

Facilitation of OSG Services

OSG School 2023

August 11, 2023



What are we going to talk about?

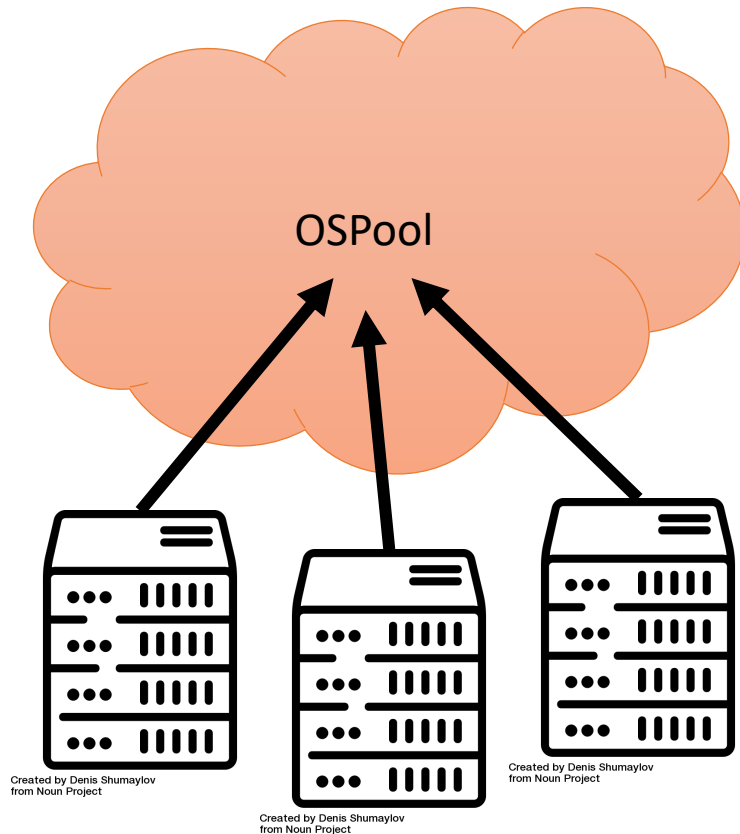
- How can you (or your institution) partner with us (OSG team)?
- How can you partner with the researchers at your institution?



Partnering With OSG Services



I want to contribute to the OSPool



- Two models:
 - A cluster; we need a user account and ssh access via that account to head node
 - OR, any group of servers with ability to start containers
- Your computers become available to run jobs from the OSPool.
- [More info in this talk](#)

