CSC

ICT Solutions for Brilliant Minds

Webinar: From Taito to Puhti

Jussi Enkovaara

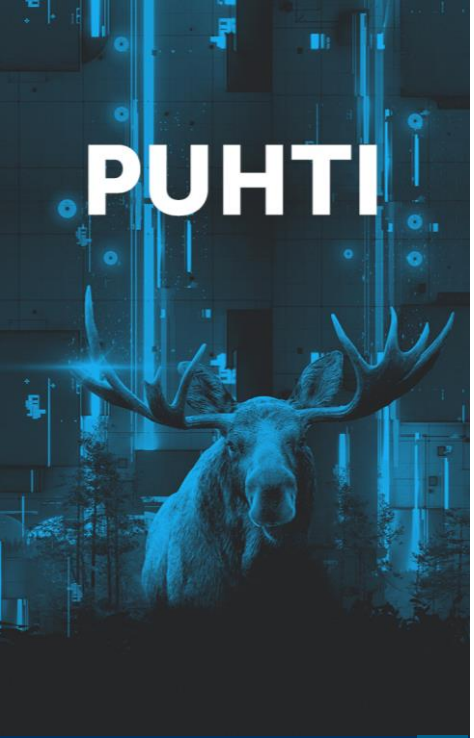
Contents

- Overview of Puhti
- Getting access to Puhti
- Disk areas in Puhti
- Module system and running jobs
- Building applications yourself

Taito general availability has ended

Limited support
Full decommissioning latest in March 2020

User documentation: docs.csc.fi

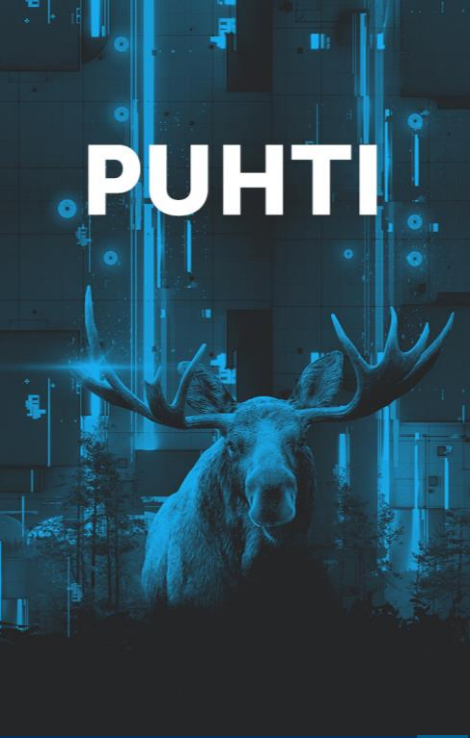


Puhti - computing cluster

Puhti has in total three partitions with 1002 compute nodes

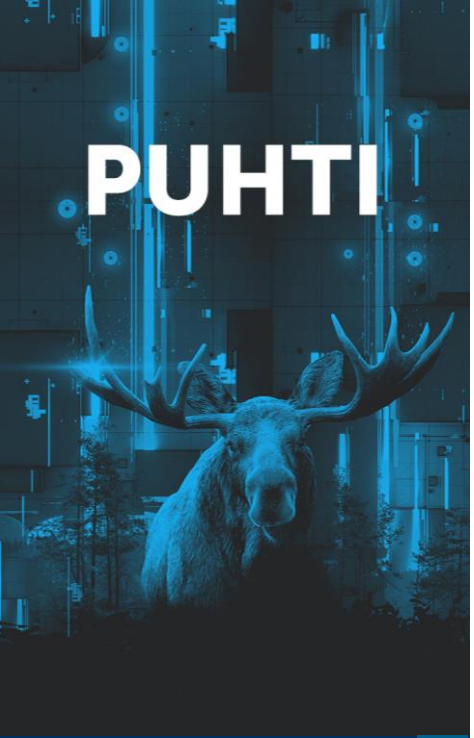
1. CPU partition (Puhti)
   - 682 CPU nodes with 40 cores per node, in total 27280 cores
   - Floating point performance is 1.8 Petaflops
2. GPU partition (Puhti – AI)
   - 80 nodes with 40 CPU cores and 4 GPUs per node, in total 3200 CPU cores and 320 GPUs
   - Floating point performance is 2.7 Petaflops
3. FMI partition (Puhti-FMI)
   - 240 CPU nodes, in total 9600 CPU cores
   - Owned by FMI and operated by CSC - Not available for normal CSC users

- 4.8 Petabytes work disk for data under active use



Puhti – latest generation technologies

- All nodes equipped with latest generation Intel Xeon Scalable processors
  - 20 cores per CPU, 40 cores per node running at 2.1 GHz
  - Supports AVX-512 instructions for vectorized computations - 2x speedup (theoretical) compared to Haswell CPUs in Sisu and Taito per core
  - Supports VNNI instructions for AI *inference* workloads – speedup up to 10x
- GPU nodes equipped NVIDIA V100 (Volta) GPUs
  - 32 GB of memory per GPU – 2x more than in current Pascal GPUs
- Infiniband HDR interconnect between nodes
  - 100 GB/s bandwidth in CPU nodes
  - 200 GB/s bandwidth in GPU nodes



Technical specifications for nodes

	CPU	CPU cores	Memory	Number of nodes
Puhti CPU partition	Xeon Gold 6230	2 x 20 cores @ 2.1 GHz	192 GB	532
	Xeon Gold 6230	2 x 20 cores @ 2.1 GHz	384 GB	92
	Xeon Gold 6230	2 x 20 cores @ 2.1 GHz	384 GB + 3.2 TB NVMe	40
	Xeon Gold 6230	2 x 20 cores @ 2.1 GHz	768 GB	12
	Xeon Gold 6230	2 x 20 cores @ 2.1 GHz	1.5TB	6
Puhti-AI GPU partition	Xeon Gold 6230 4 x V100 32 GB	2 x 20 cores @ 2.1 GHz	384 GB (Host) 128 GB (GPUs) 3.2 TB NVMe	80



Getting access to Puhti

- All users need to apply for new services via new CSC customer portal **my.csc.fi**
- Project manager of CSC project need to apply
  - Project participants need to accept terms and conditions
- Connect with ssh
  - ssh <csc_username>@puhti.csc.fi



Storage in new infrastructure

- HOME: **User specific** directory for small data.
- PROJAPPL: **Project specific** directory, for example for sharing projects own application codes.
- SCRATCH: **Project specific** area for temporary data, i.e. intermittent simulation results.
  - Similar to WRKDIR in current systems (however, WRKDIR was user specific)
  - **Requested quota consumes billing units**
  - **Automatic cleaning**: Files will be deleted 90 days from last access, relevant data should be moved to **Allas**
- PROJAPPL, SCRATCH: By default, all files and directories are accessible to all project members



Storage in new infrastructure

- Default quotas can be found in docs.csc.fi (Computing -> Disk areas)
  - **Note!** Also maximum number of files is limited
- SCRATCH directories are of the form: /scratch/<project>
- PROJAPPL: /projappl/<project>
- Project names and other information can be found at my.csc.fi
- **`csc-workspaces`** –command can be used for listing available directories in Puhti
  - In future, also project names are made available in Puhti



Running preinstalled applications

- Scientific software offering by CSC remains mostly the same
- Similar **module** system as previously
  - i.e. **module load biokit , module spider gromacs,** etc.
- Optimum runtime parameters (number of CPU cores etc.) most likely different than in Taito/Sisu
- Similar SLURM batch job queuing system as previously
  - New queues and policies (number and type of nodes, running times)
  - Recommended to write new batch job scripts starting from the templates in docs.csc.fi (Computing -> Running jobs)



SLURM configuration

- Obligatory:
  - #SBATCH --account=project_XXXXXX
- When communication is not critical:
  - #SBATCH --ntasks=120
- For minimum spread, optimal communication:
  - #SBATCH --nodes=3
    #SBATCH --ntasks-per-node=40
- Ask for the memory you need
  - Either cores or memory can run out in a node
  - In partition hugemem, placement is based on requested amount
  - Check resource usage with seff <SLURM_JOBID>



Billing is by used *or* requested resources

- Additive billing of used/requested resources
  - Core hours (used) 1h = 1 BU
  - GPU hours (used) 1h = 60 BU
  - Memory (requested) 1 GiBh = 0,1 BU
  - NMVe (requested) GiBh = 0,006 BU

    → Total BUs = (NCores * 1 + MemGiBs * 0.1 + NVMeGiBs * 0.006 + NGPUs * 60) * Walltime hours
  - Lustre (SCRATCH, PROJAPPL) quota 1 TiB = 50000 / year
    - Note, first TiB is for free

GPU or CPU?
Compare BU cost.

- Try with our calculator:
  research.csc.fi/billing-and-monitoring



Interactive jobs

- Do not run long lasting or memory intensive pre/post processing on login nodes!
  - Few minutes, few gigabyte runs are ok
- Interactive jobs can be run in compute nodes via batch system
  - See docs.csc.fi (Computing -> Running jobs -> Example job scripts)
- For GUI applications NoMachine is recommended
  - Tutorial about NoMachine in docs.csc.fi (Support -> Tutorials)

New: NVMe Fast local disk

- 40 CPU nodes + 80 GPU nodes with 3.2 TiB NVMe (SSD)
- Request 2000 GB with
  - #SBATCH --gres:nvme=2000
- More details in docs.csc.fi (Computing -> Running jobs -> Creating a batch job script)
- Typical use cases:
  - Turbomole
  - Orca



Installing applications by yourself

- New software stack
  - GNU and Intel compilers
  - Various high-performance libraries
  - HPC-X (OpenMPI based) and MPICH MPI libraries
- Applications should be rebuilt
  - Configure scripts, Makefiles etc. may need modifications
- Recommended compiler flags etc. in docs.csc.fi (Computing -> Compiling)
- Applications should be installed in PROJAPPL disk area
  - Easier sharing for the whole project
- Ask for help in servicedesk@csc.fi



Questions?

- Up-to-date information about timetables, relevant changes for users etc. : research.csc.fi/dl2021-utilization
- CSC Customer portal: my.csc.fi
- User documention: docs.csc.fi