**Title**
Sinclair Community College and OARnet Requirements Analysis Report



**Author**
Zurawski, Jason

**Publication Date**
2022-08-23

Peer reviewed

# Sinclair Community College and OARnet Requirements Analysis Report

*September 15th, 2022*

## Disclaimer

# Sinclair Community College and OARnet Requirements Analysis Report

*September 15th, 2022*

The Engagement and Performance Operations Center (EPOC) is supported by the National Science Foundation under Grant No. 1826994.

ESnet is funded by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research. Carol Hawk is the ESnet Program Manager.

ESnet is operated by Lawrence Berkeley National Laboratory, which is operated by the University of California for the U.S. Department of Energy under contract DE-AC02-05CH11231.

This is a University of California, Publication Management System report number LBNL-2001475[1].

[1] https://escholarship.org/uc/item/0th6628f

## Participants & Contributors

Intisar Bashir, OARnet
Mark Fullmer, OARnet
John Graham, OARnet
Robert Gutendorf, Sinclair Community College
Kyle Jones, Sinclair Community College
Scott McCollum, Sinclair Community College
Justin Morgan, Sinclair Community College
Angela Nixson, OARnet
Anthony Ponder, Sinclair Community College
Scott Reinemann, Sinclair Community College
Dana Rogers, OARnet
Seth Schwartz, Sinclair Community College
Alison Scott, OARnet
Pankaj Shah, OARnet
Andrew Shepherd, Sinclair Community College
Denis Walsh, OARnet
Michael Watkins, Sinclair Community College
Maryann Zbydnowski, OARnet
Ann Zimmerman, OARnet
Jason Zurawski, ESnet

## Report Editor

Jason Zurawski, ESnet: zurawski@es.net

# Contents

# 1 Executive Summary

## Deep Dive Review Purpose and Process

EPOC uses the Deep Dive process to discuss and analyze current and planned science, research, or education activities and the anticipated data output of a particular use case, site, or project to help inform the strategic planning of a campus or regional networking environment. This includes understanding future needs related to network operations, network capacity upgrades, and other technological service investments. A Deep Dive comprehensively surveys major research stakeholders' plans and processes in order to investigate data management requirements over the next 5–10 years. Questions crafted to explore this space include the following:

- How, and where, will new data be analyzed and used?
- How will the process of doing science change over the next 5–10 years?
- How will changes to the underlying hardware and software technologies influence scientific discovery?

Deep Dives help ensure that key stakeholders have a common understanding of the issues and the actions that a campus or regional network may need to undertake to offer solutions. The EPOC team leads the effort and relies on collaboration with the hosting site or network, and other affiliated entities that participate in the process. EPOC organizes, convenes, executes, and shares the outcomes of the review with all stakeholders.

## This Review

Between October 2021 and March 2022 staff members from the Engagement and Performance Operations Center (EPOC) met with researchers and staff from Sinclair Community College (SCC) and OARnet with the purpose of performing a Deep Dive into scientific and research drivers. The goal of this activity was to help characterize the requirements for a number of campus use cases, and to enable cyberinfrastructure support staff to better understand the needs of the researchers within the community.

## This review includes case studies from the following campus stakeholder groups:

- GIS Education
- Automotive Education
- Unmanned Aerial Systems
- Automation & Control of Robotics
- Cyber Security, Data Analytics, & Computer Science
- Information Technology

Material for this event included the written documentation from each of the profiled research areas, documentation about the current state of technology support, and a write-up of the discussion that took place via e-mail and video conferencing.

The case studies highlighted the ongoing challenges and opportunities that SCC has in supporting a cross-section of established and emerging research use cases. Each case study mentioned unique challenges which were summarized into common needs.

### The review produced several important findings and recommendations from the case studies and subsequent virtual conversations:

- The localized storage needs at SCC will increase in the future, and the Information Technology support infrastructure should begin to plan for this eventuality.

- The WAN connectivity to SCC (main and remote campuses) should be increased in the coming years to keep pace with demand in partnership with OARnet.

- A wholistic view of network and information security for SCC is suggested, particularly ensuring that there are controls put in place to validate potential sensitive data that is shared.

- There are a number of potentially computationally intensive use cases at SCC, but it must be established what sort of computation they require, and where it can be leveraged (e.g., local or remote).

- Several SCC use cases have an intense growth potential for data storage and processing at SCC in the coming years. Due to the collaboration space, it is recommended that SCC, OARnet, and other stakeholders, evaluate the best way to create an environment that would facilitate ease of sharing this data both at the networking layer, as well as the data mobility hardware and software layers.

- The adoption of portal software, e.g., easy to use mechanisms to search, display, and retrieve data sets, may be useful for some of the use cases that have large numbers of collaborators.

- Federated identity to enable data sharing may be required in the future years to manage collaborations.

# 2 Deep Dive Findings & Recommendations

The deep dive process helps to identify important facts and opportunities from the profiled use cases. The following outlines a set of findings from the SCC and OARnet Deep Dive that summarize important information gathered during the discussions surrounding case studies, and possible ways that could improve the CI support posture for the campus:

- The localized storage needs for the profiled science, research, and education use cases at SCC will increase in the future, and the Information Technology support infrastructure should begin to plan for this eventuality. It is recommended that a future group storage solution for campus should be easy to use (e.g., have APIs and mount points that can integrate into the user's workflow), enable sharing (e.g., via data transfer hardware the use of data mobility tools such as Globus), and be scalable to add more capacity over time.

- The WAN connectivity to SCC (main and remote campuses) should be increased in the coming years to keep pace with demand. OARnet should work with SCC, and help them consider options from a regional capability standpoint.

- A wholistic view of network and information security for SCC is suggested, particularly ensuring that there are controls put in place to validate potential sensitive data that is shared. Some data (e.g., GIS/UAV, Automotive, etc.) may be sensitive and require CUI support. A CUI environment may be required for some data sensitive use cases either maintained by SCC, or provided by a third party.

- There are a number of potentially computationally intensive use cases SCC, but it must be established what sort of computation they require: Cloud, Grid, or HPC. Some of these can be acquired from outside entities (e.g., OSC, OSG, XSEDE, or commercial entities), and others may be better to be operated by SCC for the benefit of campus users.

- The GIS/UAV use cases, as discussed in this document, have the largest growth potential for data storage and processing at SCC in the coming years. Due to the collaboration space, it is recommended that SCC, and other stakeholders, evaluate the best way to create an environment that would facilitate ease of sharing this data. A VRF between facilities, operated by OARnet, may be a potential network solution, coupled to data transfer hardware, software, and storage resources that all collaborators may be able to integrate into federated identity.

- The adoption of portal software, e.g., easy to use mechanisms to search, display, and retrieve data sets, may be useful for some of the use cases that have large numbers of collaborators (e.g., UAV, GIS). The use of portals will depend heavily on how "public" the data is designed to be, with more "private" use cases encouraged to adopt modern data mobility tools.

- Federated identity to enable data sharing may be required depending on the granularity of some of the collaborators.

The following are recommendations for future activities:

- It is recommended that discussion between SCC and OARnet occur to discuss community best practices surrounding how research funds that enter the university for projects that may require IT can be socialized and planned for. EPOC can be a resource to assist with these discussions.

- It is recommended that SCC evaluate different mechanisms to provide local storage solutions that can scale for use cases, with an emphasis on prioritizing simple mechanisms to integrate with workflows and enable sharing with collaborators

- It is recommended that OARnet work to re-establish a regional perfSONAR testing mesh across the state of Ohio and with regional partners and peers across the United States. EPOC can be a resource to assist with this work.

