



CLAIX-2023 HPC

RWTH Aachen, Germany,
26 February - 01 March 2024

<http://www.vi-hps.org/training/tws/tw44.html>

Marc-André Hermanns
RWTH Aachen University

Agenda

- Hardware Overview
- Software Stack
- SLURM Batch Scheduler
- Filesystems
- Porting Considerations

System Configurations – CLAIX @ RWTH Aachen University

	CLAIX-2018 MPI	CLAIX-2023 HPC
Nodes	1243	632
Sockets	2	2
Processor	Intel Skylake	Intel Sapphire Rapids
Cores per Socket	24	48
Cores per Node	48	96
NUMA Domains per Sockets	2	4
Main Memory per Node (GB)	192	256 (470) / 512 (160) / 1024 (2)
Network	OmniPath	Infiniband

CLAIX-2023 Login

- Login Nodes:
 - login23-1.hpc.itc.rwth-aachen.de
 - login23-2.hpc.itc.rwth-aachen.de
 - login23-3.hpc.itc.rwth-aachen.de
 - login23-4.hpc.itc.rwth-aachen.de
- VPN needed from outside RWTH Aachen Campus
- MFA setup via RegApp at <https://regapp.itc.rwth-aachen.de>
 - Second Factor Tokens (NB: Create TAN List & Smartphone Token)
 - SSH Key Upload & Activation
 - Otherwise Password & MFA on **every** login
 - Classic ./ssh/authorized_keys not enabled
 - Re-authentication needed after 10 hours
 - Needed for every distinct host.

CLAIX-2023 Software Stack

- Software Stack is completely rebuild for CLAIX-2023
 - Software is located in different paths
 - Software versions installed may differ significantly from CLAIX-2023
 - **Recommendation:** Rebuild software specifically for CLAIX-2023 where possible
- Lmod Module System
 - Main Toolchains: intel & foss
 - Default toolchain: intel/2022a (VI-HPS Workshop focuses on intel/2022b – need to reload)
 - Using hierarchical modules
 - Modules may become visible only after loading compiler and MPI modules
 - Use `module spider` to search for software
- Purge modules completely when changing toolchains
 - Use `module purge; module load <new-toolchains>`

SLURM Configuration – Test Phase

- CLAIR-2023 Queue c23test is hidden
 - Will not show in standard `sinfo` or `squeue` output
 - Needs to be specifically referenced
 - `sinfo -p c23test`
 - `squeue -u $USER -p c23test`
- Access currently restricted to testing users (and workshop participants)
 - Check if you are in `supp0006` group via `groups` command
 - Specify SLURM account explicitly: `-A supp0006`
 - Specify SLURM partition explicitly: `-p c23test`
 - (Already set in VI-HPS Tutorial environment – see last slide)

Filesystems available on CLAIX-2023

- \$HOME via NFS
 - Backup: Snapshots available vis \$HOME_SNAPSHOTS
 - Slowest filesystem
- \$WORK via NFS
 - Backup: Snapshots available via \$WORK_SNAPSHOTS
 - Potentially faster than \$HOME
- \$HPCWORK via Lustre
 - Currently not yet fully connected (**NOT RECOMMENDED!**)
- \$BEEOND via BeeGFS on Demand
 - Generated ad hoc within the batch job (needs `#SBATCH --beeond` parameter in batch script)
 - Needs staging in and out of data as part of batch job

What to consider when porting applications?

- Twice the number of cores per node
 - Adapting NODES and TASKS-PER-NODE in batch scripts
- More cores can communicate over shared-memory
 - Application behavior due to imbalances may change
- Different number of NUMA domains
 - Sub-NUMA Clustering (SNC) enabled
 - Access within NUMA domain faster, across NUMA domains more costly
 - Try to keep all threads of a process within the same NUMA domain
- E.g., for Hybrid Setups: 1 process per NUMA domain, 12 threads per process

VI-HPS Tutorial Environment on CLAIX-2023

- Login to cluster and source VI-HPS environment to set variables and load modules

```
$ source /home/hpc/vihps-tw44/setup.sh
INFO: VIHPS_ROOT=/home/hpc/vihps-tw44/
INFO: VIHPS_WORKSPACE=/work/mh269604/vihps-tw44/
INFO: SBATCH_ACCOUNT=supp0006
INFO: SBATCH_RESERVATION=vihps-tw44
INFO: SBATCH_PARTITION=c23test
INFO:
INFO: Change to your workspace with:
INFO: $ cd $VIHPS_WORKSPACE
INFO:
[INFO] Module intel/2022b loaded.
[INFO] Module likwid/5.3.0-d8fea29 loaded.
[...]
$
```