

Experimental Computing Laboratory Meeting October 24, 2024, 12:30 PM

Advanced Computing Systems Research Section

Jeffrey Vetter, Section Head

Steve Moulton, Systems Engineer

Aaron Young, Software Engineer

ORNL is managed by UT-Battelle LLC for the US Department of Energy



U.S. DEPARTMENT OF
ENERGY

Infrastructure/System Status

- Ubuntu 24.04 announced, but not approved for ORNL use. Will not start installations or upgrades until this happens.
 - Not considered stable; rolling upgrades generally broken.
 - Will be installed on request if needed
- All user facing systems but two have been migrated to Ubuntu 22.04
 - One Centos 7 system left against need – will reconfigure if no interest
 - One Rocky 9 system built for user to test in that environment
 - All rebuilt systems are configured for cgroups v2.
 - If you need v1, see ExCL staff

Plans for next three months

- Continuing documentation improvements
 - <https://docs.excl.ornl.gov>
 - System use and assignment document in progress, will soon be posted to documentation site
- Stabilize non-ORNL-managed UID & GID assignments ExCL-wide to simplify service deployment
 - Also has implications for docker groups
 - Additional requirement: olama install requires a stable UID/GID.
- Deploy Mi300a (4 gpu integrated) system. On 5100 dock awaiting installation. Will deploy in 5100 data center to ease modifications (as needed for heterogeneous accelerator environment)

Mi300a Specs

- 4 * *Mi300A system*
 - *APUs (Integrated CPUs and GPUs)*
 - *128 GB HBM3 memory on each*
 - *Infinity Fabric Links (8)*
 - *PCIe 5.0x16 for peripherals (and who knows what else).*
 - <https://www.amd.com/en/products/accelerators/instinct/mi300/mi300a.html>

Reminder: SSH Keys for Authentication

- Using SSH keys is the preferred way to authenticate your user and to authenticate with private Git repositories.
 - For security, it is recommended to use an SSH keys encrypted with a passphrase.
- Why not passwords?
 - ExCL will block your account after 3 failed attempts. Automatic login tools, e.g. VS Code, can easily exceed this limit using a cached password and auto-reconnect.
 - For git repos with two-factor authentication, an application token/password must be created, and this password must be stored externally and is more cumbersome to use.
- How to get started?
 - Set up a key pair:
 - [Visual Studio Code Remote Development Troubleshooting Tips and Tricks](#)
 - [Generating a new SSH key and adding it to the ssh-agent - GitHub Docs](#)
 - [Add key to Git Hosting Websites.](#)
 - [Setup ExCL worker node proxy via login node.](#)

Proposed Host layout Change: Milan0

- Currently has two Nvidia A100 GPUs
- Proposal
 - Move 2 AMD M100 GPUs from Cousteau to Milan0
 - Cousteau would then become general compute server
 - Deploy Xilinx U280 in Milan0
 - Requested Intel ARC GPU
- This would provide a nice heterogeneous environment, as well as moving M100s to Pcie5 (which they may or may not be able to take full advantage of).
- Comments?

IRIS Setup scripts

- Setup scripts for ExCL systems that setup the environment and tools for IRIS accelerator use are available at `/auto/software/iris` and used via:

```
source /auto/software/iris/setup_system.source
```

- To request changes to these scripts, fork the repo [Brisbane / ExCL IRIS Setup Scripts](#), make the modifications, and open a merge request. Once we review the merge request, we will merge the changes and deploy the scripts to ExCL.

Host Quarantine

- ITSD compliance enforcement mechanism mandated by DOE
- Only directly affects ExCL systems that have ORNL-visibility
 - Login
 - Infrastructure servers (service, ansible)
 - Hypervisors with VMs requiring ORNL access
 - These systems are all currently compliant, except
 - Old hypervisor (to be reimaged and dropped off of ORNL networks) and graylog and check_mk (awaiting reinstallation).
 - We are carefully tracking quarantine notifications (none yet). ITSD is showing flexibility in allowing time for corrective action.
- Internal systems are maintained to the same standards

Container Registry Moved from GitLab to Harbor

- See [Container Registry \(ornl.gov\)](https://ornl.gov/container-registry).
- Two Harbor Sites:
 - Internal (servers.moderate): camden.ornl.gov
 - External (DMZ): savannah.ornl.gov
- Training Links:
 - [Lunch & Learn](#) Recording
 - [Slides](#)
- Harbor also works with Apptainer. See [Apptainer | ExCL User Docs \(ornl.gov\)](#).
- This is an ITSD resource, not an ExCL-maintained resource.

Questions/Discussions/Comments?

- Can use Harbor for IRIS?
 - IRIS not currently have a container. Usable for this.
 - Narasinga & Aaron have identified issue that will get resolve.
 - Initially 50Gb storage limit (per user?)
- HiFive in ExCL? Will take up with Brett W.
- Monthly Newsletter