- It is recommended that OARnet and SCC work with EPOC to establish a baseline data transfer performance number via the Data Mobility Exhibition project.

- It is recommended that SCC and OARnet propose an ongoing engagement strategy to keep conversations between users, IT, and leadership ongoing as requirements change. Establishing a WG or Council will allow this effort to live on, and scale, over time. Can focus on setting policy standards for all users. EPOC can serve as a resource to assist with these.

- It is recommended that SCC explore technology intensive use cases, such as UAV and GIS, on a more routine basis to ensure that requirements are being integrated correctly at the campus layer. Due to the larger than normal technology demands, these discussions may require integration with regional partners.

- It was observed that some federal collaborations (e.g., those with WPAFB) may require higher levels of network security, information security, and/or special peering. It is recommended that SCC and OARnet explore these in greater detail.

# 3 Process Overview and Summary

## 3.1 Campus-Wide Deep Dive Background

Over the last decade, the scientific community has experienced an unprecedented shift in the way research is performed and how discoveries are made. Highly sophisticated experimental instruments are creating massive datasets for diverse scientific communities and hold the potential for new insights that will have long-lasting impacts on society. However, scientists cannot make effective use of this data if they are unable to move, store, and analyze it.

The Engagement and Performance Operations Center (EPOC) uses the Deep Dives process as an essential tool as part of a holistic approach to understand end-to-end research data use. By considering the full end-to-end research data movement pipeline, EPOC is uniquely able to support collaborative science, allowing researchers to make the most effective use of shared data, computing, and storage resources to accelerate the discovery process.

EPOC supports five main activities

- Roadside Assistance via a coordinated Operations Center to resolve network performance problems with end-to-end data transfers reactively;
- Application Deep Dives to work more closely with application communities to understand full workflows for diverse research teams in order to evaluate bottlenecks and potential capacity issues;
- Network Analysis enabled by the NetSage monitoring suite to proactively discover and resolve performance issues;
- Provision of managed services via support through the Indiana University (IU) GlobalNOC and our Regional Network Partners; and
- Coordinated Training to ensure effective use of network tools and science support.

Whereas the Roadside Assistance portion of EPOC can be likened to calling someone for help when a car breaks down, the Deep Dive process offers an opportunity for broader understanding of the longer term needs of a researcher. The Deep Dive process aims to understand the full science pipeline for research teams and suggest alternative approaches for the scientists, local IT support, and national networking partners as relevant to achieve the long-term research goals via workflow analysis, storage/computational tuning, identification of network bottlenecks, etc.

The Deep Dive process is based on an almost 15-year practice used by ESnet to understand the growth requirements of Department of Energy (DOE) facilities[2]. The EPOC team adapted this approach to work with individual science groups through a set of structured data-centric conversations and questionnaires.

[2] https://fasterdata.es.net/science-dmz/science-and-network-requirements-review

## 3.2 Campus-Wide Deep Dive Structure

The Deep Dive process involves structured conversations between a research group and relevant IT professionals to understand at a broad level the goals of the research team and how their infrastructure needs are changing over time.

The researcher team representatives are asked to communicate and document their requirements in a case-study format that includes a data-centric narrative describing the science, instruments, and facilities currently used or anticipated for future programs; the advanced technology services needed; and how they can be used. Participants considered three timescales on the topics enumerated below: the near-term (immediately and up to two years in the future); the medium-term (two to five years in the future); and the long-term (greater than five years in the future).

The case study process tries to answer essential questions about the following aspects of a workflow:

- ***Research & Scientific Background***—an overview description of the site, facility, or collaboration described in the Case Study.
- ***Collaborators***—a list or description of key collaborators for the science or facility described in the Case Study (the list need not be exhaustive).
- ***Instruments and Facilities: Local & Non-Local***—a description of the network, compute, instruments, and storage resources used for the science collaboration/program/project, or a description of the resources made available to the facility users, or resources that users deploy at the facility or use at partner facilities.
- ***Process of Science***—a description of the way the instruments and facilities are used for knowledge discovery. Examples might include workflows, data analysis, data reduction, integration of experimental data with simulation data, etc.
- ***Computation & Storage Infrastructure: Local & Non-Local***—The infrastructure that is used to support analysis of research workflow needs: this may be local storage and computation, it may be private, it may be shared, or it may be public (commercial or non—commercial).
- ***Software Infrastructure***—a discussion focused on the software used in daily activities of the scientific process including tools that are used locally or remotely to manage data resources, facilitate the transfer of data sets from or to remote collaborators, or process the raw results into final and intermediate formats.
- ***Network and Data Architecture***—description of the network and/or data architecture for the science or facility. This is meant to understand how data moves in and out of the facility or laboratory focusing on local infrastructure configuration, bandwidth speed(s), hardware, etc.
- ***Resource Constraints***—non-exhaustive list of factors (external or internal) that will constrain scientific progress. This can be related to funding, personnel, technology, or process.
- ***Outstanding Issues***—Listing of any additional problems, questions, concerns, or comments not addressed in the aforementioned sections.

At a physical or virtual meeting, this documentation is walked through with the research team (and usually cyberinfrastructure or IT representatives for the organization or region), and an additional discussion takes place that may range beyond the scope of the original document. At the end of the interaction with the research team, the goal is to ensure that EPOC and the associated CI/IT staff have a solid understanding of the research, data movement, who's using what pieces, dependencies, and time frames involved in the Case Study, as well as additional related cyberinfrastructure needs and concerns at the organization. This enables the teams to identify possible bottlenecks or areas that may not scale in the coming years, and to pair research teams with existing resources that can be leveraged to more effectively reach their goals.

## 3.3 SCC and OARnet Deep Dive Background

Between October 2021 and March 2022, EPOC organized a Deep Dive in collaboration with SCC to characterize the requirements for several key science drivers. The representatives from each use case were asked to communicate and document their requirements in a case-study format. These included:

- GIS Education
- Automotive Education
- Unmanned Aerial Systems
- Automation & Control of Robotics
- Cyber Security, Data Analytics, & Computer Science
- Information Technology

## 3.4 Organizations Involved

The Engagement and Performance Operations Center (EPOC) was established in 2018 as a collaborative focal point for operational expertise and analysis and is jointly led by Indiana University (IU) and the Energy Sciences Network (ESnet). EPOC provides researchers with a holistic set of tools and services needed to debug performance issues and enable reliable and robust data transfers. By considering the full end-to-end data movement pipeline, EPOC is uniquely able to support collaborative science, allowing researchers to make the most effective use of shared data, computing, and storage resources to accelerate the discovery process.

The Energy Sciences Network (ESnet) is the primary provider of network connectivity for the U.S. Department of Energy (DOE) Office of Science (SC), the single largest supporter of basic research in the physical sciences in the United States. In support of the Office of Science programs, ESnet regularly updates and refreshes its understanding of the networking requirements of the instruments, facilities, scientists, and science programs that it serves. This focus has helped ESnet to be a highly successful enabler of scientific discovery for over 25 years.

Indiana University (IU) was founded in 1820 and is one of the state's leading research and educational institutions. Indiana University includes two main research campuses and six regional (primarily teaching) campuses. The Indiana University Office of the Vice President for Information Technology (OVPIT) and University Information Technology Services (UITS) are responsible for delivery of core information technology and cyberinfrastructure services and support.

Sinclair Community College (SCC) Sinclair Community College was founded in 1887, and has 5 regional campus locations throughout Southwest Ohio. The three strategic priorities of the college are: Alignment with the economy and social needs of the southwestern Ohio citizens, communities, and businesses that will be served with educational programs and services. Growth for the number of college students, and the rates of success that benefit the businesses and community organizations will receive more training and development services. Equity in the student body, faculty, and staff that overall reflect the holistic diversity of the region.

The Ohio Academic Resources Network (OARnet) The Ohio Academic Resources Network (OARnet) is a state-funded IT organization that provides member organizations with intrastate networking, virtualization and cloud computing solutions, advanced videoconferencing, connections to regional and international research networks and the commodity Internet, colocation services and emergency web-hosting. The OARnet network is a dedicated, statewide, high-speed fiber-optic network that serves Ohio K-12 schools, college and university campuses, academic medical centers, public broadcasting stations and state and local/state government. OARnet delivers technology-based solutions that reduce costs, increase productivity and improve customer service – and has done so since 1987. As a division of the Ohio Department of Higher Education's Ohio Technology Consortium, OARnet serves Ohio's education, health care, public broadcasting and government communities.

## 4 SCC Case Studies

SCC presented a number of use cases during this review. These are as follows:

- GIS Education
- Automotive Education
- Unmanned Aerial Systems
- Automation & Control of Robotics
- Cyber Security, Data Analytics, & Computer Science
- Information Technology

Each of these Case Studies provides a glance at research activities, the use of experimental methods and devices, the reliance on technology, and the scope of collaborations. It is important to note that these views are primarily limited to current needs, with only occasional views into the event horizon for specific projects and needs into the future. Estimates on data volumes, technology needs, and external drivers are discussed where relevant.

## 4.1 SCC GIS Education

*Content in this section authored by Scott Reinemann, SCC, Department of Education*

### 4.1.1 Use Case Summary

Scott Reinemann's teaching and research work focuses on ways to leverage Geographic Information Systems (GIS) students understand the concepts of mapping, spatial analysis, computer aided design, and geographic understanding. This is coupled with more traditional geography education approaches for students both in the classroom and in the field. To accomplish these goals, a field course has been developed that allows students to use Unmanned Aerial Systems (UAS) and Global Positioning Systems (GPS) to accurately map the physical environment, specifically a rock glacier in Nevada.

For most of these efforts a data life cycle is important, however Scott and his colleagues are still learning the best approaches to adopt. Currently the data created is stored on personal and work computers, then later it is uploaded to a cloud sharing service for dissemination and long-term storage. In the classroom environment, a cloud-based Learning Management Service (LMS), D2Learn, is used to disseminate and collect data from students.

### 4.1.2 Collaboration Space

- ***External Data Used***:
    - Local, regional, and national governmental dataset for geographic information
    - Census Demographic Data
    - Colleagues' data stored on Cloud Services
    - Student created Data on LMS
- ***Locally Produced Data***:
    - Students
    - Colleagues from within and external to SCC
    - Public shared by request.

### 4.1.3 Instruments & Facilities

- ***Local***:
    - Computer Lab (~20 stations)
    - Student work run Computer Lab (~15 stations)
    - Microsoft OneDrive Servers (for local data)
- ***Remote***:
    - Tree Core Processing Lab (Columbus, OH)
    - Sediment Analysis Lab (Athens, GA)

### 4.1.4 Data Narrative

Data for educational purpose is mainly stored on local computers at SCC, or within cloud-based services. This data is created on a semester-by-semester, basis and usually updated as the semester progresses. When the research for a class or particular assignment is needed, the updates to the stored data is not always done in a convenient and repeatable manner. For instance, there may be lots of manual changes to the data set with little automation.

#### 4.1.4.1 Data Volume & Frequency Analysis

The current volume of data produced for this educational activity amounts to GB, produced on a weekly basis.

#### 4.1.4.2 Data Sensitivity

There are no sensitive aspects to this research data, external data resources are public and local data has no restrictions.

#### 4.1.4.3 Future Data Volume & Frequency Analysis

The expected volume of data produced for this educational activity in the future will amount to GB, produced on a weekly basis.

### 4.1.5 Technology Support

The following sections outline the technology support supporting this use case.

#### 4.1.5.1 Software Infrastructure

- ***Microsoft Office (intuitionally paid for):*** software for the collection, storage and dissemination of data.
- ***ESRI ArcGIS Pro (commercial paid yearly):*** Software used for the mapping of research data along with the software used for teaching map creation to students.
- ***Microsoft OneDrive & Google Shared Drive:*** tool used for the collection and dissemination of data.
- ***Microsoft SharePoint:*** Used to store and transfer data within a team of educators and student workers.
- ***Desire2Learn (LMS):*** tool used to share knowledge with students and help them progress in their academic career.

#### 4.1.5.2 Network Infrastructure

Very little is known about networking, beyond what is used local to the SCC environment.

#### 4.1.5.3 Computation and Storage Infrastructure

Very little is known about computing and storage, beyond what is used local to the SCC environment and the cloud services that are used for storage. Being able to leverage more computing resources for some of the GIS work is desirable, but unknown how to do this at this time.

In terms of future storage requirements, the use case values:

- Additional capacity
- Speed of the storage infrastructure (to match computing and networking needs)
- Access for students and faculty

- Sharing with collaborators

ArcGIS files and databases can become quite large, so the ability to have large capacity and access them quickly is important.

#### 4.1.5.4 Data Transfer Capabilities

The use case regularly transfers dataset in the order of 2-10 GB, and this is typically done through a cloud service as the share point. This can take 2-3 hours to both upload and then download the data.

### 4.1.6 Internal & External Funding Sources

- Western National Parks Association (Set to expire June 2022)
- Great Basin National Heritage Area (Set to expire Aug. 2023)

### 4.1.7 Resource Constraints

The most serious constraint to resources is the ability to access and share data. Typically, a user trying to exchange information with a local or remote colleague must utilize a cloud service to broker the exchange, and sometimes they must use a "personal" account versus those that are provided by the university. Addressing these issues would help to streamline the process of science and education.

### 4.1.8 Ideal Data Architecture

An easier to use repository to handle data sharing is desired, as this is a major area of friction within the process of science and education. This would allow easier uploads, sharing, and search.

### 4.1.9 Outstanding Issues

There are no outstanding issues to report at this time.

## 4.2 SCC Automotive Education

*Content in this section authored by Justin Morgan, SCC, Automotive Department*

### 4.2.1 Use Case Summary

The automotive department serves regional automotive repair facilities with interns and graduates to work as automotive service technicians. With advancements in electrification and autonomy, the automotive department has trained other automotive instructors throughout the nation on advance driver assist systems and is preparing a three day summer institute on battery electric vehicle theory/operation, service and diagnosis. The small amount of research from the automotive department focuses on those two subjects.

### 4.2.2 Collaboration Space

Collaboration with other colleges with automotive departments.

### 4.2.3 Instruments & Facilities

There are no instruments or facilities that require networking, computational or storage support to report at this time.

### 4.2.4 Data Narrative

There are no forms of external data that are shared with collaborators.

#### 4.2.4.1 Data Volume & Frequency Analysis

The current volume of data produced for this educational activity amounts to GB, produced on a yearly basis. This data is stored within appliances used by the department.

#### 4.2.4.2 Data Sensitivity

There are minimal sensitive aspects to this research data. Some technology does produce proprietary data, but it not shared.

#### 4.2.4.3 Future Data Volume & Frequency Analysis

The expected volume of data produced for this educational activity in the future will amount to GB, produced on a yearly basis. This data is stored within appliances used by the department.

