

BASE: TalTech’s HPC Infrastructure 2020–2024

Heiko Herrmann* , Toomas Kaevand, Lauri Anton
HPC Centre, Tallinn University of Technology (TalTech)
firstname.lastname@taltech.ee

DOI: 10.48726/8f0v8-1eb97 (this version)

DOI: 10.48726/zf23z-8ba50 (all versions)

Abstract

Many scientific and engineering disciplines are now relying on or are supplemented by large scale numerical simulations or data analysis. These fields include physics, chemistry, CFD, climate and ocean modeling, pollution transport, virus spreading, building modeling and training for deep learning and big data analysis. This requires access to suitable high performance, high throughput or parallel computing environments.

Some of these fields can largely benefit from data parallel architectures like GPUs, others are better suited for regular CPUs.

TalTech’s HPC center provides a heterogeneous environment suitable for diverse applications.

1 Introduction

Compute clusters have been build and operated independently by several groups at Tallinn University of Technology for about two decades. For example the Institute of Cybernetics had the first cluster since the early 2000s and replaced and enlarged it in 2007 within the Center of Excellence CENS. The group of theoretical chemistry operates their own cluster for a similar duration.

In 2013 the first central HPC cluster of the university was procured, administration was part of the IT service department.

In 2019 the HPC Centre was established as a separate unit within the IT College.

Scientific fields using the cluster are theoretical chemistry, system biology, engineering, marine

systems, computational fluid dynamics, mechanics of materials, AI, gene technology, bio medicine, etc.

2 The BASE Cluster

The BASE cluster is a part of Estonia’s national scientific research infrastructure and provides a stepping stone for the access to LUMI [1] and international supercomputing centers. The infrastructure is used to solve many scientific and engineering questions on-site, but can also be used for prototyping and scaling experiments in pre-stages for the access to supercomputing centers.

Currently, BASE is a merger of the former HPC1 and HPC2 clusters installed in 2013 and 2019, respectively, supplemented with a GPU server and visualization server. The heterogeneous cluster can now be accessed through a single login node.

Due to the different investment phases the network topology of the merged cluster environment is complicated, see Fig. 1. All nodes are connected with a 1 Gb/s management network (not shown in the figure), additionally, the gray nodes are connected with a 1 Gb/s network used for computation and filesystem access, and a subset of these (the gray-ib nodes) are additionally connected with a QDR InfiniBand network, which also connects to the viz and mem1tb nodes. The green nodes are connected with a 25 Gb/s ethernet network for computation and filesystem access and a subset of these (the green-ib nodes) are additionally connected with an FDR InfiniBand network. The 25 Gb/s ethernet network has a largely redundant layout, the nodes connecting to two 100 Gb/s switches in parallel.

In 2022 the 30 TiB \$HOME-filesystem provided

by CEPH, and a 130 TiB scratch filesystem provided by BeeGFS was replaced by a new Spectrum Scale fileserver providing 2 PiB storage using GPFS.

2.1 Software on the Cluster

The operating system on the cluster is CentOS 7 with EPEL, with the exception of the GPU server (Ubuntu 20.04) and the visualization server (Debian 10).

The application software is installed in one of three ways:

- from the package manager of the distribution (if recent version is available)
- from the SPACK package manager
- manually from installer or compiled from source

The different versions in the user environment are managed using Lmod [2].

There are three installations of SPACK maintained on our cluster:

- for CentOS, with optimized software for SandyBridge and Skylake_avx512 processors
- for Ubuntu, with optimized software for Zen2
- for Debian, with optimized software for Ivy-Bridge.

The software installed for the different operating systems differs, and is selected for the use-case of the system. Since the Debian machine is the visualization system and not a computational node, the main focus is on visualization and data analysis software, similarly software for GPU computing is only available on the Ubuntu installations.

The main compiler is GCC-10.3.0 build with OpenACC and OpenMP GPU offloading, but LLVM and Nvidia HPC SDK (PGI) are available, too.

The Centre has a “bring your own license” policy for commercial software. If users have the license and want to use the software on the cluster, we can install it with corresponding access restrictions.

2.1.1 Computational Chemistry

The following software is used on our cluster, users bring their own licenses for commercial software:

- nwchem
- cp2k
- gaussian
- Turbomole
- ORCA
- AmberTools21
- MultiWFN
- Molden

2.1.2 Engineering

The following software is used on our cluster, users bring their own licenses for commercial software:

- OpenFAOM
- ElmerFEM
- FreeFEM
- CalculiX
- SU2
- Salome
- Gmsh
- Abaqus
- LSDyna
- StarCCM+
- Comsol

2.1.3 Development tools

The following software is used on our cluster, users bring their own licenses for commercial software:

- GCC (on GPU nodes incl. OpenACC offloading)
- LLVM (on GPU nodes incl. OpenMP offloading)
- Nvidia HPC-SDK on GPU nodes

