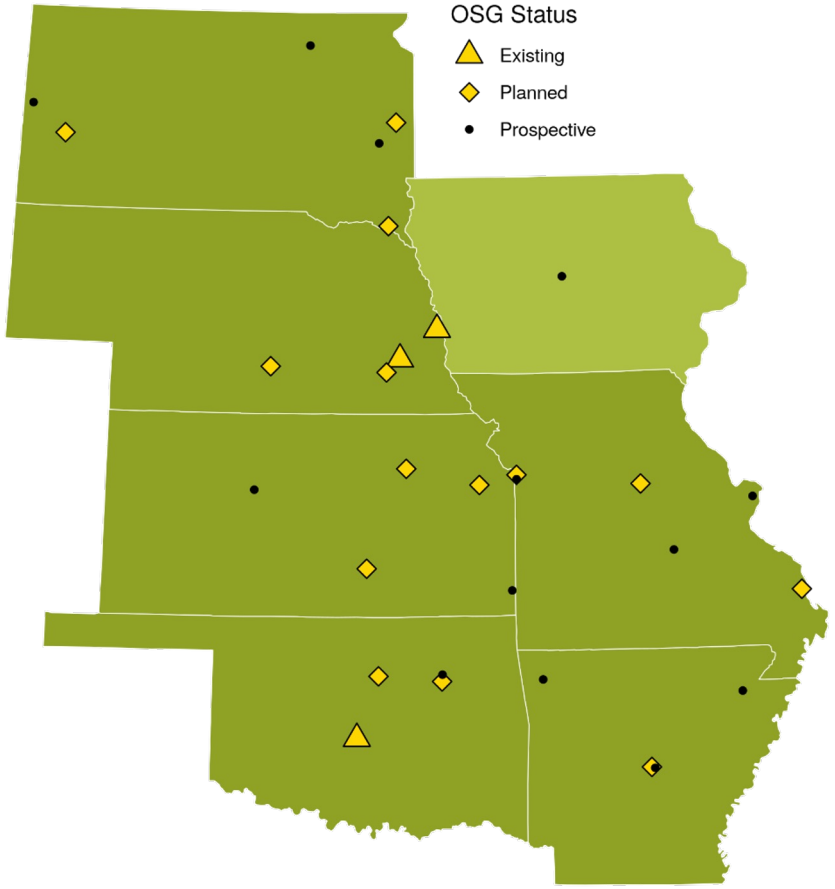
Integrating Great Plains Network Capacity into the OSPool

Derek Weitzel
University of Nebraska–Lincoln

Introduction to GP-ARGO

- You get a node, you get a node… everyone gets nodes!
- Evolved over the lifetime of the project.
- How to manage all these nodes?
- How to get “buy-in” to GP-ARGO: Visualization and Monitoring
- Nodes with NVIDIA A100 GPU



Introduction to GF

GP-ARGO (Under the name of Great Plains Network) is a major contributor to opportunistic usage

(Credit: Frank's slide)

OSG

Top OSPool

Facility	Core Hours
Syracuse University	36.0 Mil
University of California San Diego	29.7 Mil
University of Wisconsin	26.4 Mil
Lancium	14.7 Mil
Great Plains Network	9.95 Mil
University of Chicago	6.51 Mil
Indiana University	6.37 Mil

GP-ARGO Distribution

- Initially configured nodes at KSU with input from local admins
- Shipped nodes to sites, sometimes by dropping off and storing at my house
- Configured with disk imaging and local scripts, with a wireguard VPN set to connect back to KSU.

Ryan Johnson (University of South Dakota) picking up their node from my house

Ansible Playbooks

- Deployed nodes initially by hand, then with Ansible.
- Deploy the OSG pilot container with ansible
  - Working on contributing the recipe back up to OSG for others to use.
- Ansible deployed:
  - **OSG Pilot Container**
  - **CVMFS**
  - NVIDIA Container Driver
  - Chrony (Time synchronization)
  - Grafana (visualization, more on that later)

OSG Deployed on GP-ARGO

- OSG runs on the GP-ARGO through the OSG Pilot Container
- It is a docker container that once started, will accept OSPool jobs.
- Self-registration to get a token allowing the container to join the OSPool
- Supports GPU jobs if the node has a GPU

