The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
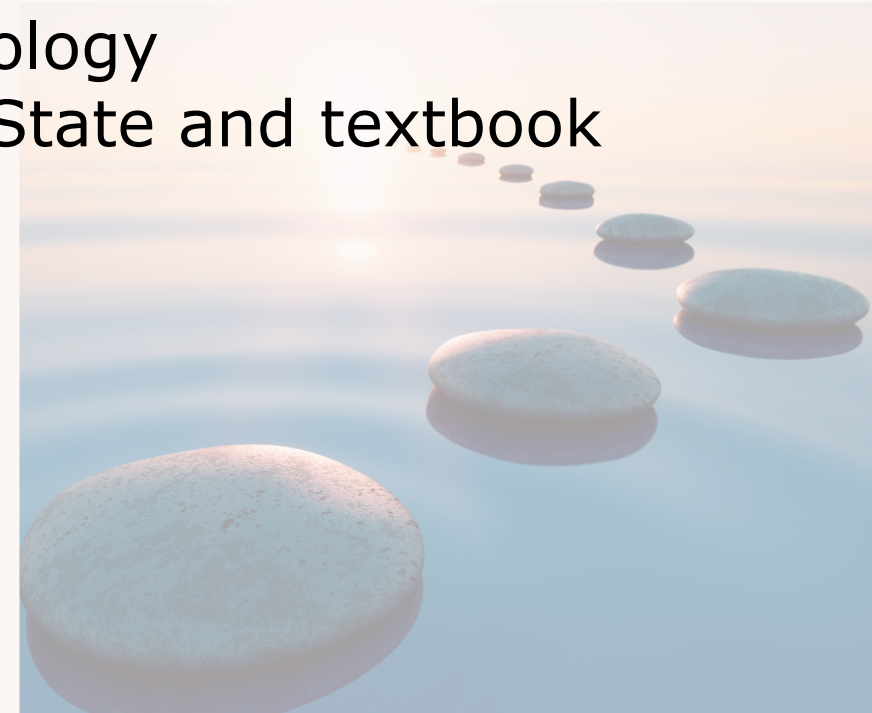
Gretta Kellogg (gretta.kellogg@psu.edu)
Asst. Director, Center for AI/ML to Industry
Penn State University

AIMI

PennState
Institute for Computational and Data Sciences

AIMI

My Research Computing Story…..

How I ended up in research computing through Enterprise IT and Computer Mediated Communication.

- Graduate Student – IBM SUR Grant for online learning platform development & feedback study
- Programmer w/Teaching and Learning in Technology
- Course developer for Online Learning for Penn State and textbook publishers
- Enterprise IT for Microsoft AD & Linux systems
- IT Project Manager
- Research Program Mgr (2 Genomic Centers)
- ACI-REF
- Director Research Facility
- Asst. Director of Research Center

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Research Engagement: the PSU RISE Team

RISE is a team of ICDS computational scientists and software engineers that merge advanced computing concepts to a broad spectrum of research projects.

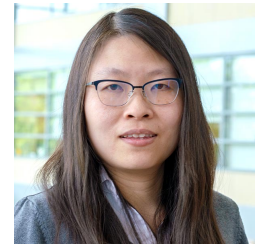Currently 13 Full Time PhD and Masters level domain specialists

https://www.icds.psu.edu/computing-services/rise/

Dr. Chuck Pavloski
RISE Team Lead
Earth Sciences

Ron Tapia
RISE Engineer
Platform Computing

Clayton Colson
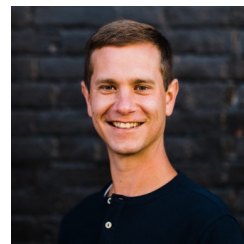RISE Engineer
Immersive Experiences

Dr. Danying Shao
RISE Engineer
Computational Biology
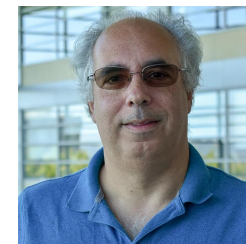
Dr. Simon Delattre
RISE Engineer
Machine Learning
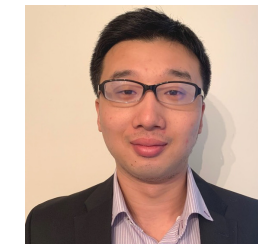
Dr. Patrick Dudas
Interim CIE Director
Data Visualization

Dave McLaughlin
RISE Engineer
Data Visualization

Diego Menendez
RISE Engineer
Software Engineering

Jeff Nucciarone
RISE Engineer
Parallelization

Dr. Weinan Chen
RISE Engineer
Materials Science

Dr. Justin Petucci
RISE Engineer
Machine Learning

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Sample Intake categories and questions:

Description
- Who is the researcher and or college/institute?
- What are the funding agencies involved?
- Can you describe your current research workflow?

Software
- What software do you use for accomplishing your research? What versioning requirements to you have?
- Do you typically run your software on your own workstations, local servers or HPC clusters? How long does the analysis run before completion ?
- What common software packages do you need for publications/presentations?

Networking
- What are your network latency or data transfer expectations?
- What is a typical data set size?
- What could you do differently if you had unlimited bandwidth?

Systems
- What operating systems are you using on workstations? Servers? equipment?
- What is a typical CPU time for individual experiments? A subset of analysis? Overall analysis for 1yr? 3years?

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Sample Intake categories and questions – files provided for your own customization and use.

PennState
Institute for CyberScience

Initial Questions to help us gather IT requirements for newly incoming faculty:

Our focus is to leverage our RISE team's collective experiences to design and manage complex research computing systems to accomplish their aims over the life cycle of their projects. Therefore we approach their needs by considering all of what their research may require from research networking to systems development, to equipment integration, automated analysis pipelines, data management and dissemination, software engineering and also storage and compute requirements. We will work with the faculty to help them develop sustainable and affordable workflows for conducting and managing their research computing needs that include feedback systems (whenever possible) for continuous quality improvement.

In-depth - Research Faculty Onboarding questions:

Description
- Who is the researcher and or college/institute?
- Where is the researcher connected? Any instruments?
- What are the funding agencies involved?
- Can you describe your current research workflow?