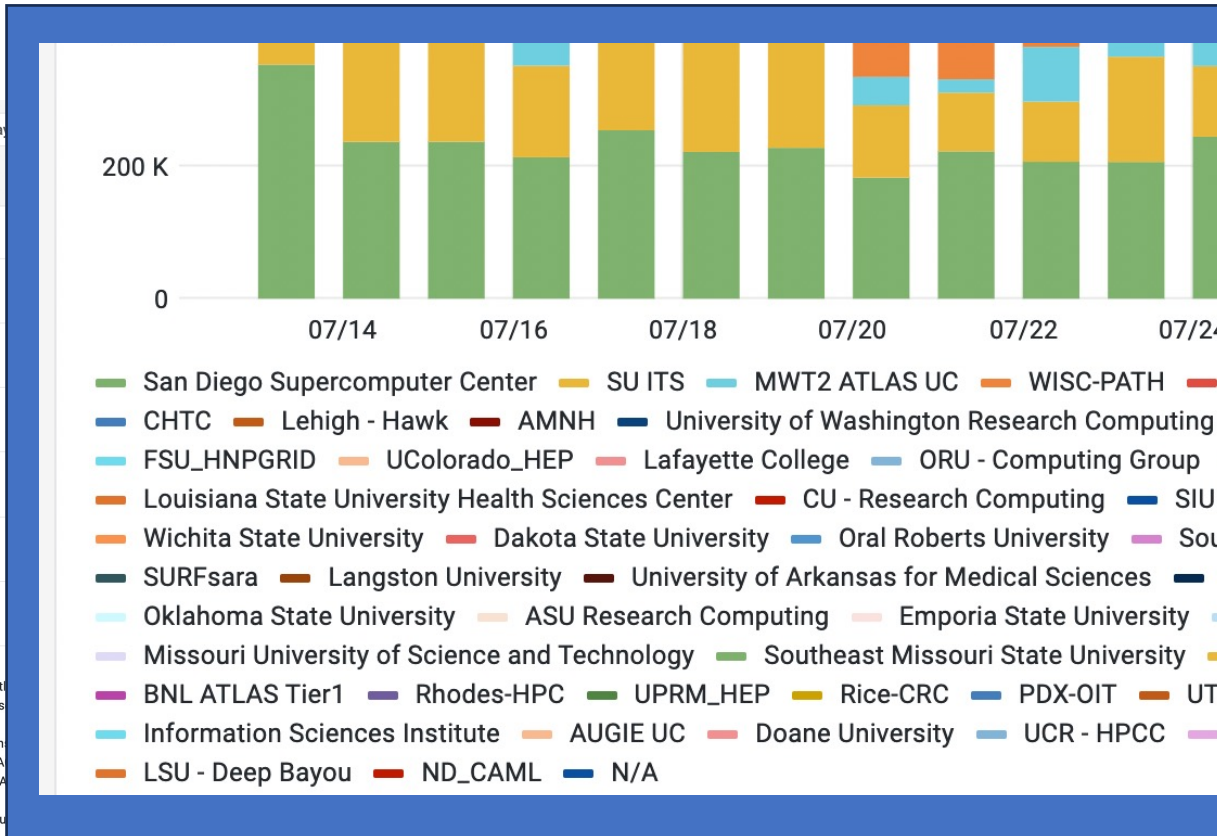
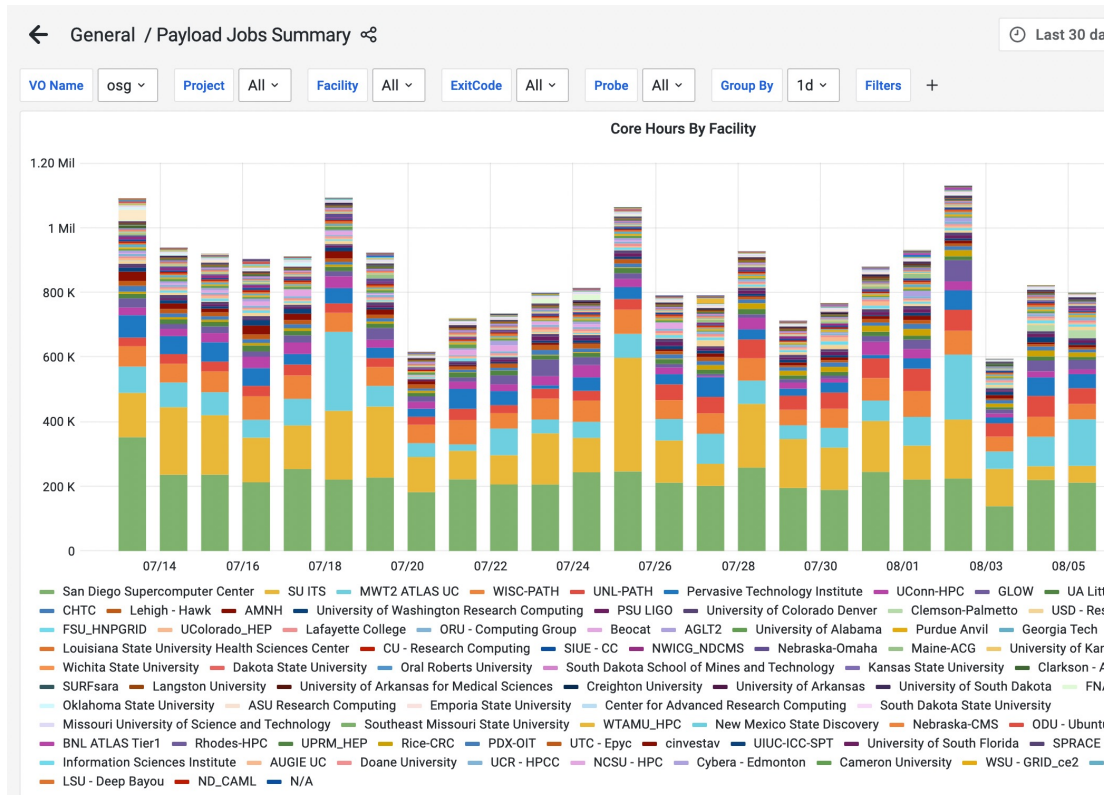


Why Contribute?