Name	#Nodes	Model	Processor model	#S/ $\frac{C}{S}/\frac{T}{C}$	Memory	#GPUs	GPU model	Network
base	1	Fujitsu	Intel Xeon Silver 4114	2/10/2	188 GiB			25 GBit/s Eth
green	14	Fujitsu	Intel Xeon Gold 6148	2/20/2	96 GiB			25 GBit/s Eth
green-ib	18	Fujitsu	Intel Xeon Gold 6148	2/20/2	96 GiB			FDR, 25 GBit/s Eth
gray	100	SuperMicro	Intel Xeon CPU E5-2630L	2/6/2	48 GiB			1 GBit/s Eth
gray-ib	48	SuperMicro	Intel Xeon CPU E5-2630L	2/6/2	64 GiB			QDR, 1 GBit/s Eth
amp	1	SuperMicro	AMD EPYC 7742	2/64/2	1 TiB	8	Nvidia A100 40 GiB	100 GBit/s Eth
amp2	1	SuperMicro	AMD EPYC 7713	2/64/2	2 TiB	8	Nvidia A100 80 GiB	100 GBit/s Eth
mem1tb	1		Intel Xeon CPU E5-4640	2/8/2	1 TiB			QDR, 1 GBit/s Eth
viz	1	Fujitsu	Intel Xeon CPU E5-2630L v2	2/6/2	64 GiB	2	Nvidia Tesla K20Xm 6 GiB	25 GBit/s Eth, 1 GBit/s Eth

Table 1: S: Sockets (CPUs), C: Cores, T: Threads; configuration

- Valgrind
- Extrae
- TAU
- Score-p
- Scalasca
- HPCToolKit

2.1.4 Programming languages, data analysis

The following software is used on our cluster, users bring their own licenses for commercial software:

- C/C++ (GCC, LLVM, Nvidia/PGI)
- Fortran (GCC, LLVM, Nvidia/PGI)
- Chapel
- Julia
- SBCL
- Matlab
- GNU Octave
- GNU R
- mawk, gawk, nawk, bioawk

2.2 The Visualization system

A special visualization node is connected to the HPC BASE cluster, this is equipped with graphic cards for remote visualization and can be used in different ways. The provision of the special visualization node means, that the produced data can be directly visualized and analyzed in the HPC environment, without the need to copy or move

data to the user’s desktop. This avoids unnecessary data duplication and speeds the analysis up. While regular ssh forwarding is of course available, more efficient methods like VirtualGL [3], TigerVNC [4] and X2GO [5] are recommended, as these can make better use of the available network bandwidth and the graphic cards in the visualization system.

The visualization system has special software installed for the visualization of large data sets, this software includes ParaView [6, 7], VisIt [8], COVISE [9] and 3D-Slicer [10].

2.3 Connected Resources of Other Departments

2.3.1 AI lab storage and software

The storage and software of the AI lab is available on the GPU nodes.

2.3.2 MSI storage

The storage server of the Marine Systems Institute is mounted on all compute nodes.

2.3.3 Kyb3 visualization

There is also the possibility to use the 3D visualization systems of the Rheology of Composites group of the Department of Cybernetics, which consist of a graphics workstation with 3D monitor and the Kyb3 [11], a semi-immersive 3-wall 3D environment.

3 The OpenStack ETAIS cloud

The HPC Centre additionally operates a small OpenSTACK cloud, where users can install their own software and services in CentOS, Debian and Ubuntu images. The cloud consists of 5 nova