Software
- What software do you use for accomplishing your research? What versioning requirements to you have?
- Do you typically run your software on your own workstations, local servers or HPC clusters? How long does the analysis run before completion?
- Do you also build, test and distribute own software for your research analysis?
- What common software packages do you need for publication/presentation? MSOffice/Adobe Illustrator, etc?

Systems
- What operating systems are you using on workstations? Servers? equipment?
- What is a typical CPU time for individual experiments? A subset of analysis? Overall analysis for 1yr? 3years?
- What types of systems are you currently using or would like to leverage for your work? Workstations? Servers? Application platforms? Science Gateways? Virtual/cloud hosting or bringing hardware to be house on premise?

Personnel Support
- Are you providing IT professional support from your own lab to support your research? Do you need to utilize PSU support, sysadmins or programs or help with grant funding?
- What sort of training will your lab members need? What external training will be required?
- Would you be willing to discuss your of research workflow so we can capture your current compute/staffing needs?

Data Management
- Where is the data generated?
- Where is the data processed?
- Where is your data stored?
- Where is your data archived?
- What are the data storage estimates for 1 year? 3 yrs? 5?

Transfer Requirements?
- How is the data transferred?
- What application is used?
- What are your existing transfer rates?
- Does your research require datasets coming into Penn State? Going out from here?

Security
- What is the risk category of your data?
- Do you receive data from or collaborate with anyone outside of Penn State?
- Do you have a data management plan from your grant?
- Do you use or operate a data management portal?
- Can you describe your current security workflow?

Share this to them first in advance. Then share again in-person or over zoom

---

fx Research Faculty Onboarding Questions

Research Faculty Onboarding Questions

Description
- Who is the researcher and or college/institute?
- Where is the researcher connected? Any instruments?
- What are the funding agencies involved?
- Can you describe your current research workflow?

Software
- What software do you use for accomplishing your research? What versioning requirements to you have?
- Do you typically run your software on your own workstations, local servers or HPC clusters? How long does the analysis run before completion?
- Do you also build, test and distribute your own software for your research analysis?
- What common software packages do you need for publications/presentations? MSOffice/Adobe Illus...

Networking
- What ... latency or data transfer...
- ...data set size?
- ...could you do differ... limited bandwidth?
- Would you be willing... of research workflow so we can capture Network path of data tra...

Systems
- ...operating systems are you using on workstations? Servers? equipment?
- What is a typical CPU time for individual experiments? A subset of analysis? Overall analysis for 1yr? 3...
- What types of systems are you currently using or would like to leverage for your work? Workstations? Servers? Application platforms? Science Gateways? Virtual/cloud hosting or bringing har...

Personnel Support
- Are you providing IT professional support from your own lab to support your research? Do you need to utilize PSU support, sysadmins or programs or help with grant funding?
- What sort of training will your lab members need? What external training will be required?
- Are you writing or planning to write proposals that may change your current compute/storage/staffing...

Then you can collect information. Check off services and rates they need.

MSWord and Excel templates

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

Let's talk about how to organize the resources that you want to have available as you work with researchers.

CREDIT: Cornell University Center for Advanced Computing

PennState Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

Systems include consortia, platforms, and groups of service providers.

Systems

Service Providers

Service Providers are still large-scale national resources.

Technologies include many of the tools they leverage from exoscale to local research computing.
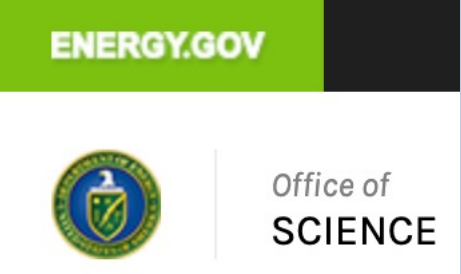
Technologies

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

NSF Cloud Computing

ACCESS

DOE Cloud Computing

ACER

Commercial Cloud Computing

Commodity

MISSING: Other examples like • NIH resources such as BioWulf, Helix and Helixweb
• Cloud resources such as CloudBank, CloudLab, GENI, emulab and Chameleon.

PennState
Institute for Computational
and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

ACCESS is an advanced computing and data resource supported by the National Science Foundation (NSF). It's a collection of integrated digital resources and services, including supercomputers, visualization, and storage systems.

ACCESS is made possible through these lead institutions and their partners — **Carnegie Mellon University**; **University of Colorado, Boulder**; **University of Illinois at Urbana-Champaign**; and **State University of New York at Buffalo**.

ACCESS Credits and Thresholds

Researchers have opportunities to request ACCESS allocations at four levels, which are described at the links in the table.

Allocation	Credit Threshold
Explore ACCESS	400,000
Discover ACCESS	1,500,000
Accelerate ACCESS	3,000,000
Maximize ACCESS	Not awarded in credits.

ACCESS replaced ~~XSEDE~~.

CREDIT: https://allocations.access-ci.org/

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

Why are there different "Credit Thresholds"?

ACCESS Credits and Thresholds

Researchers have opportunities to request ACCESS allocations at four levels, which are described at the links in the table.

Allocation	Credit Threshold
Explore ACCESS	400,000
Discover ACCESS	1,500,000
Accelerate ACCESS	3,000,000
Maximize ACCESS	Not awarded in credits.

CREDIT: https://allocations.access-ci.org/

PennState
Institute for Computational
and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

ENERGY.GOV

Advanced Scientific Computing Research (ASCR)

Office of
SCIENCE

Mission: Discover, develop, and deploy computational and networking capabilities to analyze, model, simulate, and predict complex phenomena important to the Department of Energy (DOE).

CREDIT:
https://science.osti.gov/ascr/Facilities/User-Facilities

Advanced Scientific Computing Research	**Environmental Management**
Basic Energy Sciences	**Fossil Energy**
Biological and Environmental Research	**Fusion Energy Science**
Cybersecurity, Energy Security, and Emergency Response	**High Energy Physics**
Defense Nuclear Nonproliferation R&D	**Nuclear Energy**
Efficiency and Renewable Energy	**Nuclear Physics**
Electricity	

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

Advanced Scientific Computing Research (ASCR)

The primary objectives of the [ASCR Computational Facilities Allocation Policy](#) are: (i) support DOE Office of Science mission critical projects, and (ii) provide substantial allocations to the open science community through a peer review process for a small number of high-impact scientific research projects.

CREDIT:
https://science.osti.gov/ascr/Facilities/User-Facilities

Argonne Leadership Computing Facility (ALCF)

The ALCF provides the computational science community with a world-class computing capability dedicated to breakthrough science and engineering.

Energy Sciences Network (ESnet)

The ESnet is a high-speed network serving thousands of Department of Energy researchers and collaborators worldwide.
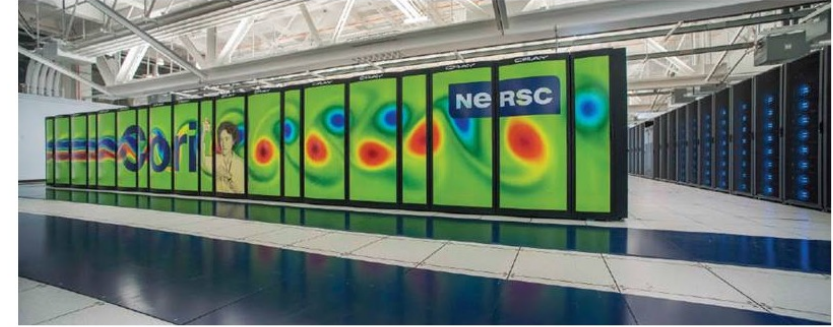
National Energy Research Scientific Computing Center (NERSC)

The NERSC is the mission high performance computing facility for the Department of Energy’s Office of Science, and is a world leader in accelerating scientific discovery through computation.

Oak Ridge Leadership Computing Facility (OLCF)

The OLCF provides the computational science community with world-class computing capability dedicated to breakthrough science and engineering.

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

ENERGY.GOV — Advanced Scientific Computing Research (ASCR)

DOE National Lab HPC Centers and Systems

o The National Renewable Energy Laboratory (NREL)

    o https://www.nrel.gov/hpc/

o Livermore Computing: HPC at LLNL, https://hpc.llnl.gov/

o Los Alamos NL High Performance Computing

    o https://www.lanl.gov/org/ddste/aldsc/hpc/index.php

o Idaho NL HPC, https://hpc.inl.gov/SitePages/Home.aspx

o Oak Ridge NL's Leadership Computing Facility, https://www.olcf.ornl.gov/

o Argonne NL's Leadership Computing Facility, https://www.alcf.anl.gov /

o Sandia NL HPC, https://hpc.sandia.gov/access/

CREDIT: https://science.osti.gov/ascr/Facilities/User-Facilities

PennState Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

ENERGY.GOV — Advanced Scientific Computing Research (ASCR)

Accessing ASCR High Throughput Network

Tools and resources deployed, developed, and maintained by the Energy Sciences Network (ESnet) are accessible through multiple mechanisms depending on the specific resource.

ESnet 100G testbed – ESnet 100 G testbed. Access is available to the research community and industry through a peer review process managed by ESnet.

ESnet Engineering Tools – ESnet develops tools and resources for network monitoring, performance and utilization in support of the DOE Office of Science's large-scale, collaborative science programs.

OSCARS – On-demand Secure Circuits and Advance Reservation System is an advanced software system for reserving time on high speed science networks.

perfSONAR – A tool developed collaboratively with multiple partners for end-to-end monitoring and troubleshooting of multi-domain network performance.

ESnet high speed network – Accessing ESnet's high-speed network occurs any time users send information to or from an ESnet endpoint, as when data is transported from the Large Hadron Collider at CERN to US collaborators or when citizens visit the webpage of a DOE national laboratory.

CREDIT: https://science.osti.gov/ascr/Facilities/User-Facilities

PennState Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

Established in 2005, the OSG Consortium operates a fabric of distributed High Throughput Computing (dHTC) services.

Open Science Pool (OSPool) provides researchers with *fair-share* access to computing and data capacity powered by *distributed high-throughput* computing (dHTC) technologies.

OSG

OSG All-Hands Meetings (AHM) provides the consortium stakeholders and the broader dHTC community with a venue to *share ideas and exchange information*.

Open Science Data Federation (OSDF) enables users and institutions to *share data files and storage capacity*, making them both accessible in dHTC environments such as the OSPool.

CREDIT: https://osg-htc.org/

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

Submit Locally, Run Globally
Researchers can run jobs on OSG from their home institution or an [OSG-Operated Access Point](#) (available for US-based research and scholarship).

The OSG Software Stack
The OSG provides an integrated software stack to enable high throughput computing; [visit our technical documents website for information](#).

Importantly, many compute tasks can take advantage of the OSPool with simple modifications, and we'd love to discuss options with you!

	Ideal Jobs!	Still very advantageous	Maybe not, but get in touch!
Expected Throughput, per user	1000s concurrent cores	100s concurrent cores	Let's discuss!
CPU	1 per job	< 8 per job	> 8 per job
Walltime	< 10 hrs*	< 20 hrs*	> 20 hrs
RAM	< few GB	< 40 GB	> 40 GB
Input	< 500 MB	< 10 GB	> 10 GB**
Output	< 1 GB	< 10 GB	> 10 GB**
Software	pre-compiled binaries, containers	Most other than →	Licensed Software, non-Linux

*or checkpointable

** per job; you can work with a large dataset on OSG if it can be split into pieces

CREDIT: https://osg-htc.org/

PennState Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

Open to All
OSG

- Open to providers at all scales
  - from small colleges to large national labs
- Open to user communities at all scales
  - from individual students to large research communities
    - domain science specific and across many campuses
    - campus specific and across many domain sciences
- Open to any business model
  - sharing, allocations, purchasing
  - preemption is an essential part of operations

CREDIT: https://osg-htc.org/

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

Nautilus is a HyperCluster for running containerized Big Data Applications. It utilizes Kubernetes for managing and scaling containerized applications in conjunction with Rook for automating Ceph data services.

The National Research Platform (NRP) is a partnership of more than 50 institutions, led by researchers and cyberinfrastructure professionals at UC San Diego. The NRP is supported in part by awards from the National Science Foundation.