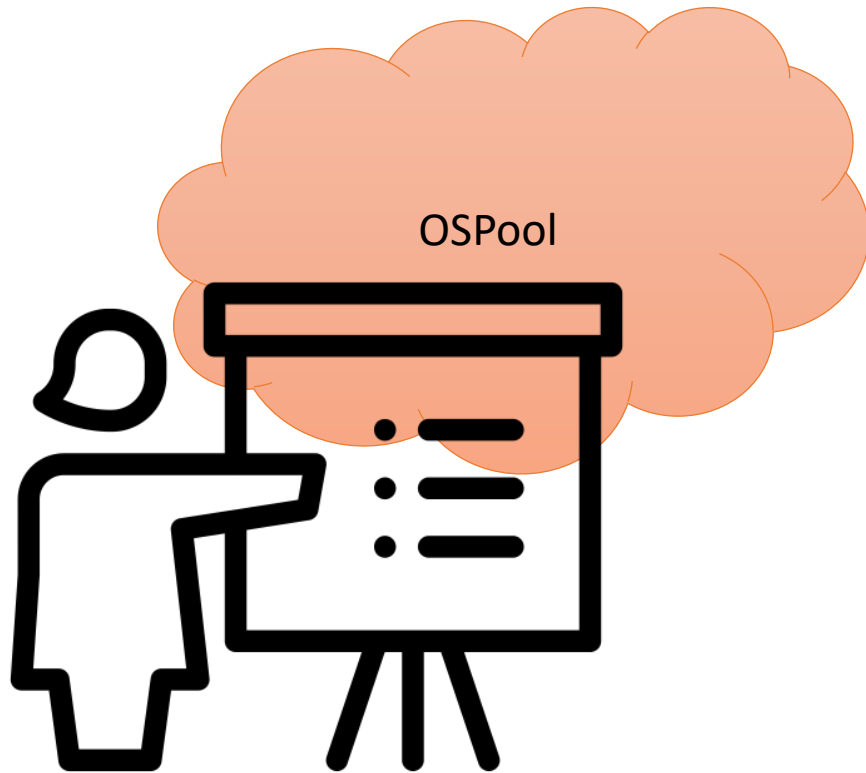
- Part of grant requirement like Campus Cyberinfrastructure (CC*)
 - We write letters of support for CC*
 - <https://osg-htc.org/campus-cyberinfrastructure.html>
- Desire to participate in the OSG Community
- Keeping otherwise idle hardware busy (hopefully)



