



AUGUST 20, 2026 | OAK RIDGE TENNESSEE

Experimental Computing Laboratory (ExCL) Monthly Meeting

PRESENTED BY

Steve Moulton, Aaron Young, Jeffrey Vetter

Advanced Computing Systems Research Section



U.S. DEPARTMENT
of **ENERGY**

ORNL IS MANAGED BY UT-BATTELLE LLC
FOR THE US DEPARTMENT OF ENERGY



Policy Changes — High-utilization Systems

ExCL resources are seeing much higher utilization as our user base grows and more work is being done in the AI space. Unfortunately, this has led to resource contentions. Over the next month we are considering the following changes.

- Popular accelerator systems will be restricted to interactive slurm access. This allows finer-grained allocation policies (e.g., allocate based on how many GPUs you need) and ease off-hour exclusive allocations when needed.
 - This will be well-documented.
- Documentation will be updated to indicate what resources are available for non-accelerator applications.
 - Case in point: one very popular system was being used for a large design space search which made no use of accelerators, but which affected accelerator access for others.
 - Example: The triple crown servers (affirmed, justify, pharaoh, secretariat) have 128 physical cores (configured 256 virtual cores) and 1 TB physical memory. These powerful systems are well suited to larger problems but are currently underutilized.

Policy Clarification — NVHPC

Systems engineering staff is being careful to minimize NVHPC (including cuda) and driver updates.

As a part of general operating system updates, nvidia and amd driver updates are automatically applied. This assures system stability and security. The downside is that sometimes the drivers are not reloaded automatically. Should you encounter this please let excl-help@ornl.gov know as soon as observe. Often this can be remediated without a reboot. Nvidia-smi will complain of driver mismatch (for example) if this is the case.

NVHPC is updated as necessary or requested. Updating NVHPC is not a part of our automatic update process, but is done on request, or as new features in either NVHPC or in Linux run time libraries mandates it.

Should a NVHPC update negatively affect you, please contact us. There are several approaches to remediating this, include rollbacks and virtualization.

Policy Changes — Login services

- New resource limits on the ExCL login server. We have implemented per-user resource limits on login to prevent individual sessions or processes from consuming too many GPU or memory resources which could impact others.

Each user is now limited to the following aggregate resources across all processes running under their login sessions:

CPU: 1 CPU core worth of aggregate CPU time

Memory pressure threshold: 1 GiB

Maximum memory: 1.5 GiB

Maximum swap: 512 MiB

Word of Caution: If you exceed your quota on the login node and will also affect the responsiveness of your connections to the worker nodes as the login node is your jump host to these other systems.

- This new limit does mean that a VS Code server with extensions will no longer be able to run on the login node. For information on how to directly connect to a worker node using VS Code, see [Visual Studio Code | ExCL User Docs](#).
 - Good systems for running a VS Code server for general use include oswald, odswald00, oswald02, and oswald03, as they are generally underutilized.

Claude Code in ExCL Updated with Codex Support



OpenAI Codex

- [Claude Code | ExCL User Docs](#) — ExCL's new Quick-Start Guide for Claude and Codex.
- Soon we will be provisioning a dedicated VM for running Claude. In the meantime, launch your Claude sessions on Explorer and use Slurm to launch jobs that need access to specific accelerators.
(Claude likes to launch sub-agents and other long running jobs that persist after exiting the TUI.)

Quickstart

Clone the repository and run Claude Code in three steps:

```
git clone git@code.ornl.gov:7ry/claude-hpc.git
cd claude-hpc
make install
```

Then launch Claude Code from any repository:

```
cd /path/to/repository
claude-hpc .
```

 To use Slurm commands (`squeue`, `sbatch`, etc.) or Lmod modules inside Claude, switch to the `slurm` branch before building:

```
git checkout slurm
```



Planning Move to TRC (3700) Lab A102

- Progress is being made on moving into the A102 lab space.
- Benchtop systems will need to be powered down and packed up to be ready for the move.
 - This will impact all benchtop FPGA systems and desktop systems including the Zenith systems.
 - The move will be planned for early to mid September. Moves using ORNL movers typically take ~14 days. I will plan to pack up loose accelerators first, then delay powering down the enclosed systems until the narrower moving window.

Shared Large Apptainer Image Space

- If you have small (<50G) container images they can be stored and shared via [Harbor](#).
- However, if you have larger than 50G, general purpose images, they can be shared via the Apptainer shared project space.
- To write to this space you need to be added to the group, but any ExCL user is able to read from this space.

nvtop

nvtop is being deployed to all systems. This is a simple performance visualization tool that is supported on NVIDIA, AMD and Intel GPUs. It is also a quick way to see what the GPU architecture (if any) is on any system. This will be deployed as a part of the excl-utils module, and is already natively available on several systems (like maxwell, shown here).



Questions/Projects/Comments/Discussions

- Checkmk monitoring of cgroup resource limit counters for login limits.