SLURM Administration and Installation

Microway Training for San Diego State Univ. 4/30/19

What is SLURM?

Stands for “Simple Linux User Resource Manager”

Allocates compute resources to users for some duration of time so they can perform work

Provides a framework for starting, executing, and monitoring work (normally a parallel job)

Arbitrates contention for resources via a queue of pending work (with optional prioritization and QOS)

SLURM Terminology

Nodes - compute resources

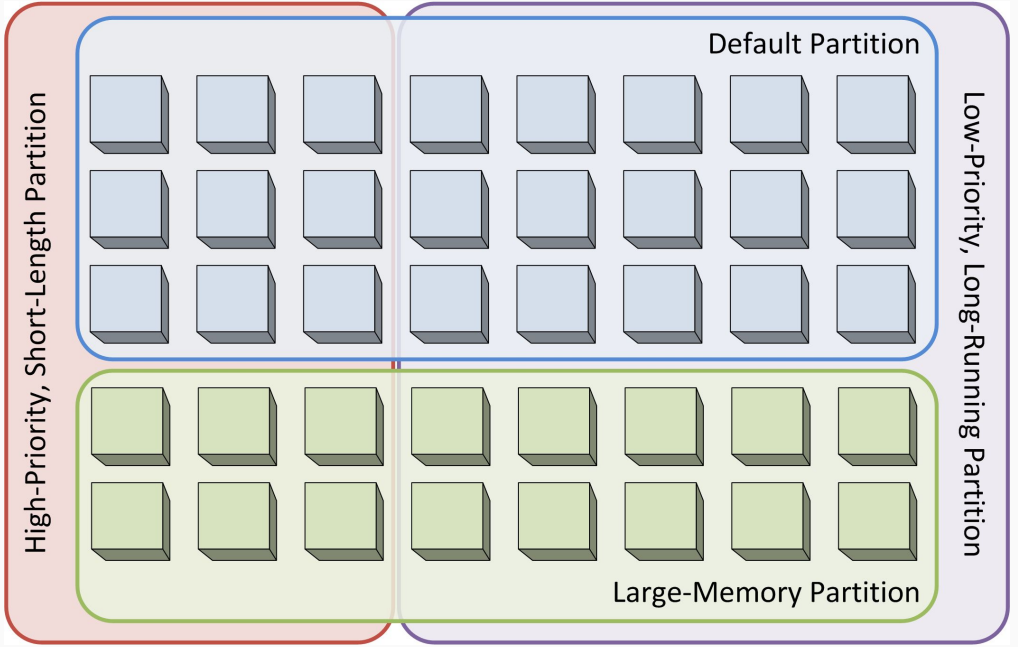
Jobs - time limited resource allocations

Job steps - any subset of the nodes and resources in a job may be used by the jobs steps

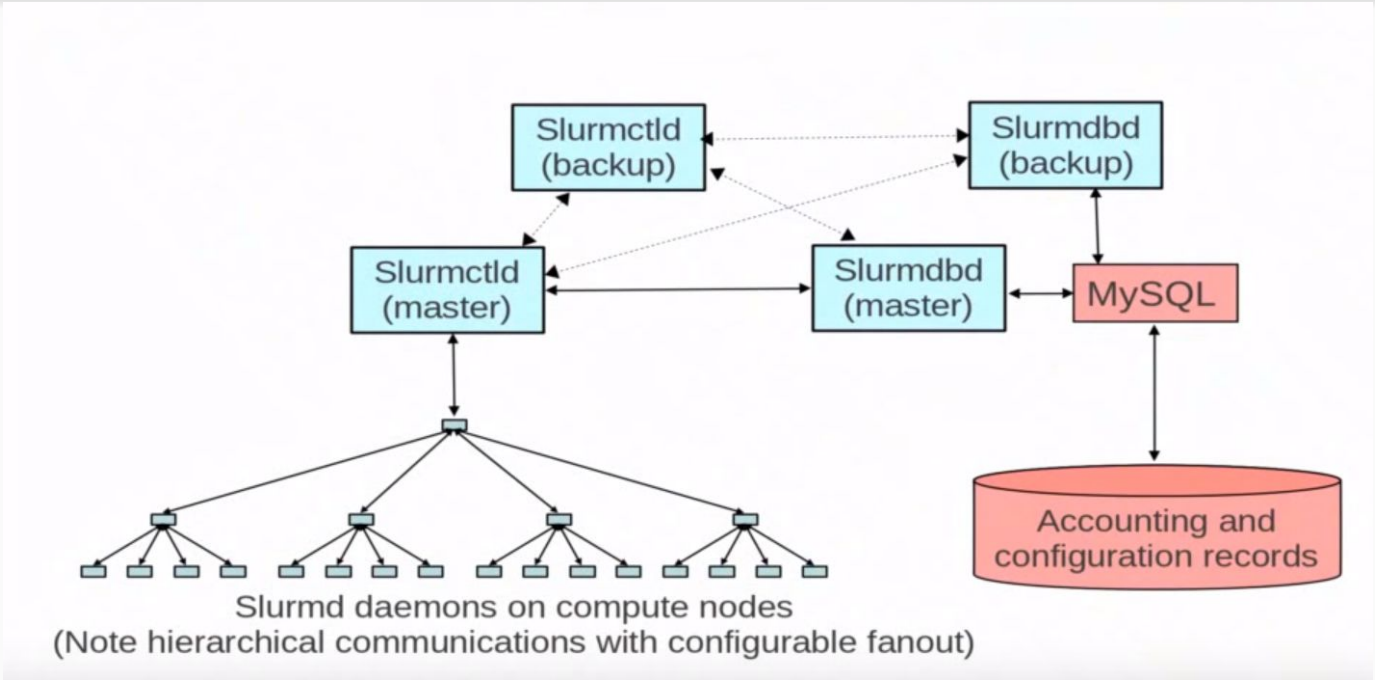
Partition - a logical group of nodes which have common features or attributes