Contributing Facilities



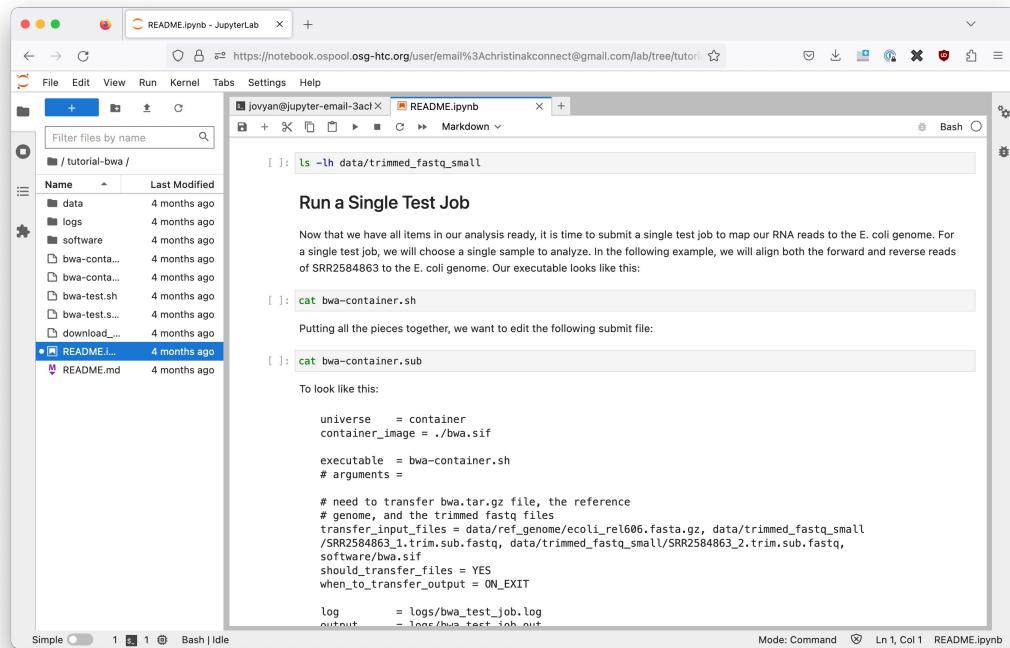
I want to use the OSPool for teaching



- Introduce batch computing or job submission in a class
- Run a workshop on high throughput methods
- Provide experience with computing in a domain science class



OSPool Notebooks



<https://notebook.ospool.osg-htc.org/>

- Starts Jupyter on an HTCondor Access Point
- Guest access (no account needed) or full accounts
- No software installation required
- GUI interface
- Can create interactive bash or python notebooks

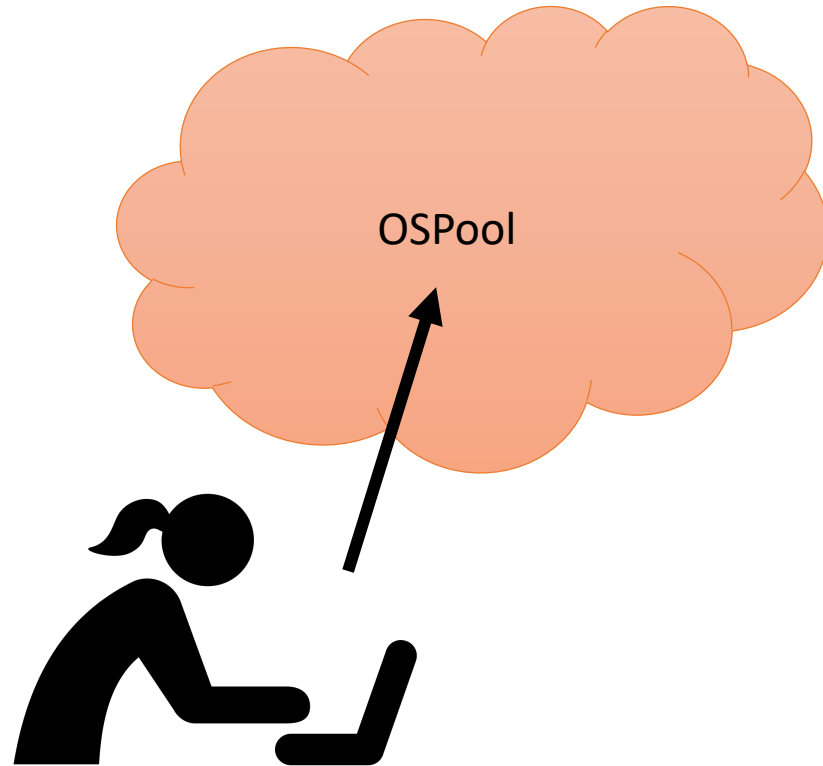


Teaching Resources

- [OSP Pool Resources for Teaching and Education](#)
- [Bring the OSP Pool to Your Event](#) (talk)
- Use any of our existing training materials:
 - [Slides from monthly training](#)
 - [Tutorials](#)
 - [OSG School materials](#)



I want to help my users use the OSPool



- How can your local users access the OSPool?
- Option 1: use an OSG-Supported Access Point (like ap40)
 - We are planning some changes to our onboarding flow to make this as seamless/easy as possible.
- Option 2: run an Access Point at your institution – talk to us
 - Allows users to use local identity for login

I want to help my users use the OSPool



Created by Adrien Coquet
from Noun Project

- [Services we provide](#)

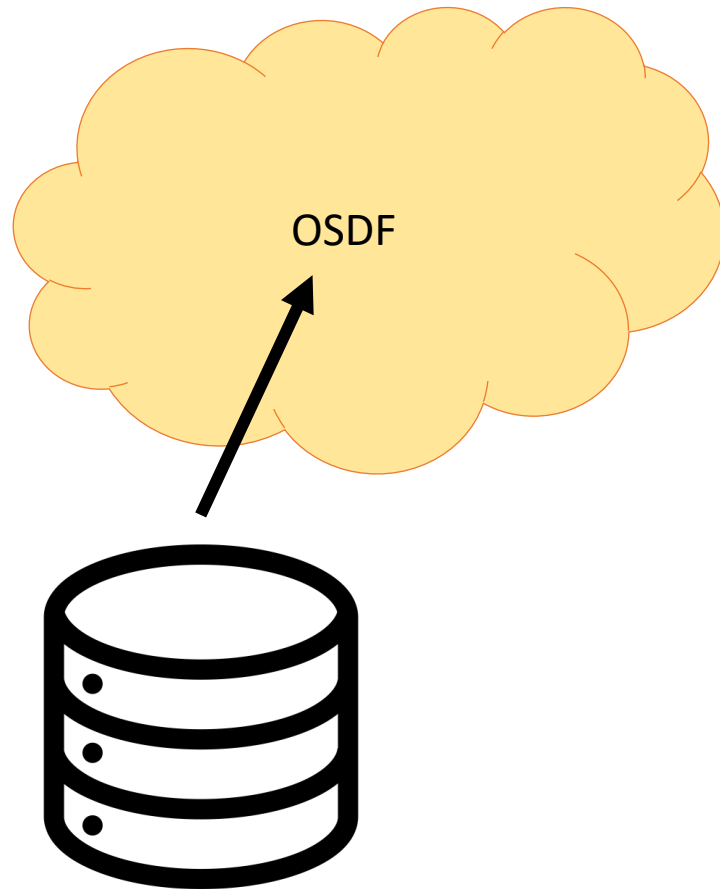
AND/OR

- Provide a set of curated containers or container build files
- Create github repositories with sample jobs
- Do training (see prev. slides)

(See also end of this presentation)

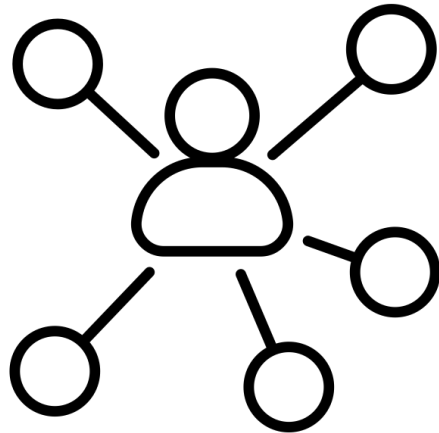


I want to share data



- Set up data origins to make data available via the OSDF
 - For example, making a local dataset available for more people to use.
- [More info in this talk](#)
- Still evolving - if you are interested, talk to us!

I want to stay connected



Created by Tini Sumiarsih
from Noun Project

We want to stay connected with YOU!

- Come to office hours
 - [User office hours are Tues/Thurs](#)
 - [Campus specific office hour on Mon](#)
- See us at conferences (e.g. PEARC)
- Invite us to a local conference or regional meeting!
- [Subscribe to our quarterly newsletter](#)
- Visit and shadow the UW Madison team



Partnering with your users



We see facilitation as:

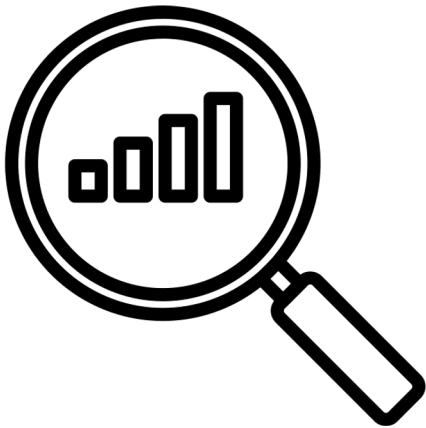
**“I want to help you achieve your research
(or teaching) goals.”**

instead of

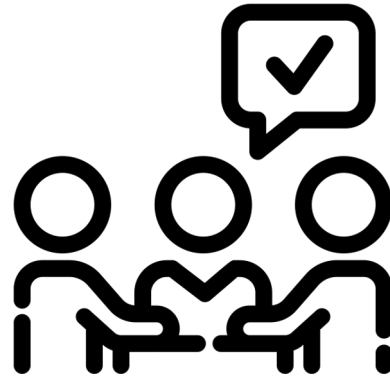
“I want to help you do computing.”



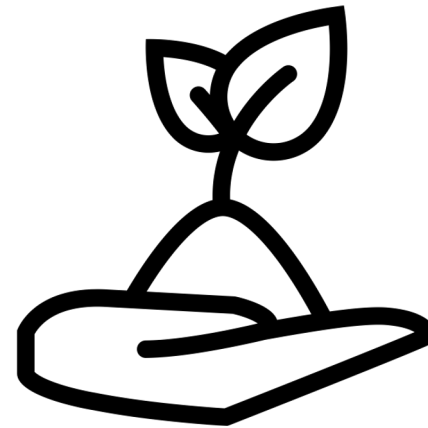
Facilitating HTC Workloads



Identify



Communicate



Support

Magnifying Glass by Adam Novantio from the Noun Project
Conversation by Maxim Kulikov from the Noun Project
sprout by lastspark from the Noun Project



Find HTC Workloads

- Where might there be HTC users at your institution?
 - User submitting 100s of jobs on local cluster.
 - “Non-traditional” computing domains, especially those doing statistical and “data science” work.
 - People with lots of pieces of data (sensors? images? samples?)
- Places to look
 - New faculty orientation, new student orientation
 - Lab/department meetings



Discuss HTC Workloads

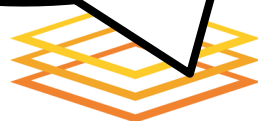
Tell me about
your research
...and how does
computing fit in?

What is the
smallest "piece" of
your workflow?

How are you running it now?

- compute/data requirements?
- your computing background?

How much
would you like
to run?



Discuss HTC Workloads

Want to understand the long-term goals of a user

Determine if problem can be broken into small pieces = HTC

Identify learning curve and potential benefit to user

Understand potential to scale



Challenges and Opportunities

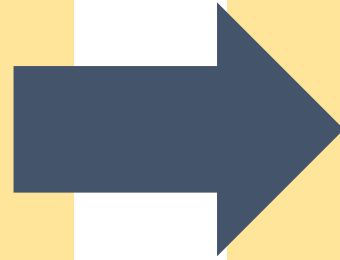
- **HTC is new**
 - Requires thinking of work in a different way (splitting jobs; optimizing for throughput rather than single-job performance).
- **(OSG specific): OSPool has unique job logistics**
 - Data is transferred (no shared filesystem).
 - Software portability.
 - Jobs can be interrupted.
- **An HTC system can provide previously impossible scalability**
 - On a small shared cluster, using 100s or 1000s of cores at once might be rare, or require long wait times. On the OSG, it's an everyday occurrence.



Communicating Outcomes

If you ...

- > execute your work ***this way*** (HTC)
- > requiring ***these learning steps*** and ***this much work***.



Then you ...

- > can reach your research goal after ***this much time and effort***, and
- > could achieve ***this much more research outcome***.



Building Relationships

- **Responsive:** Establish your presence and reliability
- **Investigative:** Get to the “real” problem
 - When a user asks for something – what do they *really* need?
 - When issues keep happening – what needs to be fixed systemically?
- **Proactive:** Look out for researcher’s interests
 - Reach out to researchers with problems; frame issues as something that prevents them from getting research done. Emphasize good citizenship!
 - Connect users doing similar work



The Big Picture

The goal is always a better research outcome for the researcher.

