

Round Table Discussions June 02, 2021

Open discussions on specific topics selected by the [Software Working Group](#) and selected from the list of [SWG Topics For Discussion](#).

Wednesday, June 2, 2021 - **HPC** - Moderated by Steve Peckins

Recording: <https://uofi.app.box.com/folder/136455936499>

Attendees

- Galen Arnold
- Luigi Marini
- Steve Peckins
- Craig Steffen
- Charles Blatti III
- Mark Van Moer
- Santiago Nunez Corrales
- Chris Navarro
- Jong Lee
- Dean Karres
- Roland Haas
- Mickolaj Kowalik
- Sandeep Satheesan
- Todd Nicholson
- Peter Groves
- Chen Wang
- Matt Berry
- Robert Brunner
- Michael Bobak
- Kaveh Karini Asli
- Bill Kramer
- Jim Phillips
- Gowtham Naraharisetty
- Michal Ondrejcek
- Elizabeth Yanello

Discussion:

- Galen shared NCSA Allocations Page: <https://wiki.ncsa.illinois.edu/display/USSPPRT/NCSA+Allocations>
- If you'd like to request time, use this link except for Blue Waters, which has its own link: <https://bluewaters.ncsa.illinois.edu/aboutallocations>
- XSEDE allocations may be found here: <https://www.xsede.org/ecosystem/resources>
- XSEDE submissions: Start here: <https://portal.xsede.org/>
- There are two kinds of allocations, continuous and quarterly through XRAC
- Steve shared the submissions page and opportunities available.
- Sandeep mentioned that there are example pages of submissions.
- Dean Karres is the Chair for Campus Champions for allocations for explorations. Your applications for these allocations need to show a need to large data needed to be scaled. Dean also mentions class based use for campus cluster.
- AWS is a very expensive, but if you have 100 or 1,000 nodes. If you want to use a small amount of nodes, the cloud is a good place for that.
- The hallmark of a true HPC system is its large capacity and large memory nodes.
- Galen gave technical information on HPC systems. Please see recording for details.
- In the cloud you tend to have storage buckets; with HPC is also extremely fast.
- It seems a lot of researchers are using SLURM sbatch for HPC submissions and workflows
- Roland Haas mentioned that slurm is a conventional system; however, containers and clusters on the cloud is much more elastic.
- If you are using containers, you don't need to upload modules. One image per node is usual.
- Another advantage of containers you are presenting to the cluster what the cluster is supposed to do. If you are using Python, if it's not in a container, each image is loaded individually (1,000's and 1,000's of little images!)
- Dean noted that those that are submitting jobs, in the conda model, be aware that you of the services in each queue, otherwise, it will spin and go nowhere. The Beckman Queue has only 9 nodes.
- Spack - A flexible package manager supporting multiple versions, configurations, platforms, and compilers.
- Dean notes that spack is "magic". For porting software, if it is on spack, it will reduce the amount of time you spend in dependency "hell".
- Steve notes that there is an ARM-based supercomputer at Stonybrook that should be allocated through XSEDE at some point in the future.
- NCSA will deploy spack onto the upcoming Delta Xsede gpu-based system.
- Craig notes that spack is integrated, at some level it will try to install new versions.
- Beyond say XSEDE there's a number of cost-recover type infrastructure at UIUC. Campus cluster was mentioned, but also the VM clusters (current and future), storage services. No HPC typical system I would be aware of (like eg the XSEDE OSG nodes).
- NCSA staff have easy access to HPC, but others need to give explanations of their need for allocations. The same is true for Campus Cluster.
- NCSA supports long term usage, but it could be slower because so many people are using this.
- There is a secondary queue if someone has left a container, you may need to wait a week to get permission to use. This queue is severely time limited.
- Peter asks the question: Is there any portability within the various HPC libraries? Galen says NO. You are committing to that container. He mentioned ARM, AMD, and X86 and Nvidia as other devices that don't use the HPC libraries.
- Sandeep says YES, sort of. There is limited portability depending on the containers you are using.
- It is a common complaint that the software old. To defend the "oldness". HPC does not use last week's latest shiny thing, because it is not a laptop. It is a complicated dependency stack that needs redundancy. Think of HPC as having "exotic" tools inside that may be 4 or 5 years old.

Links Shared During the Talk:

NCSA Allocations

http://www.ncsa.illinois.edu/user_support

<https://bluewaters.ncsa.illinois.edu/aboutallocations>

<https://www.xsede.org/ecosystem/resources>

<https://portal.xsede.org/>

https://campuscluster.illinois.edu/new_forms/user_form.php

<https://hub.docker.com/search?type=image&architecture=arm>

<https://spack.io/>

<https://www.stonybrook.edu/commcms/iacs/research/projects/Ookami>