### 4.2.5 Technology Support

This use case does not leverage technology support. Most, if not all, of the tools are designed to work by connecting point to point to a specific vehicle.

### 4.2.6 Internal & External Funding Sources

There are no sources of funding to report at this time. There have been some proposals made to the NSF to explore EV technology, but none have been awarded.

### 4.2.7 Resource Constraints

There are no reported resource constraints at this time.

#### 4.2.8 Ideal Data Architecture

There are no additional requests for data architecture at this time.

#### 4.2.9 Outstanding Issues

There are no outstanding issues to report at this time.

## 4.3 SCC Unmanned Aerial Systems

*Content in this section authored by Andrew Shepherd, SCC, Unmanned Aerial Systems*

### 4.3.1 Use Case Summary

SCC conducts applied research and development related to technology development and testing, real-world operations and activities explorations, airspace integration operations, materials, modeling and simulation, data analytics, and other efforts with respect to Unmanned Aerial Systems (UAS). This often involves flight operations; remote sensing data acquisition, transfer, and analysis; and networked telemetry for operational environment situational awareness. Activities may occur solely on the SCC campus, linked to sites in Southwest Ohio or statewide, across the U.S., or even to international partners.

### 4.3.2 Collaboration Space

SCC is active on the Ohio Federal Research Network, which includes partnerships with other academic institutions also performing work on UAS technology (e.g., The Ohio State University, University of Cincinnati) and various companies. The college is also a member of the Federal Aviation Administration (FAA) ASSURE UAS Center of Excellence, which includes 25 member academic institutions in the U.S. and internationally. SCC is also a member of a proposal team to the National Science Foundation (NSF) that will create an Industry-University Cooperative Research Centers Program (IUCRC) focused on UAS, which if funded will begin in summer 2022 and link SCC to six other research universities in CO, UT, MI, VA, PA, and TX. Data is both generated by SCC and shared with these and other entities and provided to the college for use. The data for projects may include raw or processed outputs and can range from kilobytes into the terabytes.

### 4.3.3 Instruments & Facilities

SCC leverages desktop computers in the UAS data analytics lab (26) and UAS Simulation lab (11), UAS operations lab (3), plus staff and project laptops and integrated computers in the Mobile Ground Control Station (3), Tactical Ground Control Station (3), and UAS Operations Station (1). There are also various ground control station computers dedicated to individual UAS. Some storage occurs on external solid state drives, on SCC servers, or on cloud options including Google. However, a better solution is required for data storage, archiving, access, and processing as the volume of data increases due to the passage of time, increased operations, and larger data products.

### 4.3.4 Data Narrative

Collected remote sensed data (e.g., video, still images, multi-spectral data, etc.) from UAS may be leveraged in a raw state or processed into products for analysis. Some examples include 3D terrain and building models, maps, environmental and agricultural assessment, and computer vision video analysis. Analysis may use installed or cloud software products including Pix4D and RFView among others. Artificial intelligence analysis is becoming more common in the industry and requires significant computational and storage resources to train a system and produce viable results. Additionally, Live, Virtual, and Constructive (LVC) training and operational activities require robust networks, with telemetry and video inputs drawn from UAS, traditional aircraft, ground vehicles and personnel, simulators, and computer entities. LVC events can include local sites in Southwest Ohio, or participating sites across the U.S. or abroad. LVC leverages Simlat software, as well as other supporting video feed and analysis tools.

#### 4.3.4.1 Data Volume & Frequency Analysis

Normal collections may be in the MB to GB range, but multi-flight projects can have total sets in the multi-TB range.

#### 4.3.4.2 Data Sensitivity

Some data may be ITAR controlled. Same may fall under NDA. Some may involve human subjects for testing and research projects, primarily with human factors focuses. Some may have IP ramifications.

#### 4.3.4.3 Future Data Volume & Frequency Analysis

Typical collections may be in the MB or GB range, with some total project datasets already extending into the TB range. Overtime, large multi-year projects that could occur in the future may reach hundreds of TB or even PB.

### 4.3.5 Technology Support

The following sections outline the technology support supporting this use case.

#### 4.3.5.1 Software Infrastructure

- CAD software (e.g. AutoCAD, SolidWorks, Fusion360)
- Simlat
- Pix4D
- Systems Tool Kit (STK)
- RFView
- FLIR, GeoSLAM Connect
- Google Earth
- Jira

#### 4.3.5.2 Network Infrastructure

The primary use case is to download data from UAS after flight operations have completed, but some use cases can leverage "real time" transmission back to flight stations.  Once data is retrieved from device, it is possible to leverages cellular connections for UAS field operations, including in support of LVC events. Where possible, wired internet from the field or partner sites are leveraged to connect back to SCC for data transfer or telemetry connections. In some locations, OARnet has now been connected (Springfield Beckley Municipal Airport, National Center for Medical Readiness, etc.) but it has been difficult to arrange or coordinate use so the standard Internet is often relied on and sufficient for most cases. The UAS department typically works directly with SCC IT for these issue with limited direct interaction with other departments on campus.

At this time there are no strong real-time data access requirements, but future iterations (and partners) may want this.

#### 4.3.5.3 Computation and Storage Infrastructure

This is an area in need of improvement for UAS at SCC. Storage is a combination of limited server access, external solid state drives, and Google could storage. A much larger, more organized, and professionalized system for storing, searching, accessing, transferring, processing, and sharing data, both with SCC personnel or students and external partners is needed. External partners or stakeholders may be on OARnet, otherwise in Ohio, throughout the U.S., or international. Additionally, all activities should have a cyber security component to ensure data are secure from access by those not authorized or manipulation.

At this time the analysis done for this work is not highly parallelizable, so the addition of more computing isn't a factor, but faster computing (and higher/faster memory) could be impactful.

#### 4.3.5.4 Data Transfer Capabilities

In some cases, files can be in the KB, MB, or GB range, which has allowed transfer through Google Drive, though not ideal. Larger projects that may be hundreds of GB or even TBs have had to be transferred through solid state drive and either driven or shipped between sites. Obviously, these processes are not fast or sufficient for expanding data sets and issues related to backup, security, and timeliness.

### 4.3.6 Internal & External Funding Sources

To date, UAS departmental capital or operating budgets, as well as SCC IT budgets, have supported UAS data requirements.

### 4.3.7 Resource Constraints

The current setup for UAS data is functional but not ideal. Eventually, data volumes and complexity of operations will exceed the current approach's ability to cope. Additionally, as projects become more sensitive, there are greater concerns related to cyber security and backup.

### 4.3.8 Ideal Data Architecture

More high-end hardware and software for onsite or deployed data processing and storage. Additional cloud-based storage and data processing capabilities for use by SCC and partners. Archival capabilities. Robust backup. Any solutions must also meet applicable regulations and data management plans as may be required for various projects.

### 4.3.9 Outstanding Issues

Most do not appreciate the volume of data acquired from UAS remote sensing and flight operations. These data must be collected, processed, stored, and maintained over time to be of most value. They must be searchable and accessible from anywhere, with sufficient security measures in place. The requirements will only grow in the coming years and it is imperative to address current needs and prepare for future requirements now to meet research, education, training, and consulting deliverables.

## 4.4 SCC Automation & Control of Robotics

*Content in this section authored by Michael Watkins, SCC, Automation Robotics Electronics*

### 4.4.1 Use Case Summary

The Automation and Control Technology with Robotics program builds knowledge in the application of electrical and mechanical skills for developing, installing, programming and troubleshooting the complex machinery found in the modern manufacturing and Supply Chain environments. Integrate robotic systems into manufacturing processes using robotics teach pendant programming, electromechanical system installation procedures, preventive maintenance, and calibration to ensure and maintain operational readiness. The use case leverages robotics, electronics, PLC labs.

### 4.4.2 Collaboration Space

Students, instructors, manufacturing resources. Other colleges offering similar courses

### 4.4.3 Instruments & Facilities

Robotics lab, electronics lab, computer programming software. PLC labs. Equipment connected together on Ethernet. Vision systems. Various electronic measurement equipment.

### 4.4.4 Data Narrative

Mainly computers for software and storing programs.

#### 4.4.4.1 Data Volume & Frequency Analysis

The current volume of data produced for this educational activity amounts to GB, produced on a daily basis.

#### 4.4.4.2 Data Sensitivity

There are no sensitive aspects to this research data.

#### 4.4.4.3 Future Data Volume & Frequency Analysis

The expected volume of data produced for this educational activity in the future will amount to GB, produced on a daily basis.

### 4.4.5 Technology Support

The following sections outline the technology support supporting this use case.

#### 4.4.5.1 Software Infrastructure

- PLC software
- Vision system software
- Robotics programming software
- Multisim

#### 4.4.5.2 Network Infrastructure
Ethernet connections between hardware and PC's.  Namely some automation control devices (PLCs and HMIs) that will be used for data collection and storage.

#### 4.4.5.3 Computation and Storage Infrastructure
Just local storage on computers

#### 4.4.5.4 Data Transfer Capabilities
MB size, share drive, USB memory

### 4.4.6 Internal & External Funding Sources
There are no sources of funding to report at this time.

### 4.4.7 Resource Constraints
There are no reported resource constraints at this time.

### 4.4.8 Ideal Data Architecture
There are no additional requests for data architecture at this time.

### 4.4.9 Outstanding Issues
There are no outstanding issues to report at this time.

## 4.5 SCC Cyber Security, Data Analytics, & Computer Science

*Content in this section authored by Kyle Jones, SCC, Cyber security, data analytics, computer science*

### 4.5.1 Use Case Summary

The Computer Science and Information Technology Department is involved in many National Science Foundation (NSF) and National Security Administration (NSA) projects. With these projects, we have a lot of creative curriculum and materials that we are required to share out as dissemination materials to other colleges in the state of Ohio and the United States. The department has created many extra learning materials and projects that can be implemented in other institutions. The department heavily utilizes a virtual learning environment maintained by SCC. This environment is used in a lot of our college credit plus programs and by partner institutions. This environment acts as a cyber range with class and curriculum built in to be implemented into the classroom.

### 4.5.2 Collaboration Space

Students, Faculty, and surrounding High schools and Community college. SCC is currently building out new data analytics programs and, in the future, we may be looking for potential ways to host this information as we help high schools develop new data analytics programs across the state.

### 4.5.3 Instruments & Facilities

Most of the data now that were using is currently located on the SCC network located in building 13. This includes our virtualized environment known as NetLabs that is also utilized at other local high schools in Columbus State.

### 4.5.4 Data Narrative

Currently our department is involved in several grants building out several new learning modules. In the future as we build out our new dad analytics programs, we may be looking at ways to host large amounts of data in order to have students practice real world scenarios.

#### 4.5.4.1 Data Volume & Frequency Analysis

The current volume of data produced for this educational activity amounts to TB, produced on a weekly basis.

#### 4.5.4.2 Data Sensitivity

There are no sensitive aspects to this research data.

#### 4.5.4.3 Future Data Volume & Frequency Analysis

The expected volume of data produced for this educational activity in the future will amount to TB, produced on a weekly basis.

### 4.5.5 Technology Support

The following sections outline the technology support supporting this use case.

#### 4.5.5.1 Software Infrastructure
Currently the computer science and information technology department use several technologies. These technologies include PowerBi, Tableau, and NetLabs. These are used in the educational level for the individual students in the classroom.

#### 4.5.5.2 Network Infrastructure
No additional information about the network infrastructure can be provided.

#### 4.5.5.3 Computation and Storage Infrastructure
No additional information about the computation and storage infrastructure can be provided.

#### 4.5.5.4 Data Transfer Capabilities
No additional information about the data transfer capabilities can be provided.

### 4.5.6 Internal & External Funding Sources
Expanding the Data Analytics Technician Pipeline from High School into College and High Demand Jobs in Southwest Ohio- Start - 7/1/2021 end -6/30/2024

### 4.5.7 Resource Constraints
There are no reported resource constraints at this time.

### 4.5.8 Ideal Data Architecture
There are no additional requests for data architecture at this time.

### 4.5.9 Outstanding Issues
There are no outstanding issues to report at this time.

## 4.6 SCC Information Technology

*Content in this section authored by Robert Gutendorf, SCC, Department of Information Technology*

### 4.6.1 Use Case Summary

This case study represents the entire SCC campus. SCC IT is centralized, with all 5 remote campuses connected back to Building 13 in Dayton, OH. SCC IT partners with multiple ISPs, including Spectrum, AT&T, Crown Castle, and Cinci Bell. OARnet is a huge partner and collaborator for SCC IT. SCC IT uses Microsoft Office 365 and Google Drive for cloud storage and replication, and still maintains a large number of sites and services hosted on prem. The college also utilizes a variety of software as a service applications.

### 4.6.2 Collaboration Space

perfSONAR, ShakeNet are the only ones that SCC IT is familiar with. There are no computing or storage collaborations (e.g., XSEDE, OSC) to report at this time.

### 4.6.3 Capabilities & Special Facilities

Information Technology provides datacenter and telecom closet rack space as needed on the Dayton campus as well as remote campuses. The SCC IT mission is to support all campus activities, including research. Research tends to have higher bandwidth and storage needs than our day to day activities. SCC IT provides and supports campus interconnectivity as well as internet connectivity 24x7. We have interconnectivity through OARNet, MVECA (Miami Valley Educational Computer Association) which supports schools and local governments, and Cinci Bell (Courseview Campus). Peering arrangement with Wright State - NETLab support for Columbus state. Returning Citizens with the Ohio State Correction Facilities, Ponitz High School, ZVRS Video Phones for the hearing impaired, Schneider Electric Access Control Infrastructure throughout campus.

### 4.6.4 Technology Narrative

The following sections outline the technology support supporting this use case.

#### 4.6.4.1 Network Infrastructure

The diagrams describe the WAN and LAN configuration; 40Gbps through the core, and some of the edge. SCC IT is currently upgrading edge uplinks to 40Gbps. 10Gbps fiber uplink to OARNet via Crown Castle. SCC IT pays for commodity internet by the Mbps. All circuits provide to SCC are Layer 2, and higher level peering arrangements are maintained by SCC. All network connectivity to remote campuses egresses through the main campus connection in Dayton.