Monitoring and Visualization

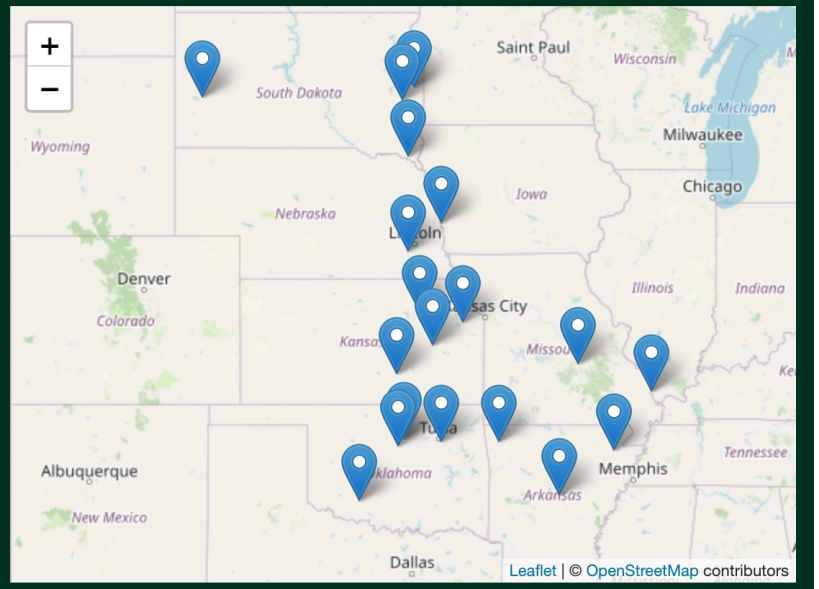
- It was important to provide information to the stakeholders so they can ask?
- What science is using GP-ARGO?
- How utilized is my institution's node?
- Who is using my institution's node?

Monitoring and Visualization

GP-ARGO

The Great Plains Augmented Regional Gateway to the Open Science Grid

GP-ARGO creates a regional distributed Open Science Grid (OSG) Gateway led by the Great Plains Network (GPN) to support computational and data-intensive research across the region through the development of specialized CI resources, workforce training, and cross-support methodologies and agreements.



Resources contributed to the Open Science Grid

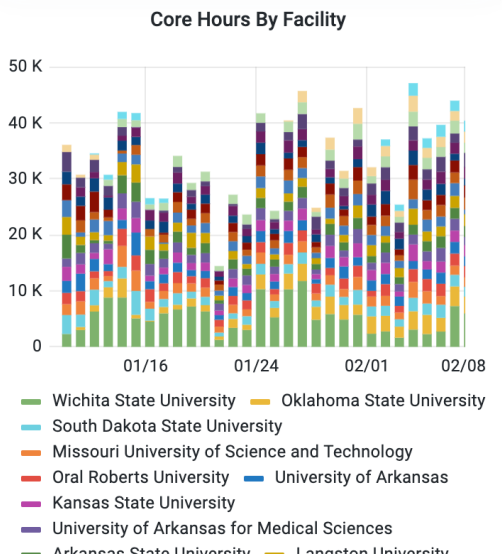
Core Hours Contributed	OSG Projects	Institutions Supported
16,034,123 ↑~34k/day	85	59

