process steps: (1) Transferring across the SC-DMZ Network MRI data from the MRI workstation to the XNAT server, where images are archived and accessible to the researchers; (2) Transferring MRI data from the XNAT server across the FIU enterprise network on to the local computer; (3) On the local computer, the MRI data is further preprocessed so that only relevant data for the study remains. Data quality control takes about three hours to perform. The data quality control process is performed on the local computer by a research student visually examining each image.

"To date, FIU's CIS has completed more than 700 participants across all of our active data collection projects. These preprocessed single subject MRI data are then typically combined into group-level analyses which can take anywhere from minutes to hours to perform."

Group-level analyses of MRI data currently take a few minutes to a day to perform and involve the following process steps: (1) Transferring across the enterprise network preprocessed MRI data from the local computer to the HPC cluster; (2) Combining single subject pre-processed data on the HPC to perform group-level analyses.

"We are not currently transferring data from external sites to our lab/cyberinfrastructure, but we do have plans to transfer data from multiple locations (University of California San Diego, Washington University, University of Minnesota, Yale University, University of New Mexico, and the National Institute on Drug Abuse [NIDA/NIH])."

There are future plans and projects that would require data transfer from multiple locations. Table 1 shows a summary of the NBC components.

Data size	Several TB per year. On average 15 GB per day are transferred to UCSD.
Collaborative sites data are transferred to/from	The research group transfers unprocessed raw data to the UCSD. The implementation of the SC-DMZ project demonstrated benefits in this science use case, because science flows from FIU to UCSD experienced less friction.
CI components identified in the science flow	MRI machine, MRI workstation, FIONA workstation, ABCD Data Analysis & Informatics Center at UCSD, XNAT server, HPC cluster, and user desktop/ laptops.
Additional researchers using the lab/science instrument	Two additional research projects (not included) are using the MRI facility for their research are 1) Cognitive neuroscience/neuroimaging research (not started yet) and 2) Adolescence, sleep, anxiety, and depression research that collaborates with researchers in Quebec City, Canada.
Future plans	Since the current projects process data that are generated on-site at the FIU MRI facility, there was no external source of data transfer to the lab/CI.

Table 1. Summary components for the NBC Laboratory

Wall of Wind Research (WOW) Facility

Science flow description: The WOW Team is conducting wind and civil engineering research at the Extreme Events Institute, International Hurricane Research Center Laboratory for Wind Engineering Research, and the WOW Research Facility. A total of 14 researchers (8 Students and 6 Research Scientists) currently use the lab. Two types of data are generated at the WOW facility: 1) Data obtained from probes, sensors and data acquisition systems (air pressure, speed, etc.) are collected on a desktop in the WOW control room (~20GB), and 2) video and photos of the experiment are collected on tapes and physically transferred to another desktop (~ 100GB). All the data is duplicated. The first copy of the data is untouched, and the second copy of the data is preprocessed on a desktop (workstation in the control center). Both copies are transferred to a file server in the data center. FIU researchers can access the data for further analysis using their own computer. All three sets of data (untouched, preprocessed, and the processed) data are copied, via a DTN, to the TACC in Austin, Texas. The preprocessed and processed data only (not the untouched data) are made available to a larger set of FIU researchers and external collaborators through the TACC's portal. Table 2 shows a summary of the WOW components.

Data size	Depending on Project ~100MB to 100GB.
Collaborative sites data are transferred to/from	To Texas Advanced Computing Center (TACC). Currently the research group does not receive data transfer from an outside collaborator.
CI components identified in the science flow	Data Sensors, control center desktop, file server, data mover (DTN), local desktop, Texas Advanced Computing Center (TACC)
Additional researchers using the lab	There are no additional researchers at this moment besides those listed above (the research scientists and the students).
Future plans	Involve more collaborators

Table 2. Summary components for the Wall of Wind Research (WOW) Facility

The WOW Facility is recognized by NSF as an important and unique scientific instrument. The availability of the SC-DMZ provides the WOW researchers greater throughput capacity to more rapidly move their data to the campus HPC or to the TACC for analysis with future collaborators.

Geographic Information Systems (GIS)

Science flow description: The Geospatial & Remote Sensing and Landscape Ecology research, conducted at the GIS Center at FIU, involves downloading of remote geospatial data, remotely sensed data and vegetation classification, satellite images, and other data collected from remote sensors sites. The GIS group consist of 16 researchers and students. At the time of the study, the GIS Center had not been connected to the SC-DMZ network because compute and storage resources are linked to several external services managed by FIU Library. As a result, more time was required to schedule the migration of the GIS Center to the SC-DMZ network not to interrupt the library's services. Table 3 shows a summary of the GIS components.

Data size	Several TB per year
Collaborative sites data are transferred to/from	The research group download data from servers at the FIU HPC, National Park Service (NPS), EPA, USDA, NASA.
CI components identified in the science flow	Data repository, local desktop, file servers, FIU HPC
Additional researchers using the lab/science instrument	At this point, no additional researchers were identified; however, the center is facilitating teaching and research activities with students, and the number of the people involved is growing.
Future plans	Future plans include addressing the complexity of the GL library facility and security in place.

Table 3. Summary components for the Geographic Information Systems (GIS)

The GIS group was not transferred to the SC-DMZ network. Future plans addressing the complexity of the FIU Library facility and security are in place. The GIS group required specific CI security upgrades, which could lead to a lengthy change of the research processes and new CI environment.

Implementation

The implemented SC-DMZ is comprised of the following components: border router, core network switch/router, labeled SDN hybrid research core network on Figure 1, a next generation firewall, a DTN, and a perfSONAR measurement node. The core network switch/router is connected to the border router at 10G, initially. Funding was approved for the 100G upgrade of the segment from the research core switch to the border router, and to the AMPATH International Exchange Point. Figure 1 shows a representation of the improved science network and SC-DMZ. New research core routers were deployed to support

10/40/100G speeds. The SC-DMZ core was initially connected to the border router at 10G. Prior to the completion of the network deployment, the network connection between the core SC-DMZ router and the border router was upgraded to 100G. The core routers are connecting to the SC-DMZ building routers at 100G. PerfSONAR measurement points were deployed at border router, data center, and building distribution points. Note that the firewall is not inline between the border router and the research core network router. This is important to remove friction to science flows. By design, the default route in the SC-DMZ for all traffic is through the Palo Alto firewall to allow for non-research flows. Traffic bypassing the firewall and using the 100G, friction free link is explicitly defined for high-volume, bulk data transfers. All infrastructure devices (switches, routers, IPS, and firewalls) send syslog, NetFlow, and threat information to Splunk. The purpose of the Splunk server is to correlate and block malicious traffic via northbound APIs (Carasso 2012). It can also reroute traffic based on performance decisions via northbound APIs. The enterprise network core also connects to the border router. A logically partitioned Palo Alto firewall is connected inline between the enterprise network core router and the border router. It performs intrusion prevention, Network address translation (NAT) server, and firewall functions. The SC-DMZ is passively inspecting research packets, whereas on the enterprise path, the IPS actively inspects packets inline.

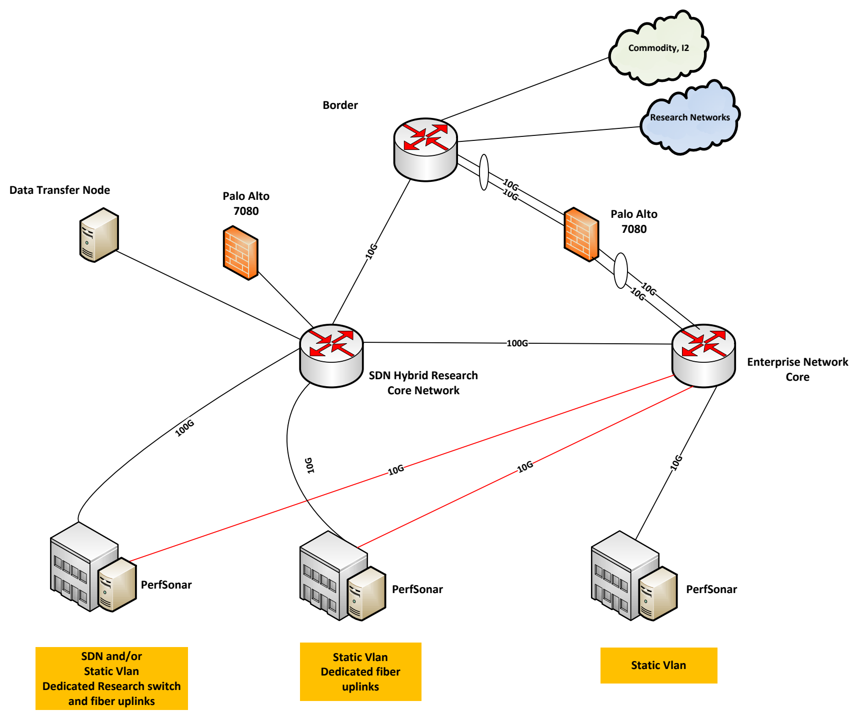

Figure 1 Science Network with support for 100G

Edge SDN switches consist of OpenFlow 1.3, enabled with 48 10G access ports and two 100G uplinks. Figure 2 represents three distinct cases for researchers to connect to the SC-DMZ. Case 1 represents the locations that have a dedicated, physical SC-DMZ router. Case 2, on the left, represents a private researcher-owned Ethernet switch that is connected to a SC-DMZ router (labeled ScDMZ router in Figure 2). Flows utilize a dedicated 100G link from the building where the researcher's lab is located to the SC-DMZ core router. In Case 3, on the right, the researcher connects to the enterprise switch uplinked to the enterprise router on a SC-DMZ VLAN. The connection to the SC-DMZ infrastructure is made at the building distribution point. The uplink between the enterprise switch and enterprise router is shared between science flows and enterprise traffic. The link between the enterprise router and SC-DMZ router is exclusively for researchers' SC-DMZ device traffic. Cases 2 and 3 leverage enterprise switches and routers at the distribution layer. Case 2 has a dedicated fiber uplink to the SC-DMZ core currently at 10G and separate fiber uplink to the enterprise core. Edge SDN switches were deployed in seven locations throughout the campus where research laboratories are located and at the FIU HPC cluster. Activation of the 100G science network and SC-DMZ was completed in August 2018.

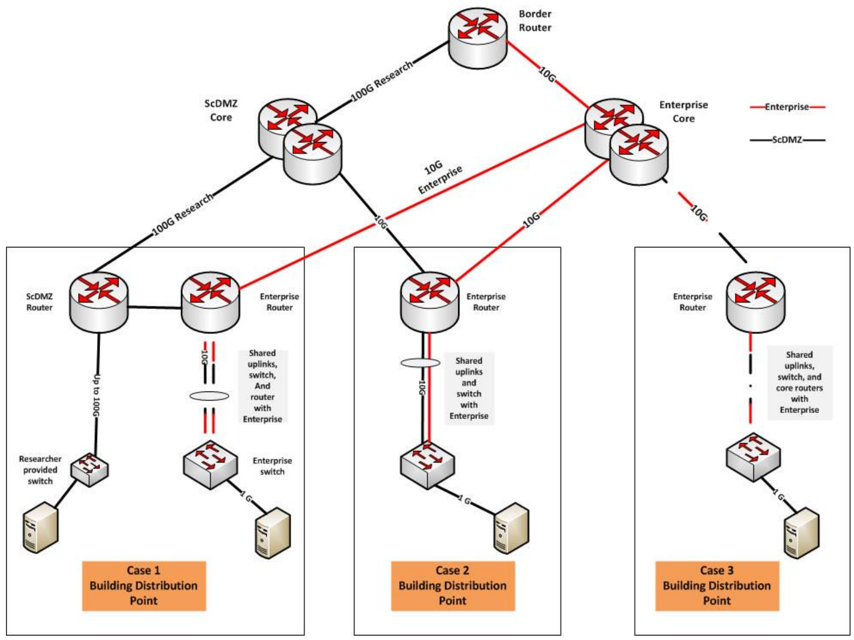

Figure 2 Options for researchers to connect to the Science network

Summary of shared and dedicated cyberinfrastructure on the Science DMZ

Table 4 summarizes by category the CI that are now connected to the SC-DMZ network at FIU. Table 4 also identifies if the CI is shared or dedicated.

	Shared	Dedicated
Compute	FIU HPC, VMWare environment	GIS Lab. Desktops, XNAT server
Storage	DDN GPFS storage for HPC Infinidat storage for XNAT, GIS and WOW	
DTN	HPC DTN	WOW DTN, FIONA DTN
File servers		GIS File Server, WOW File Server
Network	SC-DMZ Network Routers, PerfSONAR Nodes	
Data Generating Instruments		MRI Machine, MRI Control Workstation, WOW Wind Tunnel, WOW Control Desktops and Sensors

Table 4. Shared and dedicated CI on the Science DMZ

Researchers connected to the Science DMZ

This section presents the use cases for each of the science drivers that were connected to the 100G science network and SC-DMZ. It contains a brief description of the end-to-end scenario, describing the workflow of the science application, and identifying the actors (including the CI) in the workflow, and a diagram to represent the science workflow.

The Neuroinformatic and Brain Connectivity (NBC) Laboratory

There are two end-to-end dataflow scenarios for this use case as shown in Figure 3. First, projects that are part of the ABCD study have the following endpoints: MRI machine, MRI workstation, FIONA DTN, and

ABCD Data Analysis & Informatics Center (DAIC). Second, all projects, including projects that are part of ABCD, that generate DICOM data have the following endpoints: MRI machine, MRI workstation, XNAT server, HPC cluster, and personal computer. Each scan follows the Digital Imaging and Communications in Medicine (DICOM) international standard for the communication and management of medical imaging information and related data and is about 15 GB in size. The MRI machine, MRI workstation, and FIONA workstation are located at the FIU MRI facility at CIS. The XNAT server and HPC cluster are located in the FIU Data Center. The ABCD DAIC is located at University of California San Diego (UCSD). In addition to having two different workflows for this use case, the data that is transmitted is also different. The data that is transmitted to UCSD is the raw unprocessed data and is about 150GB in size for each scan. The data that is transmitted to the XNAT server at FIU is in the DICOM format. A DICOM file consists of a header and image data. The information within the header is standardized by a series of tags. By extracting data from these tags, one can access important information regarding the patient or subject.

The science workflow process steps are as follows: The MRI workstation controls the MRI machine and collects data from the MRI machine. Once the scan is complete, a technician manually transmits the DICOM data from the MRI workstation to the XNAT server in the Data Center. If the scan is part of the ABCD project, then the technician manually transmits the raw unprocessed data to the FIONA DTN. The data that resides on FIONA is manually transferred to UCSD at regular intervals. Data that resides on the XNAT server is further processed and analyzed on the FIU HPC and the researchers' personal computers. The bulk data flow in this use case was entirely migrated to the science network (blue lines in Figure 3). If a researcher accesses, processes data or analyzes the results on his/her own personal computer in the office or off-campus, the data flows through the enterprise network (red lines). This design allows for large data sets to be moved through the new science network while also providing researchers access to the data through the enterprise network. This ensures security and availability of the CI and data.

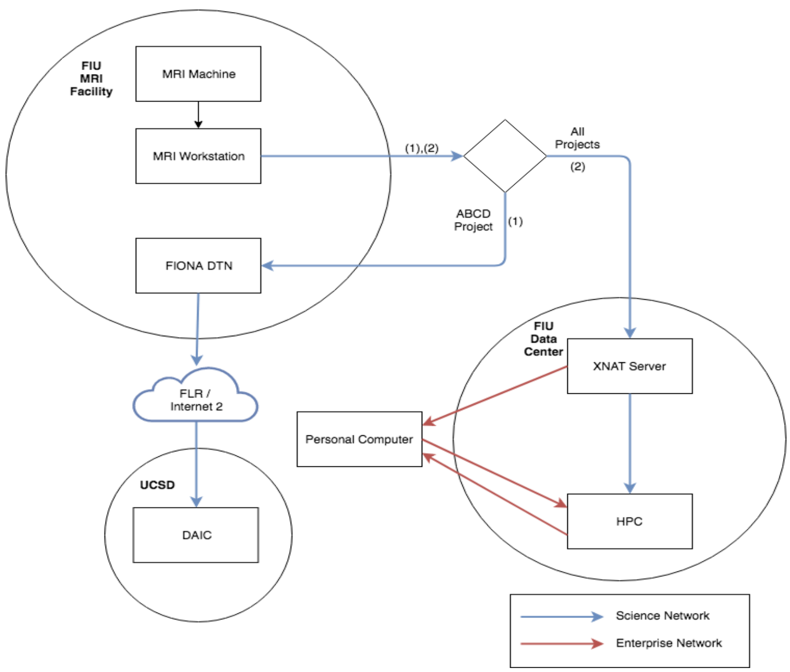

Figure 3. Endpoints of MRI science flows

The Geographic Information Systems (GIS) Center

All GIS Center projects have the following end-to-end scenario and endpoints: data repository, laboratory desktops, GIS file servers, HPC cluster, and personal computer. Data is either downloaded from an online repository (satellite and remote sensors) or physically collected on-site. The GIS lab desktops are located in the FIU Library building. The file servers that are mapped on the laboratory desktops and HPC cluster are located in the FIU Data Center. Figure 4 shows a representation of the science workflow. The bulk data flow in this use case is entirely migrated to the science network. The researchers have the option to use a local computer in the office or off-campus to access and process the data over the enterprise network. This ensures security and availability of the CI and data.

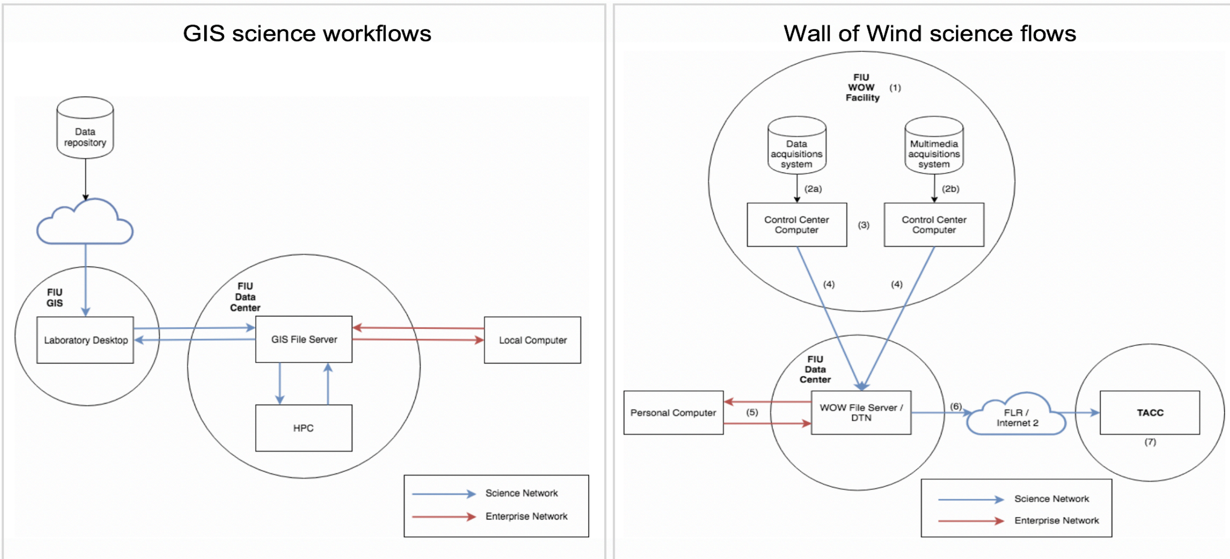

Figure 4. End points for GIS and Wall of Wind science flows

The Wall of Wind (WOW) facility

The WOW federal projects have the following end-to-end scenario and endpoints (Figure 4): data acquisitions system (sensors and probes), multimedia acquisitions system (video and photography equipment), control center computers to collect and preprocess data from the data and multimedia acquisition systems, file server and DTN, PC for processing the data, and TACC. The data and multimedia acquisition systems are located in the wind tunnel and the computers that collect the data are located in the control center at the WOW facility in the FIU Engineering Center. The WOW file server, DTN, and storage are located at FIU Data Center. The bulk data flow in this use case is entirely migrated to the science network. As with the other use cases, office and off-campus access was enabled through the enterprise network.

Additional migration of CI in the FIU data center to the new science network included several HPC critical servers hosted on non-redundant aging VMWare and KVM servers. The critical virtual servers included the HPC LDAP server, SAFE LDAP server, three license servers, two jump stations for Intelligent Platform Management Interface (IPMI) access to all machines, and HPC visualization portal. All these servers were migrated to two new redundant UCS Cisco servers that provide redundancy and mitigate against a single point of failure that could occur if any one of the servers failed. These hosts are connected to the new science network and will serve as the new HPC / Research VM environment. The HPC DTN is on a physical server that is shared by all HPC users to move data to and from the HPC, from their workstations, or from other storage devices to and from any other Globus endpoints they have access to. The HPC DTN was successfully migrated to the new science network. Other science application drivers in high-energy physics, optical astronomy, climate and weather prediction, and computer science (e.g., cloud computing and network simulation) were successfully connected to the science network during an earlier project called FlowSurge (Ibarra et al. 2012). Upon the FIU Science DMZ research project completion, performance measurements were taken to compare the performance of the SC-DMZ network with the enterprise network. As expected, the SC-DMZ network performed better across all metrics.