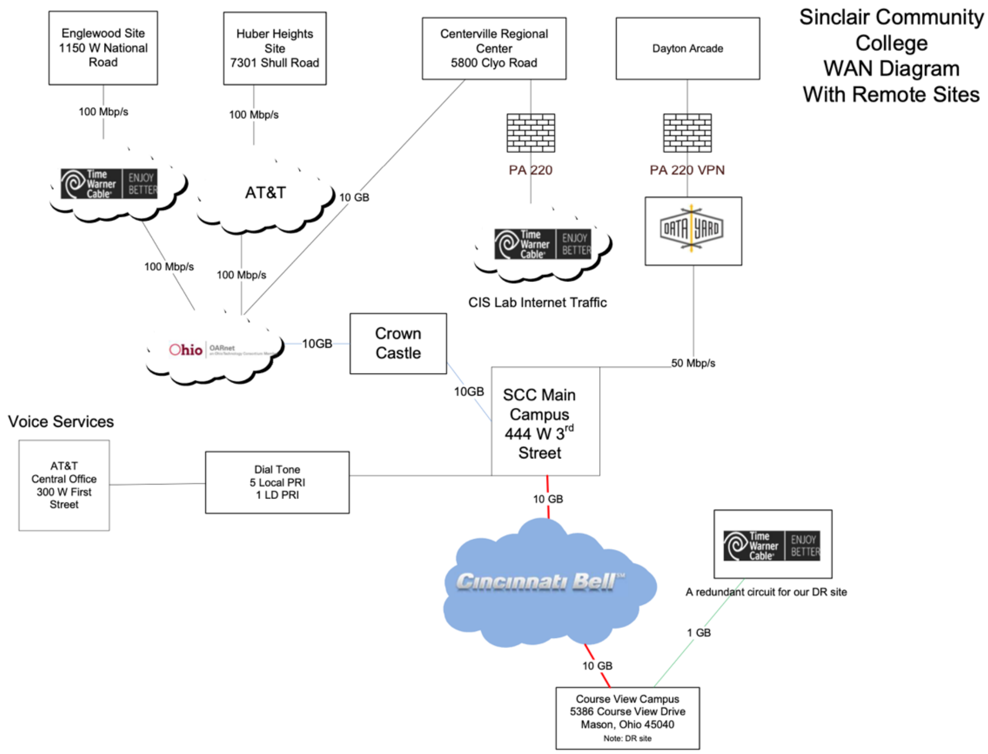


*Figure 1 – SCC WAN Diagram*

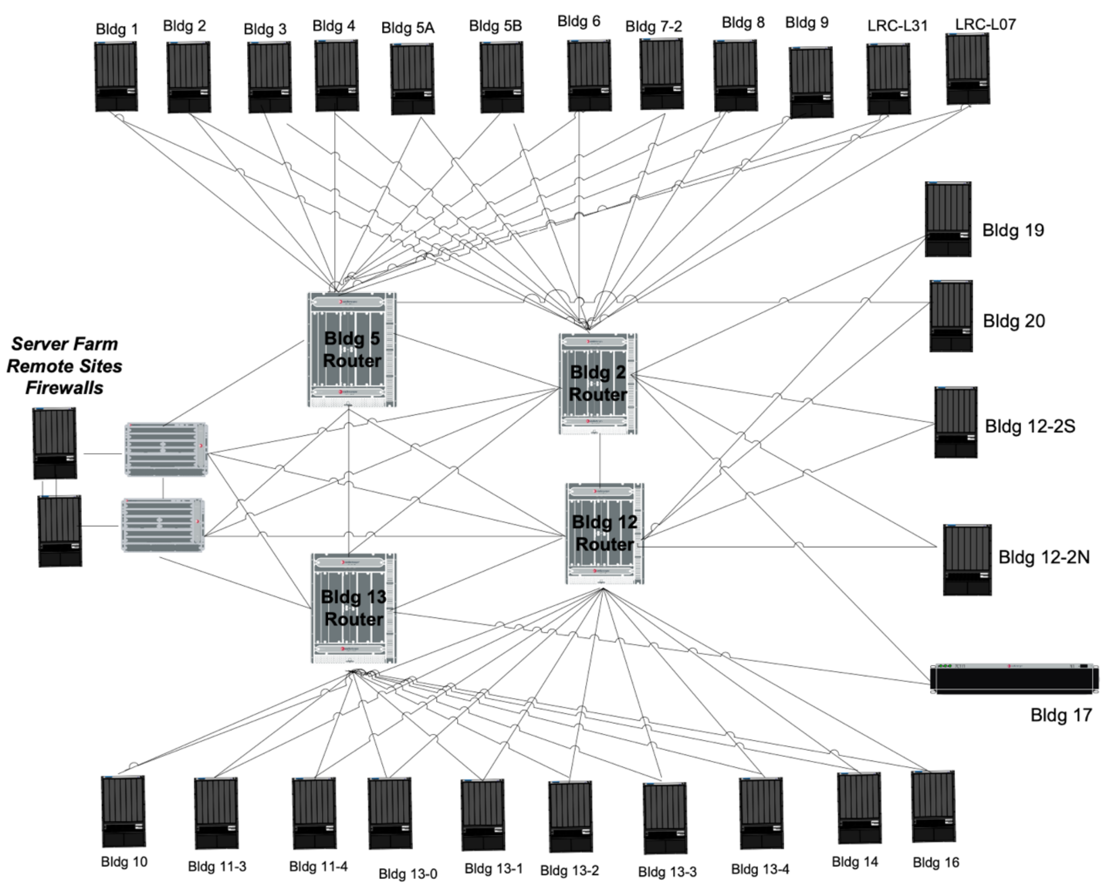

*Figure 2 – SCC (Dayton) LAN Diagram*

### 4.6.4.2 Computation and Storage Infrastructure

SAS Data and Visual Analytics Infrastructure, Elastic Stack Cluster. ~ 380 TB of available storage on-site via SAN and Hyper-Converged Infrastructure. Possible future need of ~ 400 TB to 1 PB storage for UAS data on-site. Users file tickets to acquire storage and computing (VM) resources that can be integrated into research and education use cases.