https://gp-argo.greatplains.net/

Monitoring and Visualization

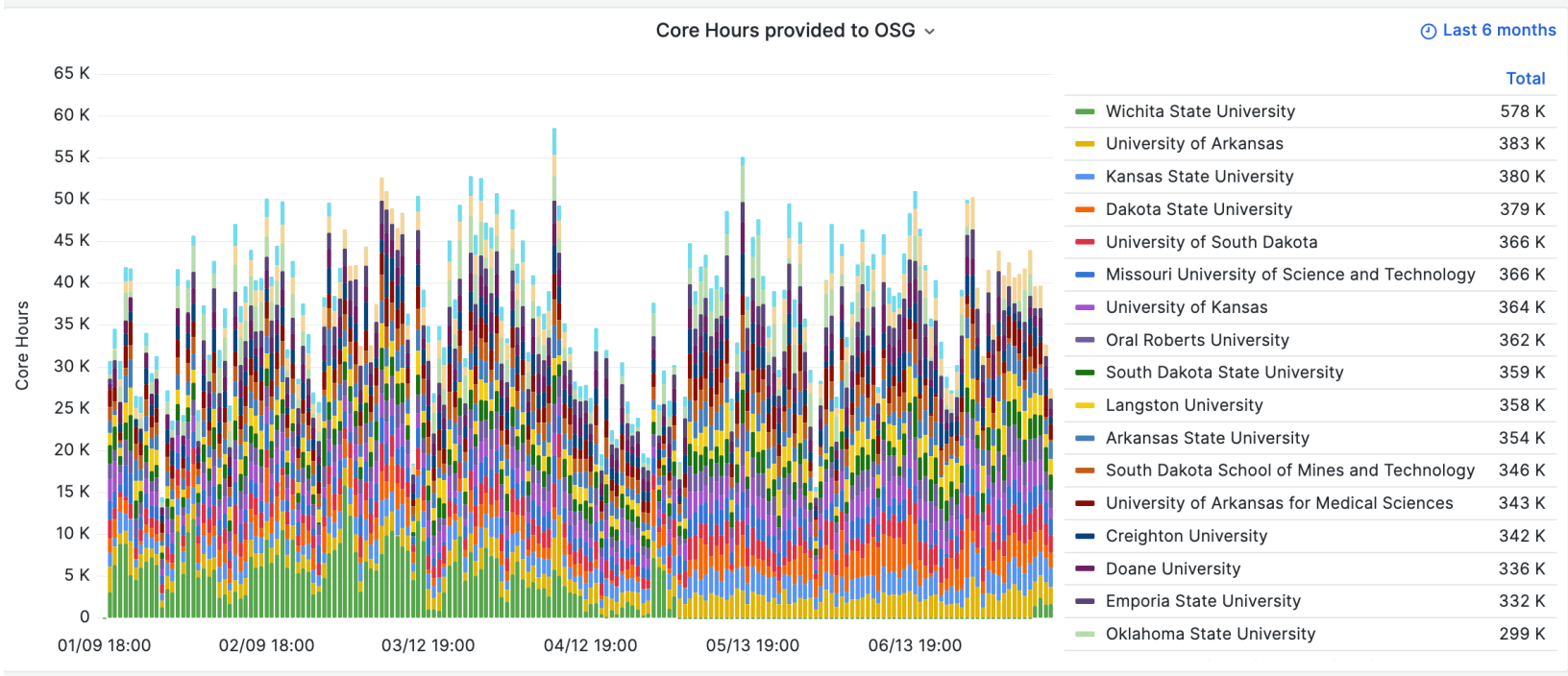
Collaborating Campuses
Arkansas State University
Cameron University
Creighton University
Dakota State University
Doane University
Emporia State University
Kansas State University
Langston University
Missouri S&T
Oklahoma State University
Oral Roberts University
South Dakota School of Mines and Technology
South Dakota State University
Southeast Missouri State University
University of Arkansas
University of Arkansas for Medical Sciences
University of Kansas
University of South Dakota
Wichita State University

OSG Project	Core Hours
Nuclear Physics Wayne State University *ID: WSU_3DHydro*	2,319,256
Gravitational Physics University of Michigan *ID: Michigan_Riles*	2,177,066
Astrophysics University of Chicago *ID: spt.all*	1,679,570
Physics Rutgers, The State University of New Jersey *ID: PixleyLab*	1,021,939
Chemical Sciences Carnegie-Mellon University *ID: TG-CHE200122*	998,385
Astronomy American Museum of Natural History *ID: AMNH.astro*	732,805
Astronomy and Astrophysics Rochester Institute of Technology *ID: CompBinFormMod*	618,884
Astrophysics	480,233