Conclusion

Two research groups, WoW and NBC, were successfully connected to the SC-DMZ network by the end of 2018. The GIS group required specific CI security upgrades, which could lead to a lengthy change of the research processes and new CI environment, and it was not connected. The qualitative approach used to research the science flows in this paper could help other universities and research labs with similar science instruments that are considering the Science DMZ approach for friction-free data transfer over the network. Future plans include SC-DMZ 100G network outreach activities among other FIU research teams whose research has the potential to generate a significant amount of data can benefit from such an infrastructure.

Acknowledgment

This Science DMZ project is supported by the National Science Foundation through the CC*DNI Engineer: A CI Engineer Lead for a Hybrid Campus Network project (NSF Award #1541402) and by the FIU World's Ahead Research Experiences for Students (WA-RES) project. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

REFERENCES

Aly, A. M., Bitsuamlak, G., and Chowdhury, A. G. 2011. "Florida International University's Wall of Wind: A Tool for Improving Construction Materials and Methods for Hurricane-Prone Regions," in *Vulnerability, Uncertainty, and Risk: Analysis, Modeling, and Management*. pp. 352-359.

Carasso, D. 2012. *Exploring Splunk*. CITO Research New York, USA.

Charmaz, K., and Belgrave, L. L. 2007. "Grounded Theory," *The Blackwell encyclopedia of sociology*).

Dalkey, N., and Helmer, O. 1963. "An Experimental Application of the Delphi Method to the Use of Experts," *Management science* (9:3), pp. 458-467.

Dart, E., Rotman, L., Tierney, B., Hester, M., and Zurawski, J. 2013. "The Science Dmz: A Network Design Pattern for Data-Intensive Science," *SC '13: Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, pp. 1-10.

Fattal, A. B. 2018. "NIH Releases First Dataset from Unprecedented Study of Adolescent Brain Development."

Hanemann, A., Boote, J. W., Boyd, E. L., Durand, J., Kudarimoti, L., Łapacz, R., Swany, D. M., Trocha, S., and Zurawski, J. 2005. "Perfsonar: A Service Oriented Architecture for Multi-Domain Network Monitoring," *International conference on service-oriented computing*: Springer, pp. 241-254.

Ibarra, J., Morgan, H. L., Johnson, E., and Drake, M. 2012. "Cc-Nie Network Infrastructure: Flowsurge: Supporting Science Data Flows Towards Discovery, Innovation and Education." Florida International University: National Science Foundation (NSF).

Koziol, Q. 2014. "Texas Advanced Computing Center," in *High Performance Parallel I/O*. Chapman and Hall/CRC, pp. 121-130.

Krefft, J. V. 2009. "MIUS News: Maps and Imagery User Services @ Fiu Green Library," *Map and User Imagery Services, Florida International University* (3:2).

Leatherman, S. P., Gan Chowdhury, A., and Robertson, C. J. 2007. "Wall of Wind Full-Scale Destructive Testing of Coastal Houses and Hurricane Damage Mitigation," *Journal of Coastal Research*), pp. 1211-1217.

Marcus, D. S., Olsen, T. R., Ramaratnam, M., and Buckner, R. L. 2007. "The Extensible Neuroimaging Archive Toolkit," *Neuroinformatics* (5:1), pp. 11-33.

Pordes, R., Petravick, D., Kramer, B., Olson, D., Livny, M., Roy, A., Avery, P., Blackburn, K., Wenaus, T., and Würthwein, F. 2007. "The Open Science Grid," *Journal of Physics: Conference Series*: IOP Publishing, p. 012057.

Ramsey, D. 2014. "FIONA: Innovative Network Appliance for Big Data." University of California in San Diego.

Stewart, C. A., Simms, S., Plale, B., Link, M., Hancock, D. Y., and Fox, G. C. 2010. "What Is Cyberinfrastructure," *Proceedings of the 38th annual ACM SIGUCCS fall conference: navigation and discovery*: ACM, pp. 37-44.

Towns, J., Roskies, R., Gaither, K., Peterson, G., Wilkins-Diehr, N., Boisseau, J., Kovatch, P., and Andrews, P. 2011. "XSEDE: Extreme Science and Engineering Discovery Environment." University of Illinois at Urbana-Champaign: National Science Foundation (NSF).

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T., Bucholz, R., Chang, A., Chen, L., Corbetta, M., and Curtiss, S. W. 2012. "The Human Connectome Project: A Data Acquisition Perspective," *Neuroimage* (62:4), pp. 2222-2231.