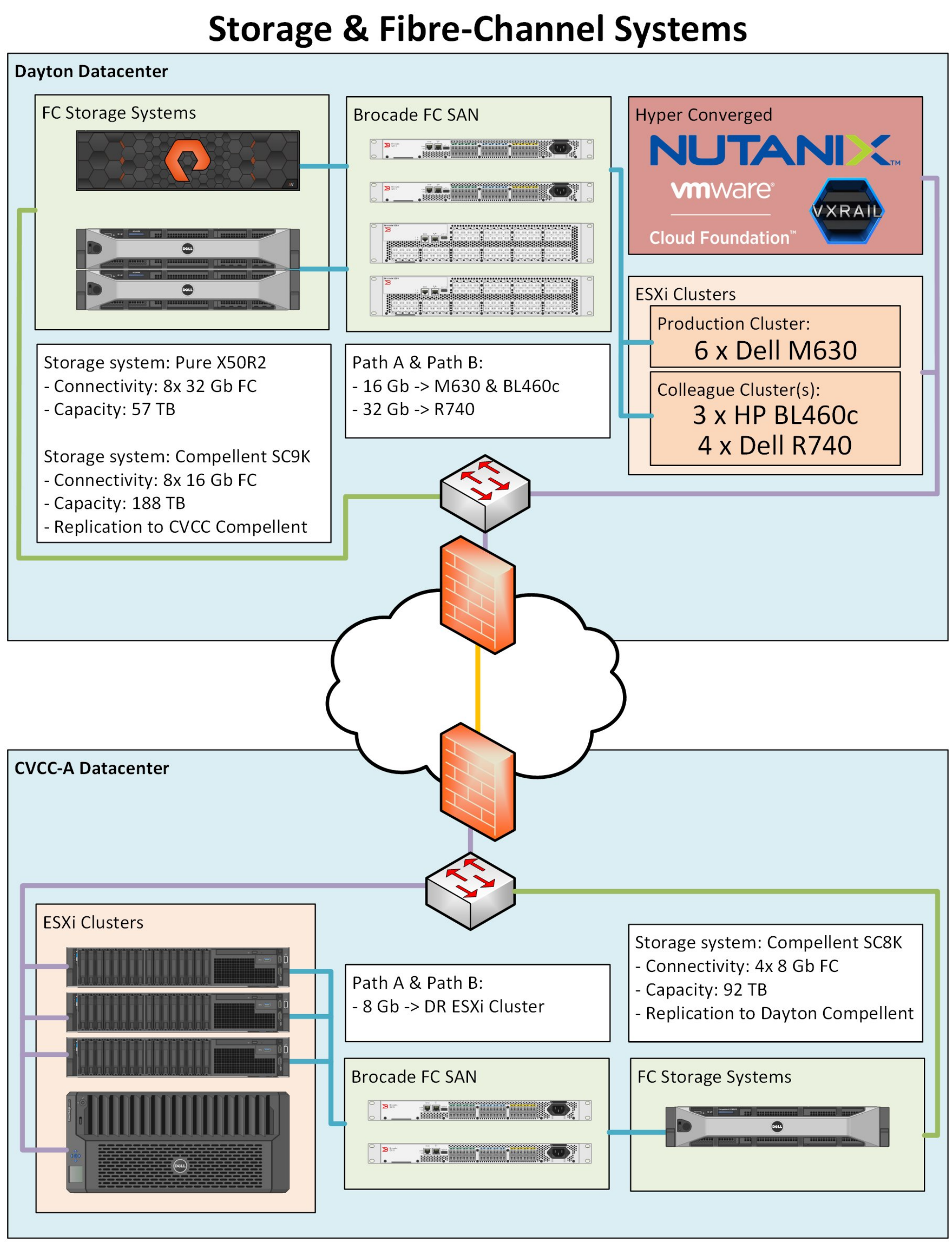


*Figure 3 – SCC Storage & Fibre-Channel System Diagram*

#### 4.6.4.3 Network & Information Security

WAN: DHS External Penetration (NESUS) Scans Weekly, Palo Alto Firewalls with WildFire and Threat Protection subscriptions Edge: Extreme NAC Port Policies, VLAN Islands, 802.1x authentication, Global Protect VPN connectivity Device: Malware Bytes, Microsoft Defender, Internal Vulnerability Scans done weekly on new server builds and post software installs, MineMeld Threat Intelligence Sharing Tool for Palo Alto, Mimecast Email Security, ELK Stack, Extreme Networks RADAR rogue AP detection, Running as least privileged user on PCs.

#### 4.6.4.4 Monitoring Infrastructure

- perfSONAR
- OBM Server Monitoring (Operations Bridge Manager)
- SiteScope
- NETSNMP
- Mrtg
- Prtg
- Elastic Stack Share data
- PowerShell scripts
- Extreme Management Console - user for internal SLAs for service availability reporting

SCC IT can always use additional assistance measuring and monitoring the infrastructure. SCC IT would be interested in discussing some possibilities with OARnet.

#### 4.6.4.5 Software Infrastructure

SAS Data and Visual Analytics, UAS uses a variety of tools SCC IT is not completely familiar with, MAT Labs (Physics), ARC GIS (Geography and Sociology).

Most departments will run their own software infrastructure, without SCC IT being aware of the requirements.  On occasion SCC IT will help to perform evaluations or security audits.