The NRP is an all-in-one system that includes computing resources, research, and education. It is designed for distributed growth and expansion.

CREDIT: https://nationalresearchplatform.org/

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Nautilus

Nautilus is a powerful nationally distributed computer system with CPUs, GPUs, and FPGAs, in two types of subsystems ("high-performance FP64/FPGA" and "FP32-optimized"), specialized for a wide range of data science, simulations, and machine learning or artificial intelligence, allowing data access through a federated national-scale content delivery network.

CREDIT: https://nationalresearchplatform.org/

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023
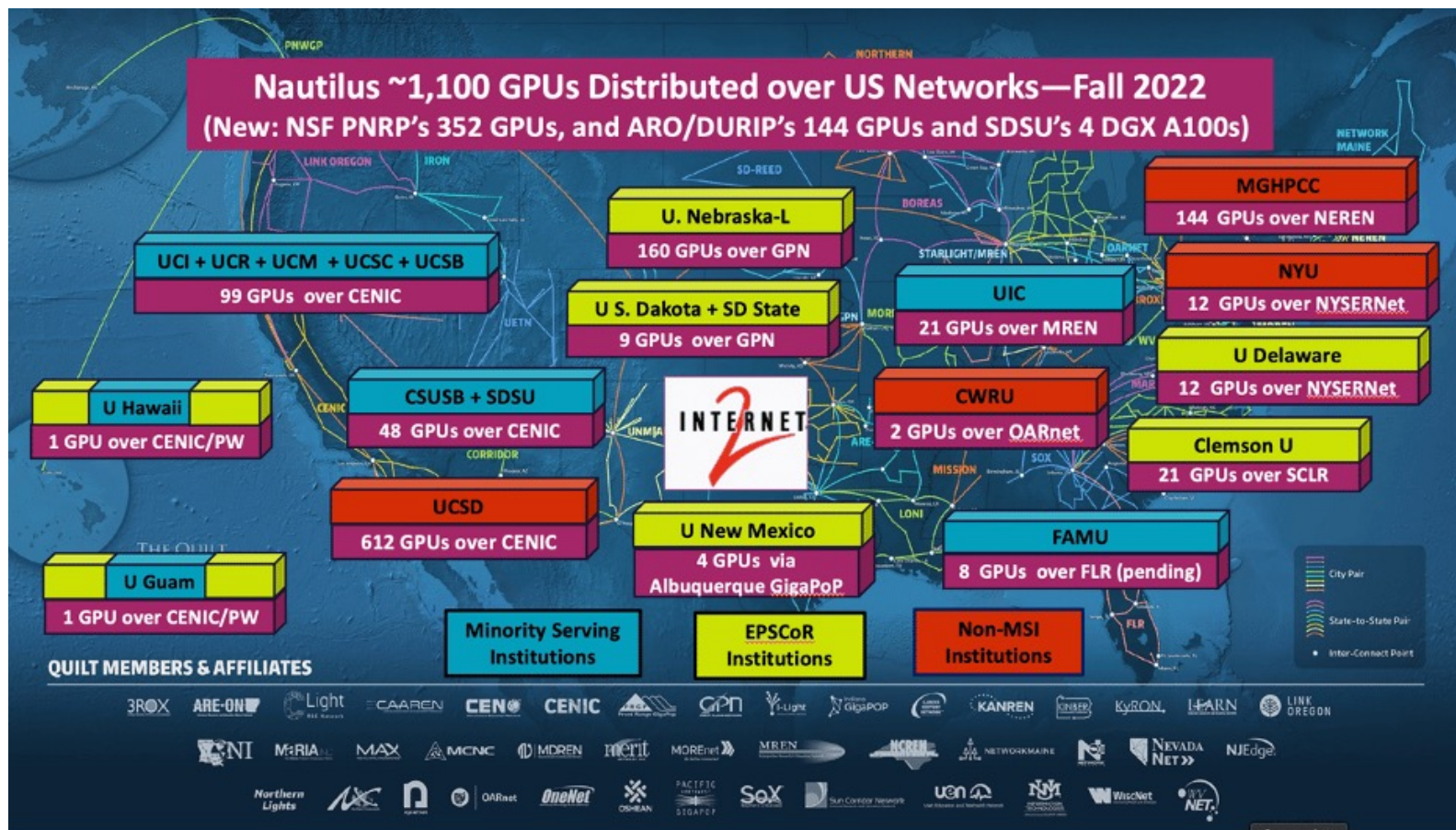
Systems, Service Providers, and Technologies

To get started on NRP/Nautilus, users go to https://nationalresearchplatform.org/ and start at join/contact.

There are JupyterLab notebooks available that address certain classes of need and reduce the pain of getting started a lot, if they do what is needed. They may need to get a cyberinfrastructure professional programmer to assist. Matrix is the online chat room we run; the developers, cyberinfrastructure professionals, and sysadmins prefer to monitor Matrix rather than e-mail.

Create account

Host account on

matrix.nrp-nautilus.io

≡ Intro ☀ 🔍 Search

Using the right region for your pod

Latency is significantly affecting the I/O performance. If you want optimal access speed to Ceph, add the region affinity to your pod for the correct `region` (`us-east`, `us-west`, etc):

```
spec:
  affinity:
    nodeAffinity:
      requiredDuringSchedulingIgnoredDuringExecution:
        nodeSelectorTerms:
        - matchExpressions:
          - key: topology.kubernetes.io/region
            operator: In
            values:
            - us-west
```

You can list the nodes region label using: `kubectl get nodes -L topology.kubernetes.io/region`

CREDIT: https://nationalresearchplatform.org/

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

CREDIT: https://www.psc.edu/resources/bridges-2/

Bridges-2
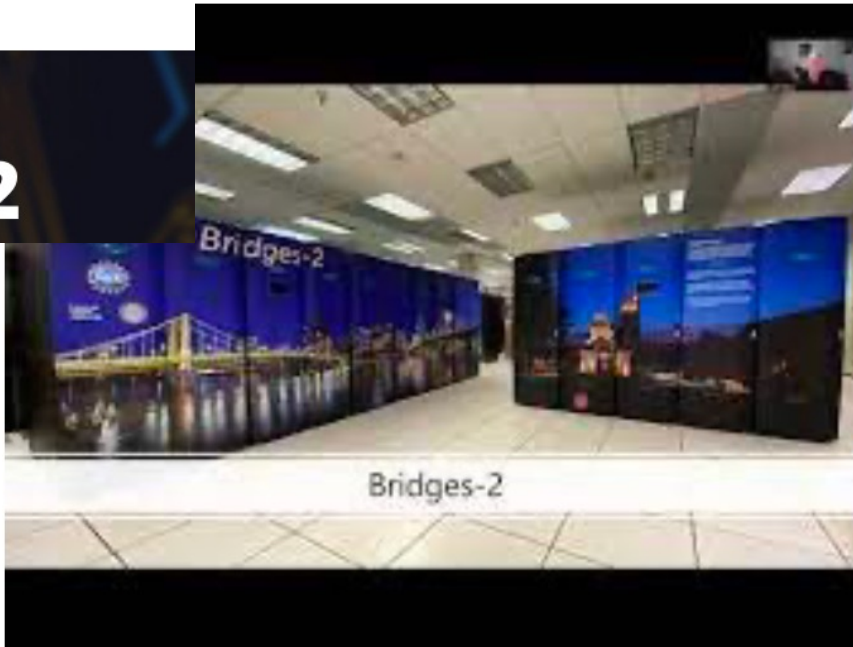
Provides transformative capability for rapidly evolving, computation-intensive and data-intensive research, creating opportunities for collaboration and convergence research.

Connecting new communities to advanced research computing.

More Science:	Approximately 3x larger than Bridges
Faster Computing:	Latest AMD EPYC processors
Faster Storage:	Fast flash array and tiered data management
Smarter Science:	Designed for Full System AI and data-centric computing
Scalable:	Interoperability with cloud and campus resources

"PSC's newest supercomputer"

Carnegie Mellon University   PSC   University of Pittsburgh

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

CREDIT: https://www.psc.edu/resources/bridges-2/

Bridges-2

- Interactivity
- Popular languages and frameworks: Python, Anaconda, R, MATLAB, Java, Spark, Hadoop
- AI frameworks: TensorFlow, Caffe2, PyTorch, etc.
- Containers and virtual machines (VMs)
- Databases
- Gateways and distributed (web) services
- Large collection of applications and libraries

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

With a focus on ease of use and broad accessibility, Jetstream2 is designed for those who have not previously used high performance computing and software resources.

The system is particularly geared toward 21st-century workforce development at small colleges and universities—in EPSCoR states.

⊘ ACCESS to Jetstream2 SUs

For simplicity, we've aligned our SU value such that *1 ACCESS Credit = 1 Jetstream2 SU.*

SUs are consumed at a rate of:

- Jetstream2 (CPU) - 1 SU per vCPU_core-hour (use of one virtual core of a CPU per hour).
- Jetstream2-LM (Large Memory) - 2 SUs per vCPU_core-hour
- Jetstream2-GPU - 4 SUs per vCPU_core-hour

Please refer to VM Sizes and configurations to see available VM flavors and per hour cost on Jetstream2.

- **SUSPENDED** instances will be charged .75 of their normal SU value. (75%)
- **STOPPED** instances will be charge 0.50 of their normal SU value. (50%)
- **SHELVED** instances will not be charged SUs. (0%)

For Large Memory and GPU allocations, the vCPU core hour cost is 2x and 4x respectively as noted above.

CREDIT: https://jetstream-cloud.org/about/index.html

PennState Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Core capabilities
- Interactive virtual machines (VMs)
- Globus Transfer
- Virtual Desktops
- Digital Object Identifiers (DOIs)

CREDIT: https://jetstream-cloud.org/about/index.html

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

AIMI

Systems, Service Providers, and Technologies

SGX3 | SGCI
A Center of Excellence and a Software Institute Serving the Science Gateways Community
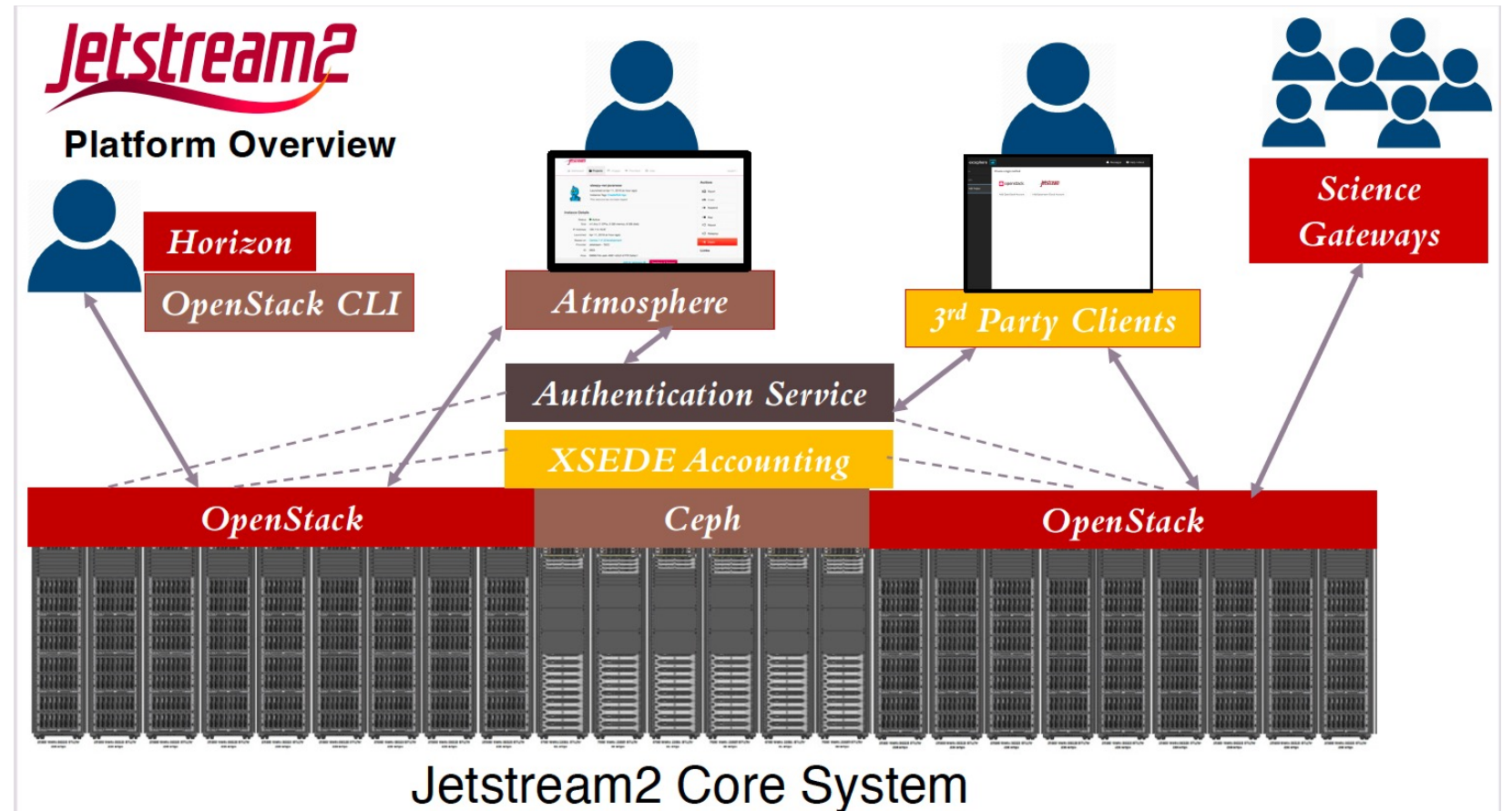
SGX3
SGX3 Funded Services