- **Choosing a “best-fit” computing approach that scales as much as possible**
 - HTC workloads compatible with the OSG can get a lot of computing.
- **Communicating the benefits and effort investment of that approach**
 - Is running on the OSG as “simple” as your desktop? Nope! Is it more powerful? Yes!
 - Quantify things! “You can achieve 10,000 runs in 2 days.” “You’re starting with 4,000, but you could easily run 40,000!”
- **Always keep this goal in mind when troubleshooting and providing guidance.**
 - Before solving a researcher’s question, stop and ask: is this the *real* problem?
 - Introduce skills that will help researchers help themselves.



Final Thoughts



Facilitation (academia?) can feel like this...



Gif by Aardman Animations (@aardman) on giphy.com
<https://media.giphy.com/media/3oz8xtBx06mcZWoNJm/giphy.gif>



Making the Best of It

- Prioritize, prioritize, prioritize!
- Invest time in scalable resources (like documentation)
- Make it a community effort - students? volunteers? study group? a wiki?



print by Star Shaped Press (<https://www.starshaped.com/>)

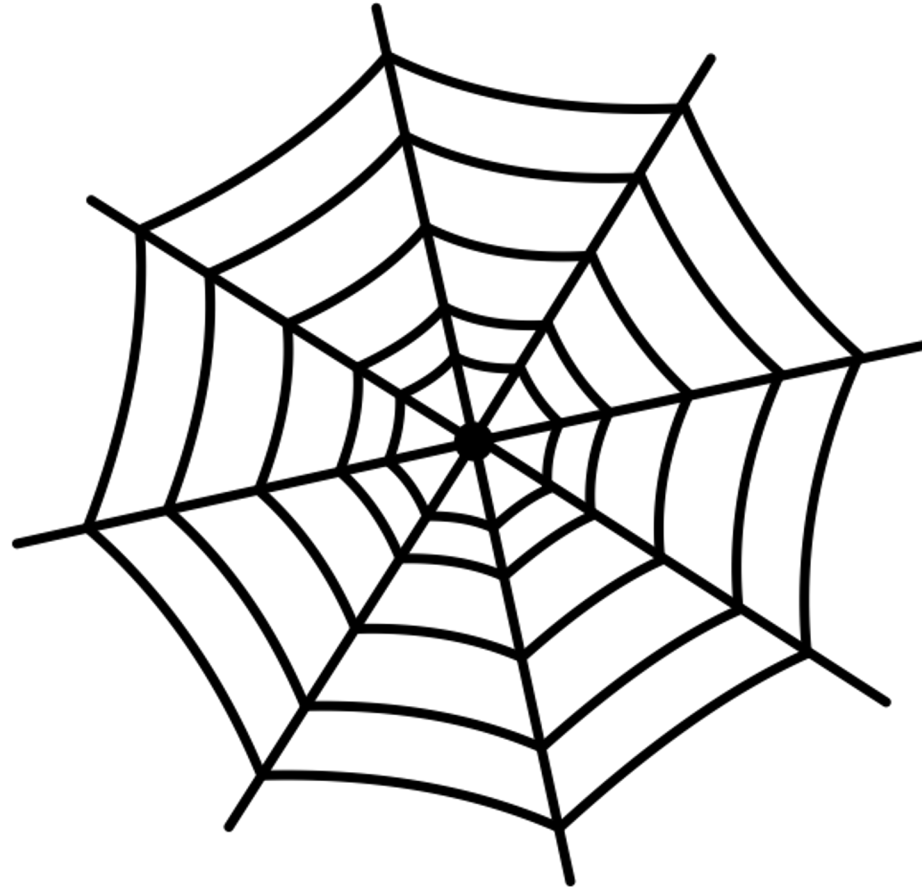


Join Networks of Support

Group

- Colleagues
- Department or College

Campus/Institution



National

- Campus Champions
- CaRCC People Network
- MS - CC

“Regional”

- e.g. Great Plains Network



Working Together

- You are part of OUR network of support and community of practice in high throughput computing
- We can be more effective in transforming science and improving HTC technologies with individual advocates like you.
- Let us know how we can help you!!

support@osg-htc.org



Acknowledgements

- This work is supported by [NSF](#) under Cooperative Agreement [OAC-2030508](#) as part of the [PATh Project](#). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

