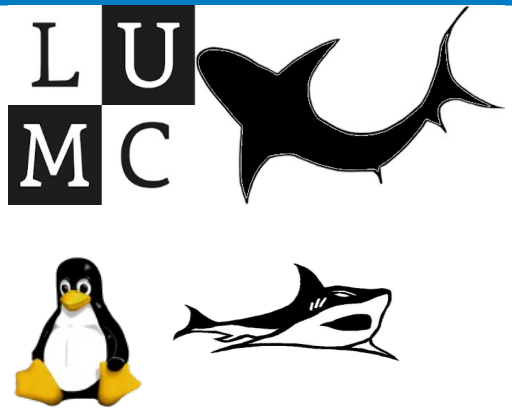
Leiden University
Medical Center

Shark SLURM cluster introduction

IT&DI / Infrastructuurdiensten / Team Linux

Pieter van Vliet
Tom Brusche
John Berbee
Michel Villerius



Overview

1. What is a compute cluster (Shark)
2. Shark cluster overview
3. Howto connect to the Shark cluster
4. Copying files to/from the Shark cluster
5. Linux commands
6. SLURM workload manager
7. Module environment & SLURM
8. SLURM commands
9. SLURM batch script examples

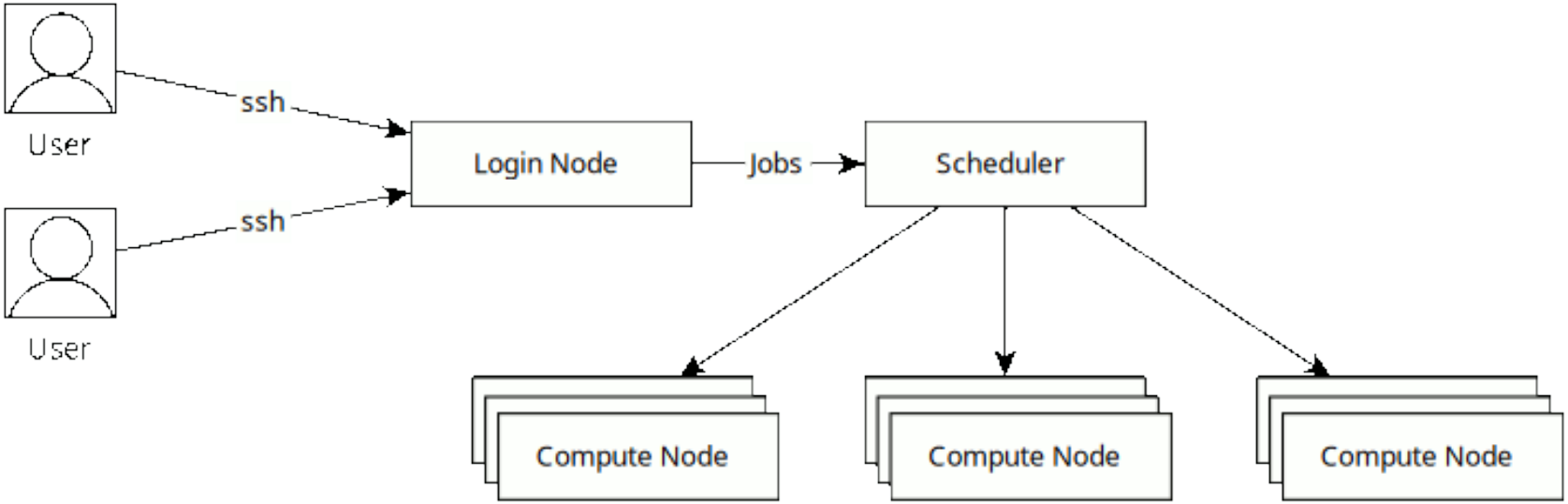
What is a compute cluster (Shark)

A “HPC” cluster is a group of servers called nodes
All nodes are connected through a fast network
There are different type of nodes with different tasks

- Head node, backup node, database node
- Login node
- Execution/compute node, high memory node , GPU node
- Storage node

All nodes have the same disks/exports mounted
A cluster is controlled by a workload manager
Shark cluster uses SLURM as a workload manager / scheduler

Shark cluster overview



Howto connect to the Shark cluster

- The default connection method for all clusters is ssh
- WINDOWS ssh clients
  - Putty (ssh)
  - Mobaxterm (ssh, sftp + graphical X11 env)
- Linux / Mac
  - ssh (+ X11 forwarding) command line

- GUI connections for all operating systems
  - X2go client
  - vncviewer (client)

Copying files to/from the Shark cluster

- The default connection method for all clusters is sftp
- WINDOWS sftp clients
  - psftp (putty sftp)
  - Mobaxterm (with sftp)
- Linux / Mac
  - sftp command line

- GUI connections for all operating systems
  - FileZilla

Linux commands

- cd
  - With absolute path cd /home/<your_username>
  - Relative path cd ~, cd ..