Are you building websites that serve your discipline?

Do you wish you could connect with and learn from others, doing the same thing?

Cyberinfrastructure Center of Excellence called, "SGX3 - A Center of Excellence to Extend Access, Expand the Community, and Exemplify Good Practices for CI Through Science Gateways," under NSF award #2231406.

SGCI
Client Funded Services

This community driven information was used to form the Science Gateways Community Institute (SGCI) in 2016 under NSF award #1547611. To date, the SGCI has served over 200 clients with a variety of services, had over 3,400 attendees to its community building events, supported 1,200 students and faculty, and produced 195 research products.

CREDIT: https://sciencegateways.org/about

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

SGX3 | SGCI
A Center of Excellence and a Software Institute
Serving the Science Gateways Community
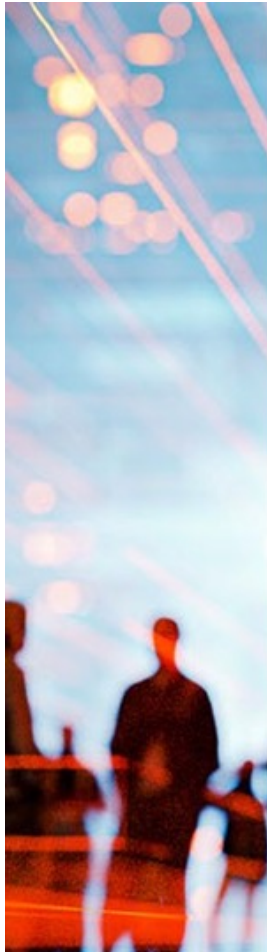
GATEWAY SERVICES

Our consulting services are tailored to meet the needs of gateway projects and have demonstrated an acceleration of gateway efforts, saving significantly on funds and time-to-science. We turn ideas into reality or enhance existing gateways. Our team consists of experts in all areas of gateway development, operations, and long-term sustainability.

Comprehensive Consulting Services
- Embedded technical support
- User experience design
- Usability
- Graphic design
- Cybersecurity
- Sustainability
- Marketing

EDUCATION & TRAINING

We offer a range of learning opportunities for the gateway community, from training events to skilled internships to a searchable database of resources and educational materials. The participants in our student-focused, workforce development programs have contributed actively to gateway projects and have gone on to start new careers.

Sustainability Training
- Gateway Focus Week
- Jumpstart Your Sustainability Plan

Student Programs
- Internships
- Hackathons
- Young Professionals Network

NETWORKING & COMMUNITY

With an annual conference, a variety of workshops and webinars, a community discussion forum, and a collaborators program, we have built a vibrant and engaging gateway community—and before the SGCI was founded, many members did not realize such a community existed.

Opportunities for Engagement
- Gateways annual conferences
- Workshops and webinars
- Community discussion forum
- Collaborators Program

CREDIT: https://sciencegateways.org/app/site/media/files/SGCI_storybook_interactive_v2022-12-13.pdf

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies
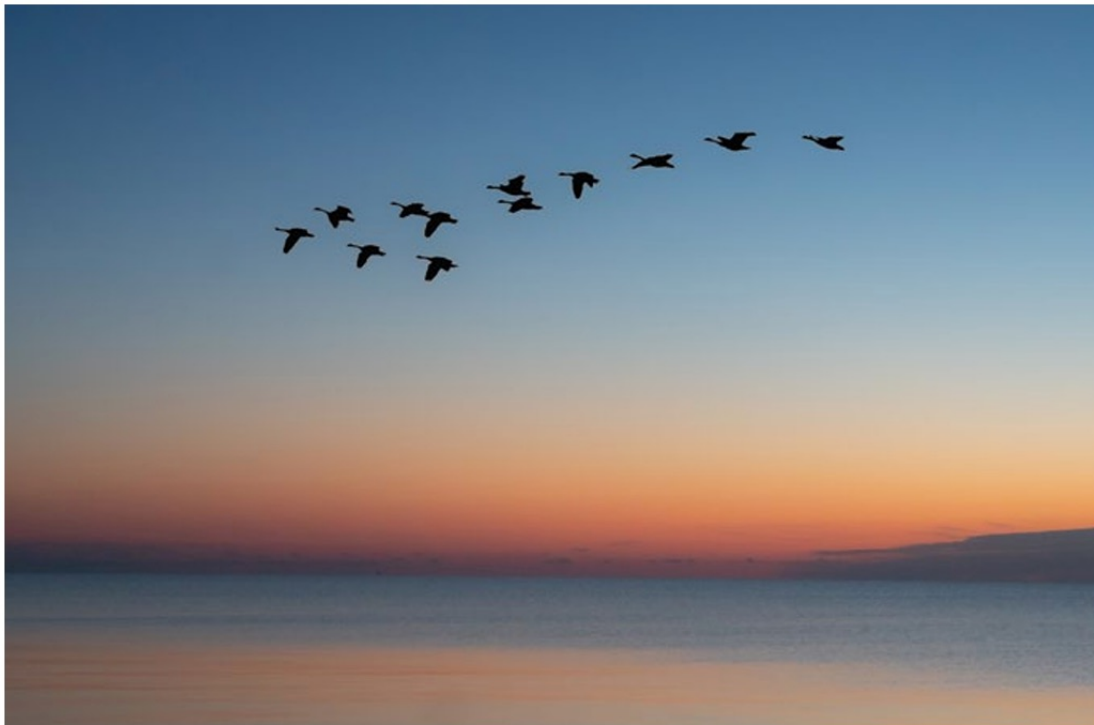
ETAG