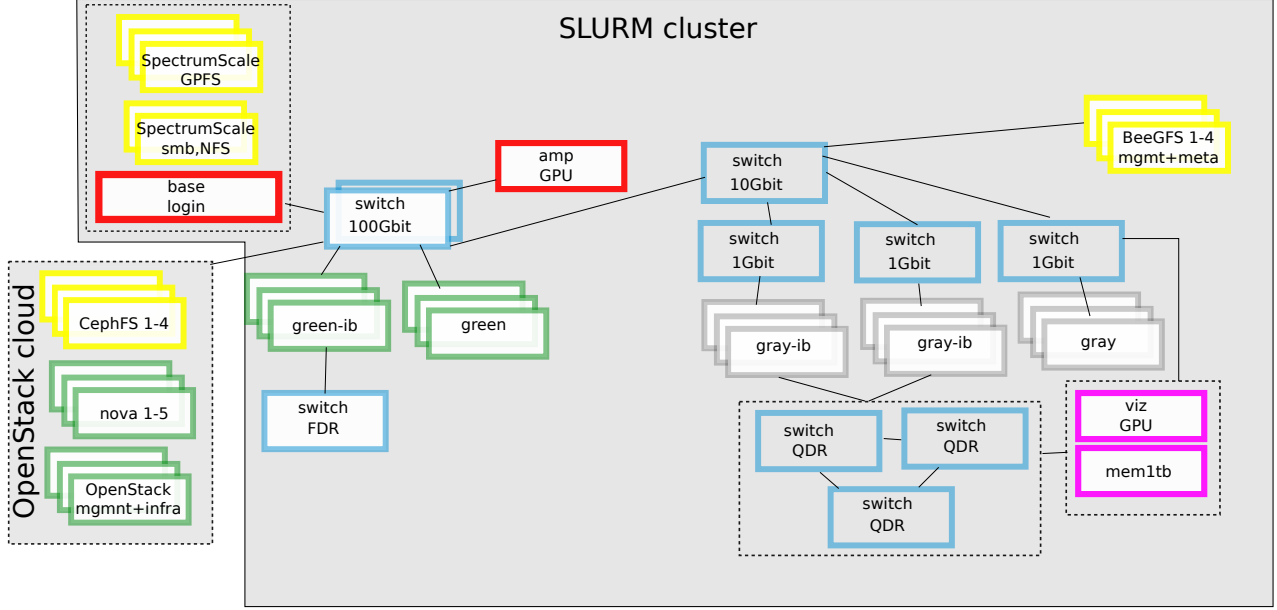


Figure 1: The topology of the HPC infrastructure. Color coding: blue: switches, green: “green” nodes, gray: “gray” nodes, red: login nodes, magenta: special purpose nodes, yellow: file system services

Type	#Nodes	Model	Processor model	CPU threads	#S/ $\frac{C}{S}$ / $\frac{T}{C}$	Memory	Network
Nova (compute)	5	Fujitsu	Intel Xeon Gold 6148	80	2/20/2	768 GB	25 Gbit
management storage	4						

Table 2: S: Sockets (CPUs), C: Cores, T: Threads; configuration

servers and an attached 65 TB CEPH storage system with 4 servers, see Table 2.

faster on an HPC system, even single-threaded software.

4 Research Outcomes

Research outcomes include CFD simulations of composite materials [12], COVID-19 [13] and computational chemistry [14].

5 Observations from user-support and administration

New users have difficulties to get used to the command-line interface.

Novice users often expect, that software runs

In cases of GPUs, they assume that they get “magically” used and do not know that the software needs to be specifically programmed to be able to make use of GPUs.

In case of parallel computing, the predominant idea among novice users is more threads (more tasks) make the program faster, they don’t believe that a program running with fewer tasks can actually be faster than one with many tasks. The concept of having to do test runs with different task counts is very difficult to teach to novice users.

Acknowledgement

Support by the EuroCC2 project.

References

- [1] LUMI supercomputer. <https://lumi-supercomputer.eu/>.
- [2] R. McLay, K. W. Schulz, W. L. Barth, and T. Minyard. Best practices for the deployment and management of production hpc clusters. In *State of the Practice Reports, SC11*, page 9:1–9:11, November 2011.
- [3] VirtualGL. <https://virtualgl.org/>.
- [4] TigerVNC. <https://tigervnc.org/>.
- [5] X2GO. <https://wiki.x2go.org/>.
- [6] James P. Ahrens, Berk Geveci, and C. Charles Law. Paraview: An end-user tool for large-data visualization. In *The Visualization Handbook*, 2005.
- [7] Utkarsh Ayachit. *The ParaView Guide: A Parallel Visualization Application*. Kitware, Inc., Clifton Park, NY, USA, 2015.
- [8] Hank Childs, Eric Brugger, Brad Whitlock, Jeremy Meredith, Sean Ahern, David Pugmire, Kathleen Biagas, Mark Miller, Cyrus Harrison, Gunther H. Weber, Hari Krishnan, Thomas Fogal, Allen Sanderson, Christoph Garth, E. Wes Bethel, David Camp, Oliver Rübel, Marc Durant, Jean M. Favre, and Paul Navrátil. VisIt: An End-User Tool For Visualizing and Analyzing Very Large Data. In *High Performance Visualization—Enabling Extreme-Scale Scientific Insight*, pages 357–372. Oct 2012.
- [9] Covise: collaborative visualization and simulation environment. <https://github.com/hlrs-vis/covise>.
- [10] Steve Pieper, Michael Halle, and Ron Kikinis. 3d slicer. IEEE International Symposium on Biomedical Imaging ISBI 2004, April 2004.
- [11] Emiliano Pastorelli and Heiko Herrmann. A small-scale, low-budget semi-immersive virtual environment for scientific visualization and research. *Procedia Computer Science*, 25(iii–iv):14–22, September 2013.
- [12] Oksana Goidyk, Mark Heinštein, and Heiko Herrmann. CFD comparison of the influence of casting of samples on the fiber orientation distribution. *Fibers*, 11(1), 2023.
- [13] Jaan Kalda, Mart Ratas, Taavet Kalda, Azer Ramazanli, Heiko Herrmann, and Juri Belikov. The importance of superspreaders on the spread of the global covid-19 pandemic. *Research Square*, 2021.
- [14] Irina Osadchuk, Hanna-Eliisa Luts, Karolis Norvaiša, Victor Borovkov, and Mathias O. Senge. Supramolecular chirogenesis in a sterically hindered porphyrin: A critical theoretical analysis. *Chemistry – A European Journal*, 29(49):e202301408, 2023.