### 4.6.5 Organizational Structures & Engagement Strategies

#### 4.6.5.1 Organizational Structure

## INFORMATION TECHNOLOGY

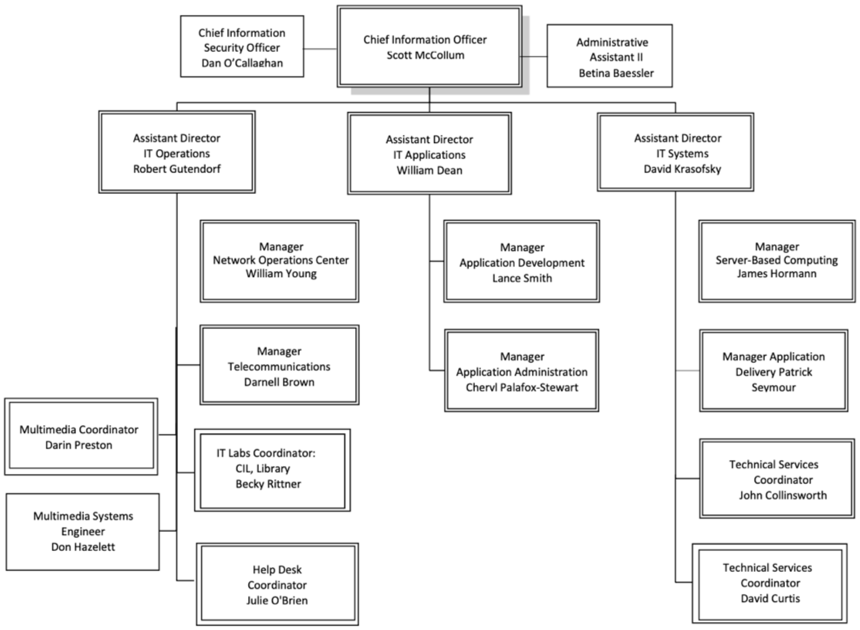


*Figure 4 – IT Organizational Chart*

## IT OPERATIONS

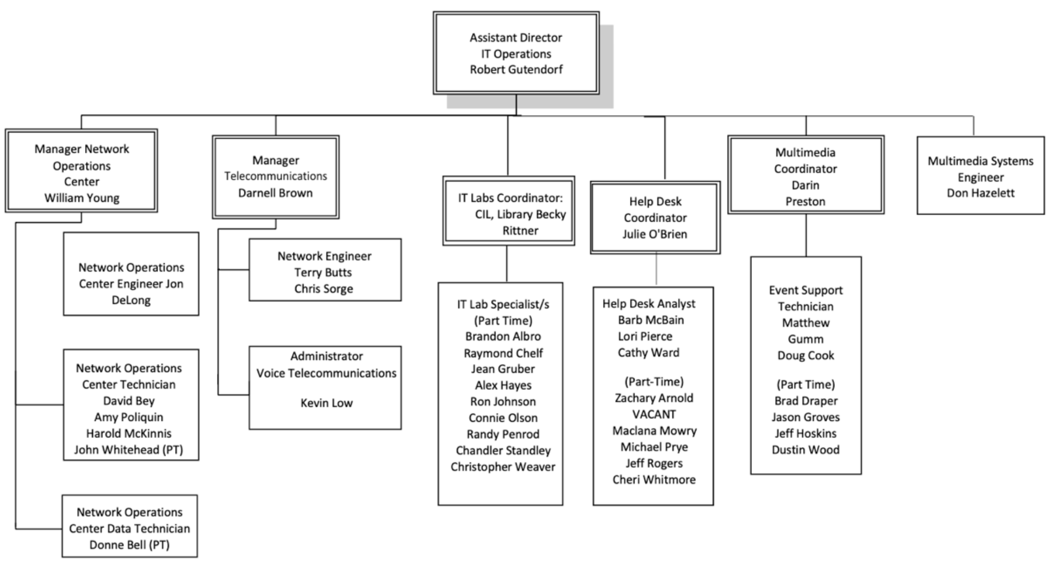


*Figure 5 – IT Operations Organizational Chart*

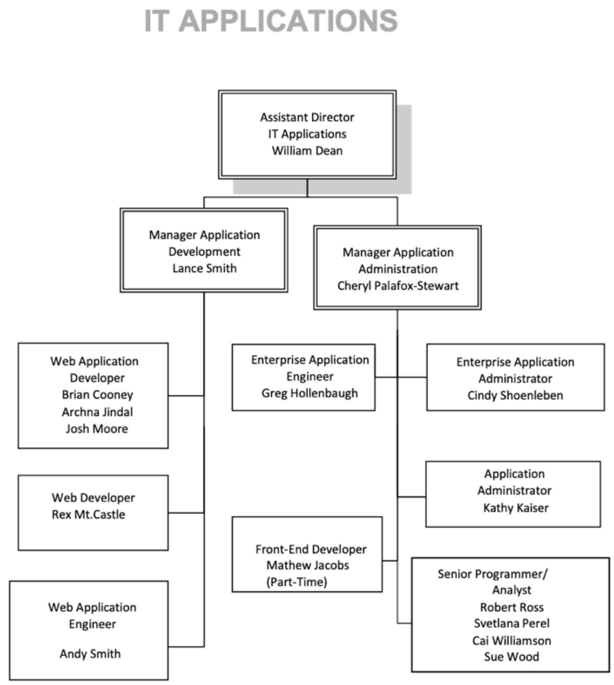


*Figure 6 – IT Applications Organizational Chart*

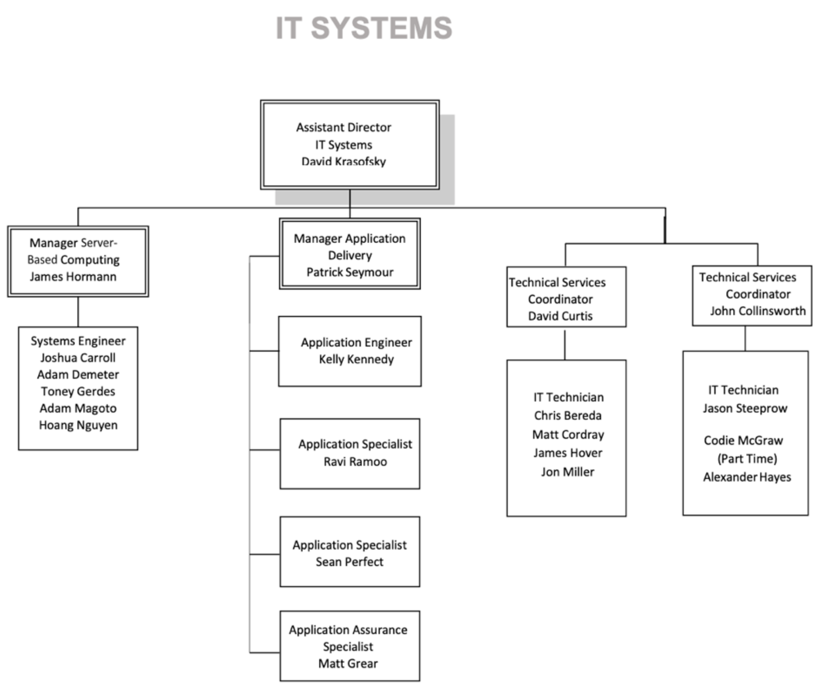


*Figure 7 – IT Systems Organizational Chart*

#### 4.6.5.2 Engagement Strategies

Object Storage Research with UAS, perfSONAR use case discussions with OARNet,

### 4.6.6 Internal & External Funding Sources

There are no sources of funding to report.

### 4.6.7 Resource Constraints

Insufficient Storage Space, Insufficient Bandwidth between remote sites, Concerns about proper security around sharing data, proper balance for freedom to research but secure production network,  other server/infrastructure resource limitations that aren't known at this time

### 4.6.8 Outstanding Issues

Proper communication around research/grants and technology support needed prior to grant award or start of services.  In years past SCC IT has not known about technology being received for grants until the technology arrives on site.  SCC IT definitely has current network constraints between remote sites that would need to be addressed, 100 MB connectivity to Huber and Englewood.

# Appendix A – OARnet Network Diagrams

The OARnet backbone network, shown in Figure 8, features more than 5,500 miles of fiber-optic cable, currently featuring six major rings that extend across the state to most of Ohio's population.

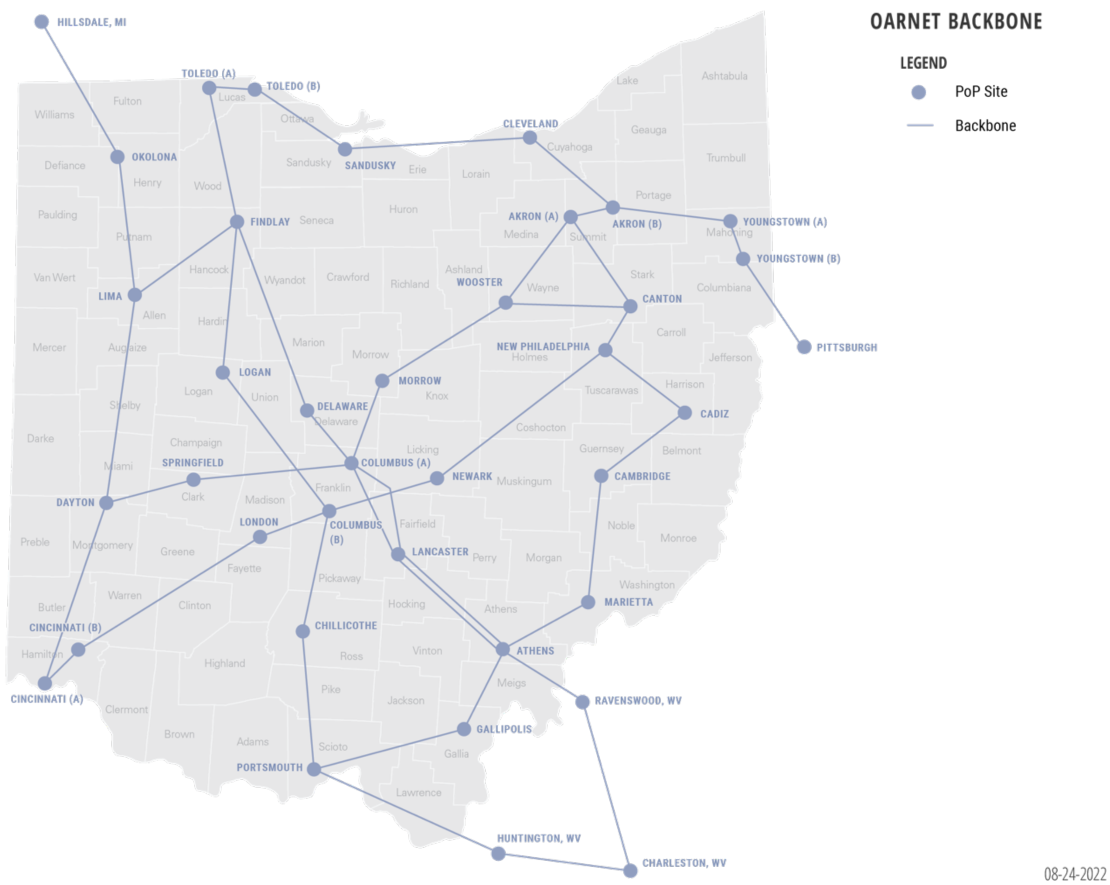


*Figure 8 – OARnet Backbone*

Beyond the in-state reach of the OARnet network, partnerships with other Regional Optical Networks (RONs) extend OARnet connectivity to Michigan's Merit research and education network, the Pittsburgh Supercomputing Center, and the OmniPoP research collaboration network in Chicago.