ARE BIRD MIGRATION PATTERNS CHANGING due to climate change? What is the impact of habitat loss on sloths? How far do wolf packs travel? These are the types of questions that scientists can answer by using wirelessly readable tags and Radio Frequency Identification (RFID) technology. The **Electronic Transponder Analysis Gateway (ETAG)** is a web tool that allows scientists to collect and share data for animal behavior in near real-time. The gateway provides a central store for researchers to upload data that can allow for collaboration and visualizations when tracking individual animals.

Screenshot

CREDIT: https://sciencegateways.org/app/site/media/files/SGCI_storybook_interactive_v2022-12-13.pdf

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

CYVERSE®

Tiers & Features

	Basic	Regular	Pro	Commercial
Discovery Environment	Yes	Yes	Yes	Yes
Data Store	Yes	Yes	Yes	Yes
Webinar Access	Yes	Yes	Yes	Yes
Advanced Features & APIs	-	-	Yes	Yes
Access to GPU	-	-	**	**
Data Storage Limit	5 GB	50 GB	3 TB	5 TB
Compute Units / Year*	200	1,000	20,000	200,000
Concurrent Jobs	1	2	4	8
Sharing Data & Apps	None	100	Unlimited	Unlimited
DOI for Data	None	5	10	20
Workshop Seats	0	2	4	10
Support	Email	In App Chat†	Screen Share Support	Screen Share Support
Price / Year	Free	$200	$340	$2,000

Vision: Transforming Science through Data-Driven Discovery

Mission: To design, deploy, and expand national Cyberinfrastructure for research, and to train scientists in its use.

About: CyVerse provides scientists with powerful platforms to handle huge datasets and complex analyses, thus enabling data-driven discovery. Our extensible platforms provide data storage, bioinformatics tools, data visualization, interactive analyses, cloud services, APIs, and more.

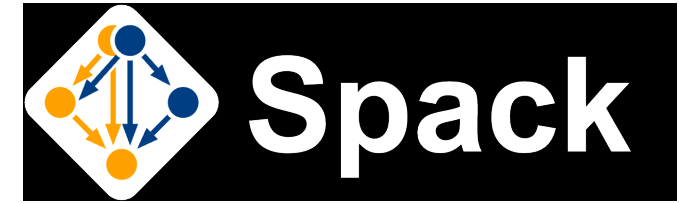CREDIT: https://www.cyverse.org/about

PennState
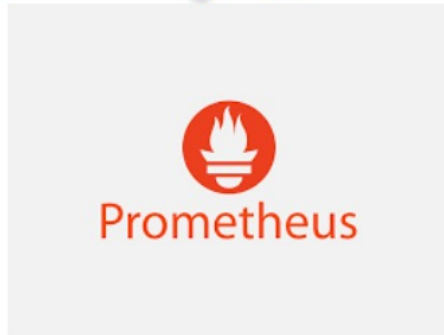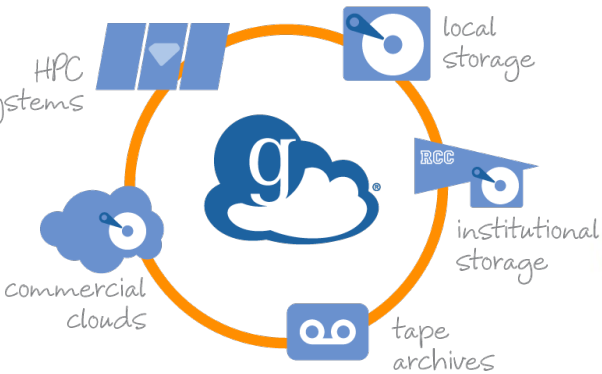Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

- **Internet2 is a community organization**: Higher education and Research institutions - Government entities, corporations and cultural organizations

- **Cyberinfrastructure provider**: secure high-speed network, cloud solutions, research support, services tailored for research and education

- **Through InCommon**: Security and privacy, IAM tools for research and education, single sign-on (SSO) for access to cloud and local research computing services and roaming wi-fi w/ eduroam

CREDIT: https://internet2.edu/

Cloud Solutions	**Security**
When you look to the cloud, we've got you covered. Our cloud services and solutions are designed for research and education.	Internet2 security services and solutions contribute to your secure infrastructure, including roaming WiFi access, single sign-on and analytical services.
Research Engagement	**Advanced Networking**
Collaborate and access key technology and research support resources. Connect to partners across the community.	Learn about the nation's largest and fastest coast-to-coast research and education network.
Community Collaboration	**Support**
Now more than ever, we need to collaborate and share. Our community works together to ask questions and turn the answers into a robust network, and software and services that serve your interests.	Do you have a question or suggestion? Please reach out so we can continue to serve you and the community.

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

The Minority Serving – Cyberinfrastructure Consortium (MS-CC) and Internet2 have been awarded a nearly $15 million grant from the National Science Foundation (NSF) to support the acceleration of cyberinfrastructure-centric research capacity at HBCU and TCU campuses.
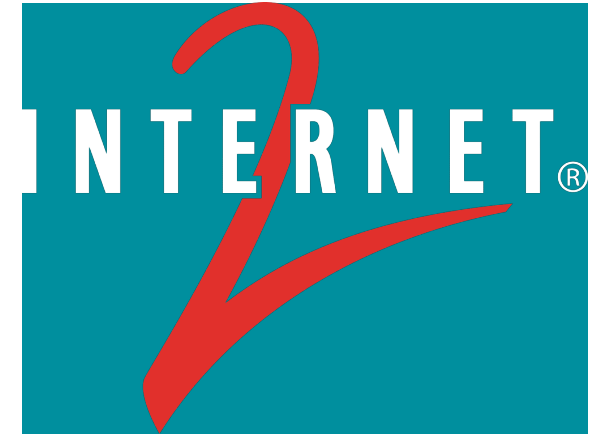
Funding supports accelerating cyberinfrastructure-centric research capacity at historically Black colleges and universities (HBCUs) and tribal colleges and universities (TCUs) through proof-of-concept grants and shared resources

CREDIT: https://internet2.edu/

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies
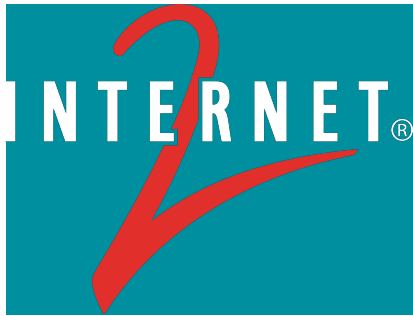
BY THE NUMBERS

ADVANCING SCHOLARSHIP • ACCELERATING DISCOVERY

320+ HIGHER EDUCATION MEMBERS

950+ EDUROAM SUBSCRIBERS

350+ NET+ SUBSCRIBERS

1000+ INCOMMON PARTICIPANTS

80,000+ COMMUNITY ANCHOR INSTITUTIONS

50+ INDUSTRY MEMBERS

60 AFFILIATE & GOVERNMENT MEMBERS

100+ COUNTRIES & RESEARCH NETWORKS CONNECTIONS

750+ NET+ CLOUD CONTRACTS

46 REGIONAL & STATE NETWORKS

800G+ WAVELENGTHS OF NETWORK CAPABILITY

32 Tbps CAPACITY PER LINK

CREDIT: https://internet2.edu/

PennState Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

Online tools like ask.CI foster more than code snippet copy/paste, it involves discussions of best approaches.

CREDIT: https://ask.cyberinfrastructure.org/

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

Search the [ACCESS Knowledge Universe.](#) Search is an evolving feature and will improve with time. If you have suggestions to help train this search please use [this form](#).

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

CI links have been crowd-sourced from the ConnectCI community and represent a "vetted" list of useful websites, training modules and tutorials.

CI links show up in a tag search if they have the relevant tag attached. Affinity groups can include relevant CI links on their respective affinity group pages. Additional CI links are always welcome, click the "Add New CI Link" button to suggest one.

CREDIT: https://campuschampions.cyberinfrastructure.org/ci-links

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

CAMPUS CHAMPIONS

Collaboratively employing advanced computing and data to tackle society's grand challenges

Champions Nationionwide	Institutions Represented	Institutions in EPSCoR States	Minority Serving Institutions
784	359	86	61

The mission of the Campus Champions is to foster a dynamic environment for a diverse community of research computing and data professionals sharing knowledge and experience in digital research infrastructure.

CREDIT: https://campuschampions.cyberinfrastructure.org/

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

CREDIT: https://campuschampions.cyberinfrastructure.org/

PennState Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

Systems, Service Providers, and Technologies

People who use or support people who use ACCESS resources and the ACCESS Resource Allocation System.

ACCESS Affinity Groups encourage community members to gather based on common interests, and they provide connections to AG-specific slack channels, Q&A Forums, events, and relevant knowledge base resources. Joining ACCESS Resource Provider Affinity Groups (AGs) will add you to the distribution list for notifications about outages, news, and training events specific to the Resource Provider named.

CREDIT: https://support.access-ci.org/affinity_groups

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

CONVERSATIONS

Thanks for your attention!

Instead of asking for questions per se, I am asking the group here – in the chat – put your favorite "go-to" resource – anything that was touched on here or not mentioned at all and you just cannot see living without it!

Share a link or a brief call-out.

ACKNOWLEDGEMENTS:
- 2022 Virtual Residency speaker - "The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies", Jacob Fosso Tande, The University of North Carolina at Greensboro
- 2021 Virtual Residency speaker - "The CI Milieu: Systems, Service Providers, Technologies", Douglas Jennewein, Arizona State U

PennState Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

PennState
Institute for Computational
and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023

What Is Cyberinfrastructure? What is Cyberinfrastructure (CI)? That depends on who you ask: Network engineers: CI is the network. Supercomputing professionals: CI is supercomputers, plus the network. Cybersecurity professionals: Cyberinfrastructure is everything that needs cybersecurity. Everyone else: CI is my thing, plus supercomputers, plus the network.

What is a CI Facilitator? "Advanced Cyberinfrastructure Research & Education Facilitator" (ACI-REF –term coined by Miron Livny) Work with users –researchers and educators –to help them improve their research and/or education productivity and aspirations via advanced Cyberinfrastructure (CI). Typically, one or a few CI Facilitators have responsibility for an entire institution, or even multiple institutions. At some institutions, CI Facilitation is part time; at others, it's full time. Some CI Facilitators are or used to be: faculty (current or former); postdocs (current or former); research staff (current or former); IT professionals, including from Enterprise IT (current or former); graduate or undergraduate students (current or former).

The Five Facings
- Researcher-facing (e.g., CI Facilitator)
- System-facing (e.g., cluster sysadmin, network engineer)
- Data-facing (e.g., Research Data Librarian)
- Software-facing (e.g., Research Software Engineer)
Strategy/policy-facing (e.g., institutional CI leader)

PennState
Institute for Computational and Data Sciences

The Cyberinfrastructure Landscape: Systems, Service Providers, Technologies
Virtual Residency Introductory Workshop 2023 – Tues June 27, 2023