VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER

ADOPTING SLURM

TRANSITIONING FROM PBS SCHEDULER

Ondřej Meca

Senior Innovation Technology Developer

ADOPTING SLURM

git clone git@code.it4i.cz:infralab/presentations/slurm23.git

/scratch/project/dd-23-116/slurm

ADOPTING SLURM

simple applications:

- src/work.cpp
  - sequential application (duration 100sec)
- src/parallel.cpp
  - hybrid application (duration 40sec)
- src/mapping.cpp
  - hybrid application (just print mapping of procs/threads to cores)

build: sh make.sh

BASIC WORK WITH SLURM

Two modes:

- Interactive
  - `salloc -A DD-23-116 –p skoleni-Adopting-Slurm-CPU`
- Batch
  - `sbatch script.slurm`

BASIC WORK WITH SLURM

Basic parameters

- `-N, --nodes`
- `-n, --ntasks`
  - total number of processes in the job
- `-c, --cpus-per-task`
  - number of threads per process
- `-t, --time`
  - Acceptable time formats include "minutes", "minutes:seconds", "hours:minutes:seconds", "days-hours", "days-hours:minutes" and "days-hours:minutes:seconds"
- `-D, --chdir`
- `-A, --account`
- `-J, --job-name`

BASIC WORK WITH SLURM

hostname.slurm

- set account, partition, name
- print hostname
- output to slurm-JOBID.out

- try to run it with different parameters (N, n)

BASIC WORK WITH SLURM

parameters.slurm

- print environment parameters of the job
- output file with many parameters available in Slurm

BASIC WORK WITH SLURM

work.sequential.slurm, work.sequential.long.slurm

- simple sequential work (2x)
- check job info
- how many steps in the job?
- how to kill the job/step?

PBS TO SLURM

job.pbs, job.slurm

- check how PBS can be set in Slurm

- where are the jobs executed?
- how to set OMP_NUM_THREADS?

JOB SETTING WITH SLURM

work.parallel.slurm

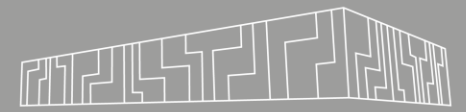
- simple parallel computation
- how to set number of MPI procs?
- how to set number of threads?
- check the job by htop

JOB SETTING WITH SLURM

work.multiple.slurm

- several sequential computation on the same node
- check the syntax
- check the htop
- why parameter –c is used?
- check how steps are executed by sacct

JOB SETTING WITH SLURM



work.array.slurm

- array job
- several independent jobs spawned by Slurm
- the behavior is the same as creating several separated jobs
  - each job gets it own resources (compute node)

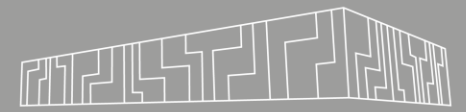
- check JOB_ID, ARRAY_JOB_ID, ARRAY_TASK_ID

JOB SETTING WITH SLURM

work.dependent.slurm

- job dependency
- run the script and check execution by sacct and squeue

JOB SETTING WITH SLURM



work.partial.slurm

- ask for a particular number of GPUs (only on Karolina)

- docs.it4i.cz/general/karolina-slurm

JOB MAPPING WITH SLURM

mapping.percore.slurm

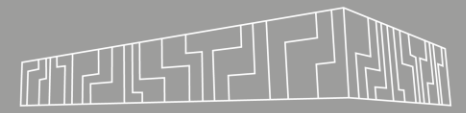
- default mapping (with –N and –n parameters)
- each process is pinned to single core
- can be used for compute bound applications

JOB MAPPING WITH SLURM

mapping.pernode.slurm

- without –n parameter
- cores of whole node is available
- be aware of NUMA
- set the number of threads (default is the number of cores)

JOB MAPPING WITH SLURM

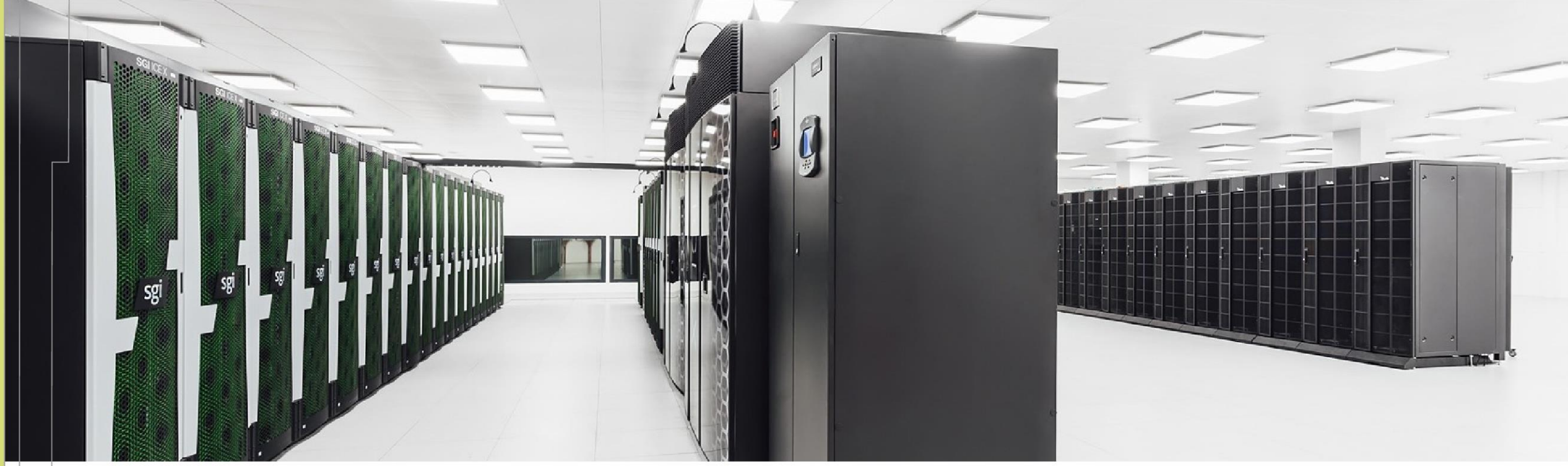


mapping.persocket.slurm

- use –c, --cpus-per-task parameter
- check the number of cores per socket

mapping.persocket.cyclic.slurm

- different mapping of MPI processes



Ondřej Meca
ondrej.meca@vsb.cz

IT4Innovations National Supercomputing Center
VSB – Technical University of Ostrava
Studentská 6231/1B
708 00 Ostrava-Poruba, Czech Republic
www.it4i.cz

VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER