

**Foundations of Distributed Systems HS 2021 (45402-01)****Fall Semester 2021**

# Usage of miniHPC

**Objectives:**

1. Understand the system architecture of the miniHPC cluster
2. Manage job submissions on the miniHPC cluster

## What is miniHPC?

It is a small HPC cluster that has a peak performance of 28.9 double precision TFLOP/s. The miniHPC has two types of nodes (see Table 1), Intel Xeon nodes and Intel Xeon Phi Knights Landing (KNL) nodes (see Table 2). The Intel Xeon nodes consist of 22 computing nodes, 1 login node, and 1 node for storage. The Intel Xeon Phi nodes consist of 4 computing nodes.

Table 1: miniHPC nodes information

Count	CPUs
22	2 Intel Xeon E5-2640
4	1 Intel Xeon Phi 7210

Table 2: CPU information

CPU model	Count	CPU speed (GHz)	Cores	Threads	RAM (GB)	Cache
Intel Xeon E5-2640 v4 <sup>a</sup>	44	2.4	10	20	64	L3: 25MB
Intel Xeon Phi 7210 <sup>b</sup>	4	1.3	64	256	128	L2: 32MB

<sup>a</sup> [http://ark.intel.com/products/92984/Intel-Xeon-Processor-E5-2640-v4-25M-Cache-2\\_40-GHz](http://ark.intel.com/products/92984/Intel-Xeon-Processor-E5-2640-v4-25M-Cache-2_40-GHz)

<sup>b</sup> [http://ark.intel.com/products/94033/Intel-Xeon-Phi-Processor-7210-16GB-1\\_30-GHz-64-core](http://ark.intel.com/products/94033/Intel-Xeon-Phi-Processor-7210-16GB-1_30-GHz-64-core)

All nodes are interconnected through two different types of interconnection networks (see Table 3). The first network is an Ethernet network with 10 Gbit/s speed; reserved for users and administrators access. The second network is the fastest network, an Intel Omni-Path network with 100 Gbit/s speed; reserved for the high-speed communication between the computing nodes.

The 28 nodes of the miniHPC cluster are interconnected in a two-level fat-tree topology, illustrated in Figure 1.

Table 3: miniHPC interconnection networks

Network fabrics	Network speed
Intel Omni-Path	100 Gbit/s
Ethernet	10 Gbit/s

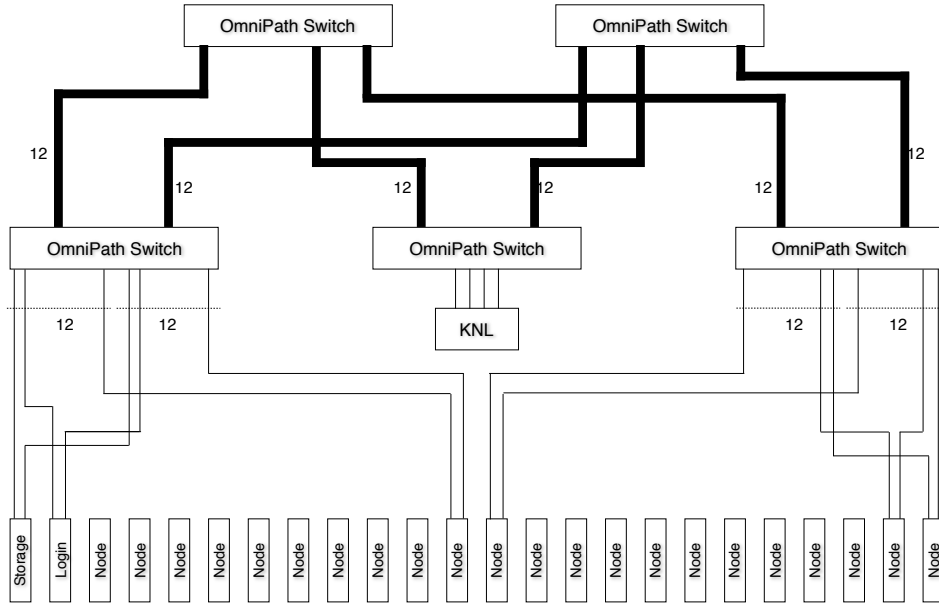


Figure 1: A graphical illustration of the miniHPC fat-tree topology. Number of nodes: 28 nodes (24 Intel Xeon and 4 Intel Xeon Phi (KNL)), number of switches: 5 switches, number of links: 196 links.

## How to login to the miniHPC cluster?

Prerequisites:

- a. Internet access
- b. ssh client software
  - For Mac and Linux machines, one can just use the terminal
  - For Windows machine, one can install PuTTY or MobaXterm

Login Procedure:

Type into your ssh client

- `ssh -X UserName@cl-login.dmi.unibas.ch`
- It will ask for a password, you must use the same password as your university email password

After successfully login, you will be at your home directory on the miniHPC login node.

Note that miniHPC is a shared computing facility with limited resources. Therefore, we employ a batch scheduling software to deploy the execution of user jobs (e.g., your applications) on it.

## What is a batch scheduling system?

The batch scheduling system (usually called job scheduler) is a single point to interface with the cluster, to which users can request one or more of the cluster resources to execute their programs (jobs). The job scheduler allocates resources to different users based on an allocation policy defined by the cluster administrators. There are many job schedulers, and on miniHPC we use Slurm.

## How to use Slurm to manage job submission?

There are 3 basic commands:

Command	Format	Description
sbatch	<code>sbatch [ options ] [ scriptfile   [ script args ] ]</code>	To submit jobs to the scheduler
scancel	<code>scancel [ -u user ] [ Job_ID ] ...</code>	To cancel running or queued jobs
squeue	<code>squeue [ -u user ] ...</code>	To display the status of jobs

# miniHPC usage hands-on:

- a. To copy files: **These commands run only on your local machine**
  - To miniHPC: `scp certain_file username@cl-login.dmi.unibas.ch:destination_path`
  - From miniHPC: `scp username@cl-login.dmi.unibas.ch:certain_file destination_path`
- b. To login
  - `ssh -X UserName@cl-login.dmi.unibas.ch`
  - Once logged in, you will be redirected to your home directory
- c. To use Intel tools including the Intel compiler and Intel MPI
  - Type `ml intel`
  - Compiling with Intel compiler has to be done on the login node using `icc` or `mpiicc` commands
- d. To generate a job script
  - You can use the following tool <https://courses.cs.unibas.ch/cluster/index.html>
  - To run on a separate node and avoid interference from other jobs, check the exclusive box.
  - Copy the generated script into a new file `jobscriptname.sh` on miniHPC (see step a above)
- e. To submit a job
  - `sbatch jobscriptname.sh`
  - The system will assign an ID to your submitted job
- f. To check job status
  - `squeue`
  - possible statuses *R* for running, *PD* for pending, *CA* for canceled, and *CG* for completed.
- g. To delete a job
  - `scancel Job_ID`
- h. To check the output of your submitted job
  - Upon completion of your job, you will find a new file, `slurm-Job_ID.out` that contains the output and any errors