- ls -l -la -lh -ltr -ltrh [list]
- cp [copy] , rm [remove], mv [move/rename]
- pwd [print current/working directory]
- mkdir [make directory] , rmdir [remove directory]

Linux commands view/convert/redirect text

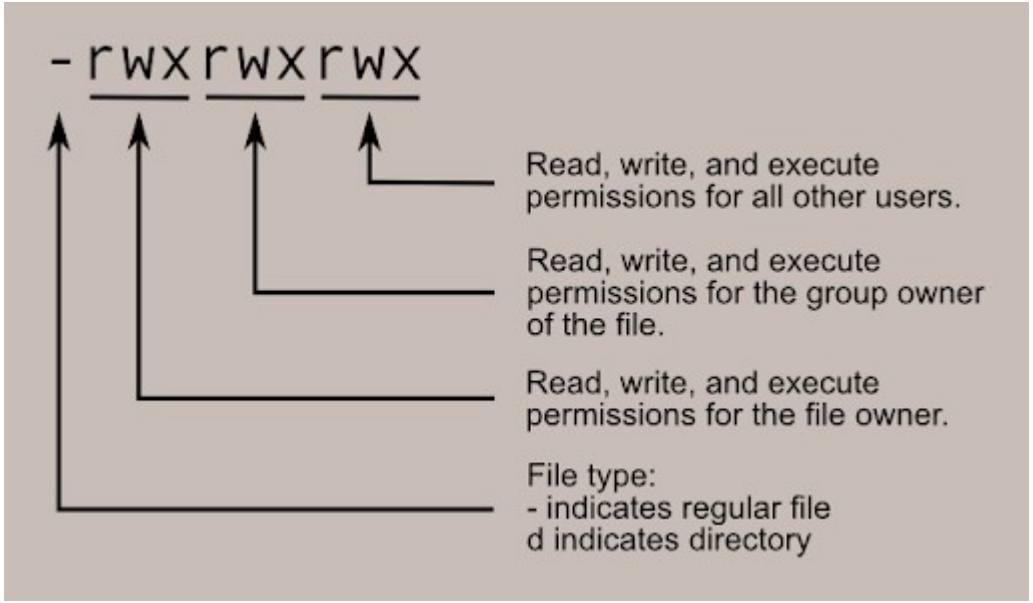
- cat [concatenate], cat -v
- more, less
- head, tail
- dos2unix, unix2dos

Redirect

- > [overwrite]
- >> [add/append]
- | [pipe]

Linux permissions and owners

- chmod [change file mode bits]
- chown [change file owner and group]



Linux commands edit/manual

- vi, vim
- emacs
- Nano
- echo [display a line of text]
- touch [creates an empty file]

- man [an interface to the on-line reference manuals]
  - man man
  - man ls

SLURM workload manager

- SLURM is an open source cluster management and job scheduling system
- SLURM allocates access to resources (compute nodes)
- SLURM provides a framework for starting, executing and monitoring jobs
- SLURM resources
  - Runtime (--time=)
  - Number of CPU’s
    - --ntasks= Number of "tasks" (MPI workers)
    - --ntasks-per-node= Number of "tasks" per node (use with distributed parallelism)
    - --cpus-per-task= Number of CPUs for each task, use this for threads/cores in single-node jobs.
  - Total memory (--mem=) or (--mem-per-cpu=)
  - GPU (--tres=gpu:1)

Modules environment & SLURM

- **Module environment**
  - module avail
  - module list
  - module add / load
  - module del / unload
  - module purge

- **module whatis**
  - module whatis library/boost/1.72.0/gcc-8.3.1
  - module spider eyes

SLURM commands

sinfo	View information about Slurm nodes and partitions
squeue	View information about jobs located in the Slurm scheduling queue

- sinfo
- sinfo -a
- sinfo -l
- sinfo -N -l
- sinfo -N -l -a

- squeue
- squeue -a
- squeue -l
- squeue -u <username>

SLURM commands

salloc	Obtain a Slurm job allocation (a set of nodes), execute a command, and then release the allocation when the command is finished
sbatch	Submit a batch script to Slurm
srun	Run parallel jobs
scancel	Used to signal jobs or job steps that are under the control of Slurm

- salloc –-nodes=1 (defaults to NumNodes=1 NumCPUs=1 NumTasks=1 CPUs/Task=1)
- srun --pty bash
- srun –-ntasks=1 –-pty –-x11 xclock

SLURM commands

scontrol	Used view and modify Slurm configuration and state
sacct	Displays accounting data for all jobs and job steps in the Slurm job accounting log or Slurm database
sstat	Display various status information of a running job/step

- sstat -j <jobID>
- sacct -j jobid --format=User,Jobname,partition,time,start,end,elapsed,MaxRss,MaxVMSize,ncpus,nodelist
- scontrol show job <jobID>
- Scontrol show node <node name>
- Scontrol show partition <partition name>

SLURM commands

-N	--nodes=<minnodes[-maxnodes]>	Request this many nodes on the cluster. Use 1 core on each node by default
-n	--ntasks=<number>	The allocation will launch a maximum of number tasks and to provide for sufficient resources the default is one task per node
	--ntasks-per-node=<ntasks>	Request that ntasks be invoked on each node
-c	--cpus-per-task=<ncpus>	Ensuring job steps will require ncpus number of processors per task

SLURM batch script examples

- #/usr/bin/cat ~/hello.sh

```
#!/bin/bash
#

echo "Hello from $(hostname)"
echo "It is currently $(date)"
echo ""
echo "SLURM_JOB_NAME: $SLURM_JOB_NAME"
echo "SLURM_JOBID: " $SLURM_JOBID
```

SLURM batch script examples

- #/usr/bin/cat ~/hello2.sh

```
#!/bin/bash
#
#SBATCH --partition short
#SBATCH –-nodes=1 --ntasks=1 --cpus-per-task=1
# SBATCH --time=0:5:0
#SBATCH --output=hello-%j.out

echo "Hello from $(hostname)"
echo "It is currently $(date)"
echo ""
echo "SLURM_JOB_NAME: $SLURM_JOB_NAME"
echo "SLURM_JOBID: " $SLURM_JOBID
sleep 90
echo "Goodbye!"
date
```

References

- SLURM https://slurm.schedmd.com/documentation.html
- Wiki https://git.lumc.nl/shark/shark-centos-slurm-user-guide/-/wikis/home
-

Leiden University
Medical Center