https://gp-argo.greatplains.net/

Monitoring and Visualization



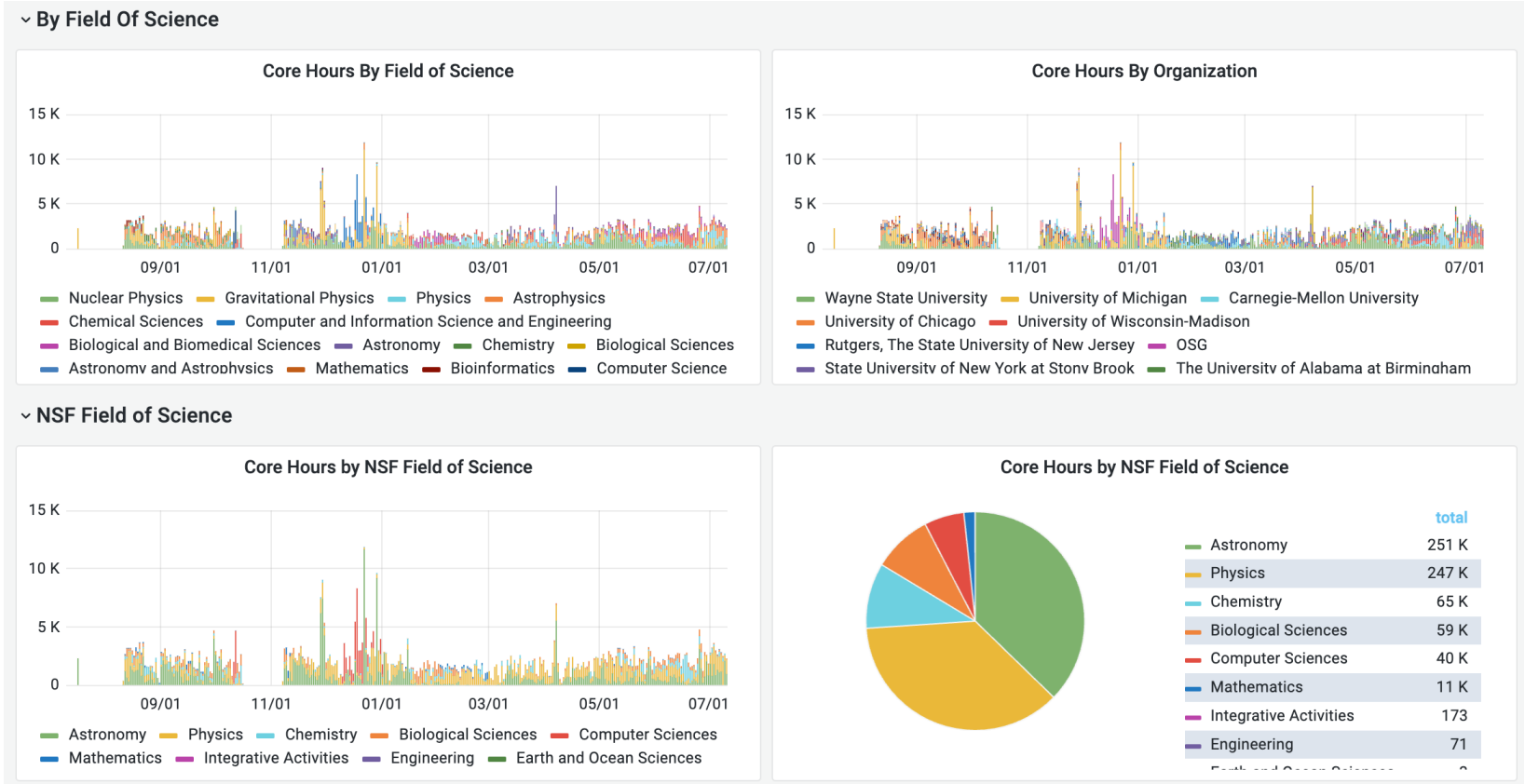
Monitoring and Visualization

OSG Science Enabled — Last 6 months

PI Name	Institution	Field of Science	OSG Project Name	Core Hours
Chun Shen	Wayne State University	Nuclear Physics	WSU_3DHydro	777 K
Jedediah Pixley	Rutgers, The State Universit...	Physics	PixleyLab	608 K
Olexandr Isayev	Carnegie-Mellon University	Chemical Sciences	TG-CHE200122	565 K
Francis Halzen	University of Wisconsin-Ma...	Astrophysics	IceCube	474 K
John Carlstrom	University of Chicago	Astrophysics	spt.all	455 K
Keith Riles	University of Michigan	Gravitational Physics	Michigan_Riles	400 K
Jiangyong Jia	State University of New Yor...	Physics	SBU_Jia	383 K
Summer Thyme	The University of Alabama a...	Biological and Biomedical S...	UAB_Thyme	250 K
Thu Le	University of Connecticut	Physics	UConn_Le	227 K

Monitoring and Visualization

- What ran at my institution? Example is University of Arkansas



Difficulties

- Working with many different universities
  - Differing policies– Especially with networking
- For example, one university (who shall be unnamed) blocked all outgoing connections.
  - The OSG requires outgoing connections for data transfers, monitoring, management of resources
  - The site offered to whitelist IPs, which is very difficult for OSG with arbitrary users

Difficulties

- GPUs from host all the way to container
- OSG Pilot Container already had GPU support ✅
- NVIDIA software to export container changed during the project ❌
- Originally, if running on a GPU-enabled node, only GPU-enabled jobs would run, frequently leaving the node idle.
  - Now the pilot container can optionally run non-GPU jobs, if there are no GPU jobs available.

The Future

- Great Plains Extended Network of GPUs for Interactive Experimenters (**GP-ENGINE**)
- CC* Regional Computing award #2322218
- Just starting **July 1, 2023**

- **Integrate resources with the National Research Platform**
- Contribute to the OSG through the integration of the NRP