Partitions

- Partitions can be thought of as "job queues"
- Partitions group nodes into logical sets
  - A node may be a member of multiple partitions
  - Constraints may be set for each partition, e.g.:
    max # of nodes, max runtime, memory limits, permissions for specific user groups



SLURM Cluster Architecture



Daemons

- **slurmctld** – Central controller (typically one per cluster)
  - Optional backup with automatic fail over
  - Monitors state of resources
  - Manages job queues
  - Allocates resources
- **slurmd** – Compute node daemon (typically one per compute node)
  - Launches and manages tasks
  - Small and very light-weight (low memory and CPU use)
  - Quiescent after launch except for optional accounting
  - Supports hierarchical communications with configurable fanout
- **slurmdbd** – database daemon (typically one per enterprise)
  - Collects accounting information
  - Uploads configuration information (limits, fair-share, etc.)
  - Optional backup with automatic fail over

Debugging Daemon activity

- -c Clear previous state, purge all job, step, partition state
- -D Run in the foreground, logs are written to stdout
- -v Verbose error messages, each “v” roughly doubles volume of messages

Typical debug mode command lines

```
> slurmctld -Dcvvvv
> slurmd -Dcvvv
```

Slurm configuration file - slurm.conf

/etc/slurm/slurm.conf - contains information required for slurm to run.

- Details about the cluster (cluster name, partitions, nodes, resources)
- Can specify Timers, Priorities,
- Can enable logging and accounting, interaction with database

An identical copy of slurm.conf must be present on all nodes (head and compute)

After making changes to slurm.conf on the head node, remember to copy it to all of the compute nodes and restart the daemons

Use slurmd to generate slurm.conf info

- Execute *slurmd* with *-C* option to print the node's current configuration and exit
- This information can be used as input to the SLURM configuration file

```
> slurmd -C
NodeName=jette CPUs=6 Sockets=1 CoresPerSocket=6 ThreadsPerCore=1
     RealMemory=8000 TmpDisk=930837
```

Installing SLURM: Building on a CentOS 7.x cluster

- Download the latest stable release from https://www.schedmd.com/downloads.php
- Refer to installation document (too many steps to list here)

We will need to use to following MCMS commands to perform cluster-wide commands:

```
-scom, scom-parallel scom-nodes, scom-nodes-parallel

-scpf, scpf-nodes
```

Installing Slurm on an Ubuntu workstation

- Slurm now has support for installing via the regular Ubuntu package manager (`apt-get install slurm`)
- This process is mostly straightforward and easier than building from source, but there are some details prerequisite steps which will be covered in the next webinar.

Restrict access to Partitions by user group

- If you have different logical user groups (e.g. Math, Physics, CS departments) that you would to exclusively allocate resources for, this can be done easily
- Use the `AllowGroups=<group>` parameter in the Partition definition
- For example, in `/etc/slurm/slurm.conf`:

```
PartitionName=PHYSICS      Priority=10000 MaxTime=3000:00 State=UP Nodes=node2
AllowGroups=physics
```

Additional resources/references

- "Rosetta Stone" of Resource Managers

- SLURM Tutorials
- SLURM man pages
- Full list of SLURM Documentation