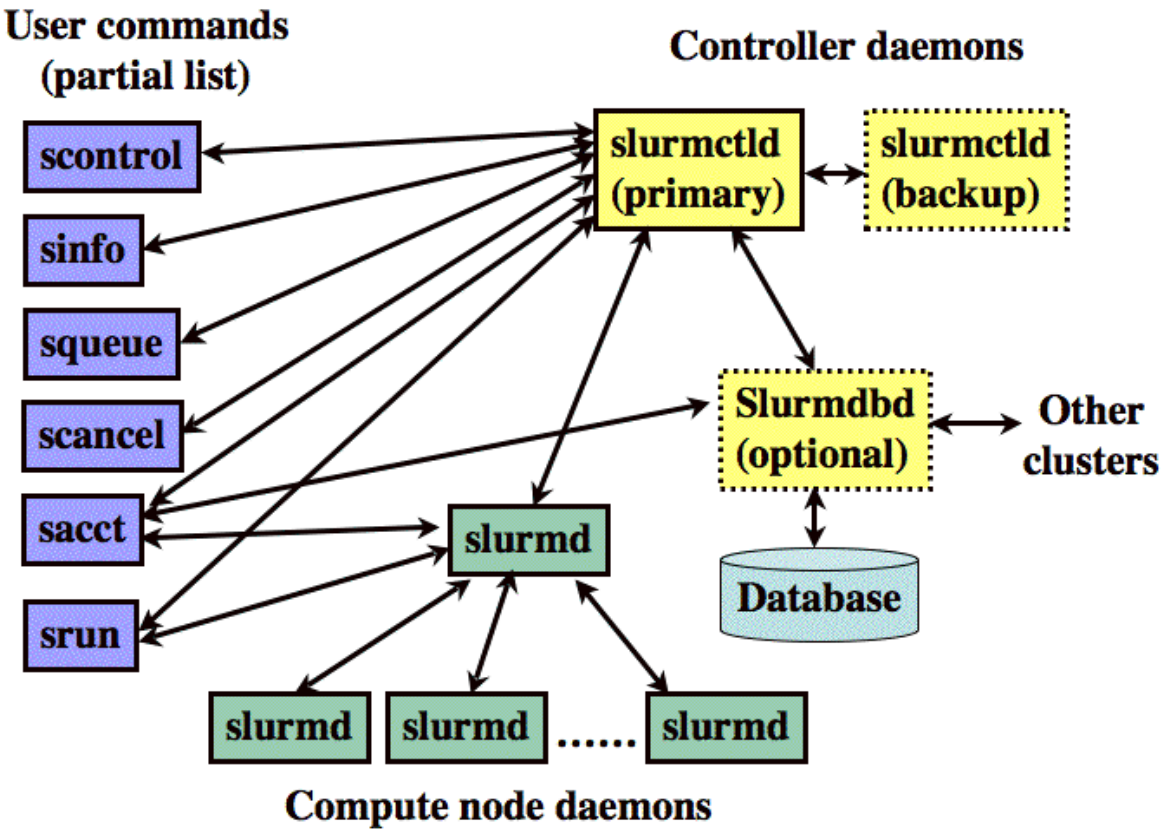
Slurm Installation



Slurmctld-slurm controller

- Node heat
- Scheduling
- Node demon communication
- DBD

_____ SYSTEM REQUIREMENT _____

 # create 3 vm of centos image

Server Type	Server Hostname	Specs
Master	slurm_master	2GB Ram, 2vcpus
Worker	slurm_node1	2GB Ram, 1vcpus
Worker	slurm_node2	2GB Ram, 1vcpus

 # edit host file and write all ip-address of all machine on all system

vi /etc/hosts

```
127.0.0.1   localhost localhost.localdomain localhost4 localhost4.localdomain4
::1         localhost localhost.localdomain localhost6 localhost6.localdomain6
192.168.174.166 master
192.168.174.167 node1
192.168.174.168 node2

~
```

 # make sure all 3 pc ping to each other

 # create keygen on master node

Ssh-keygen

 # copy keygen to all client node

Ssh-copy-id

 # Update all system

yum update

 # disable selinux on all system

setenforce 0

'or'

vi /etc/selinux/config

SELINUX=disabled

```
# This file controls the state of SELinux on the system.
# SELINUX= can take one of these three values:
#     enforcing - SELinux security policy is enforced.
#     permissive - SELinux prints warnings instead of enforcing.
#     disabled - No SELinux policy is loaded.
SELINUX=disabled
# SELINUXTYPE= can take one of three values:
#     targeted - Targeted processes are protected,
#     minimum - Modification of targeted policy. Only selected processes are protected.
#     mls - Multi Level Security protection.
SELINUXTYPE=targeted
```

 # firewall disable on all system
systemctl disable firewalld
systemctl stop firewalld

 # restat all system on all machine
init 6

```
on master:-
----------------
yum install nfs-utils                          #install nfs-utils
yum install nfs4* -y                           #install nfs-utils version 4
rpm -qa | grep nfs                             #show nfs file which u
install
mkdir shital                                   #create a directory
cd shital
      touch file1 file2                        #create file in directory
systemctl start nfs                            #start nfs service
systemctl stop firewalld                       #stop firewall service
vi /etc/exports                                #create a file
      /root/shital *(rw,sync,no_root_squash)   #write this in file as same
systemctl restart nfs                          #restart nfs service
exportfs -arv                                  #show file export or not

on client:-
-----------
yum install nfs-utils                          #install nfs-utils
yum install nfs4*                              #install nfs-utils version 4
rpm -qa | grep nfs                             #show nfs file which u
install
mkdir surya
cd shital
      touch file3 file4
systemctl start nfs
systemctl stop firewalld
mount -t nfs master:/root/shital /root/surya  #mount client and master
directory


 ls
file1 file2 file3 file4                              #show on both system
```

Install chrony(for set time)

——————————

On server—

 # install chrony
yum install chrony* -y

vi /etc/chrony.conf
26(write server-ip network) & 29→ uncomment
Ex-192.168.174.0/24
    local stratum 10

```
# Use public servers from the pool.ntp.org project.
# Please consider joining the pool (http://www.pool.ntp.org/join.html).
server 192.168.100.129 iburst
#server 1.centos.pool.ntp.org iburst
#server 2.centos.pool.ntp.org iburst
#server 3.centos.pool.ntp.org iburst

# Record the rate at which the system clock gains/losses time.
driftfile /var/lib/chrony/drift

# Allow the system clock to be stepped in the first three updates
# if its offset is larger than 1 second.
makestep 1.0 3

# Enable kernel synchronization of the real-time clock (RTC).
rtcsync

# Enable hardware timestamping on all interfaces that support it.
#hwtimestamp *

# Increase the minimum number of selectable sources required to adjust
# the system clock.
#minsources 2

# Allow NTP client access from local network.
allow 192.168.100.0/24

# Serve time even if not synchronized to a time source.
local stratum 10

# Specify file containing keys for NTP authentication.
#keyfile /etc/chrony.keys

# Specify directory for log files.
logdir /var/log/chrony

# Select which information is logged.
#log measurements statistics tracking
~
~
"/etc/chrony.conf" 38L, 1105C
```

:wq

 # copy the file to all node

[root@master]# rsync /etc/chrony.conf node1:/etc/chrony.conf

[root@master]# rsync /etc/chrony.conf node2:/etc/chrony.conf

 # start chrony service
systemctl start chronyd

On client:-
——————————
 # install chrony
yum install chrony* -y

 # start chrony service
systemctl start chronyd

 # for request to server time querry
ntpdate -q ‘server-ip’ -

 # for request to server time update
ntpdate -u ‘server_ip’ -

 # check all machine time status
chronyc sources

```
[root@node1 ~]# ntpdate -q 192.168.174.174
server 192.168.174.174, stratum 10, offset 0.540681, delay 0.02629
16 Jan 18:35:00 ntpdate[10759]: step time server 192.168.174.174 offset 0.540681 sec
[root@node1 ~]# ntpdate -u 192.168.174.174
16 Jan 18:35:29 ntpdate[10760]: step time server 192.168.174.174 offset 0.540677 sec
[root@node1 ~]# chronyc sources
210 Number of sources = 4
MS Name/IP address         Stratum Poll Reach LastRx Last sample
===============================================================================
^? 64.227.167.110                0  10     0     -     +0ns[   +0ns] +/-    0ns
^? 162.159.200.123               0  10     0     -     +0ns[   +0ns] +/-    0ns
^? 178.215.228.24                0  10     0     -     +0ns[   +0ns] +/-    0ns
^? 192.46.215.60                 0  10     0     -     +0ns[   +0ns] +/-    0ns
[root@node1 ~]#
```

 # download slum tar file from browser
https://download.schedmd.com/slurm/slurm-22.05.7.tar.bz2

 # install mariaDB on all machine
yum install mariadb-server mariadb-devel -y

 # epel-release on all machine
yum install epel-release -y

 # install munge on all machine
yum install munge munge-libs munge-devel -y

 # rpm-build on all machine
yum install rpm-build -y

 # rpm build of slurm file on master
rpmbuild -ta slurm-22.05.7.tar.bz2

 # install perl-extutils on master
yum install perl-ExtUtils-Install* -y
yum install python3 readline-devel pam-devel gcc -y

 # Generate munge key on master
/usr/sbin/create-munge-key -r

 # change permission of munge.key on master
chmod 400 /etc/munge/munge.key

 # send key to all node
scp munge.key root@node1:/etc/munge/
scp munge.key root@node2:/etc/munge/

 # change owner on all node
chown -R munge: /etc/munge/munge.key

 # start and enable munge service
systemctl enable munge
systemctl start munge
systemctl status munge

 # Go to this path
cd /root/rpmbuild/RPMS//x86-64

 # Install slurm on all node
yum install slurm* -y

 # copy all rpm file to all node
scp -r /root/rpmbuild/RPMS/x86_64/ root@node1:/home/
scp -r /root/rpmbuild/RPMS/x86_64/ root@node2:/home/

 # Export slumuser on all node
export SLURMUSER=992

groupadd -g $SLURMUSER slurm

useradd -m -c "SLURM workload manager" -d /var/lib/slurm -u $SLURMUSER -g slurm -s /bin/bash slurm

 # change file name on master
cp /etc/slurm/slurm.conf.example /etc/slurm/slurm.conf

 # edit file on master (line no 11&12&52&57)
 vi /etc/slurm/slurm.conf
        clusterName=surya   #controlmachine=surya
        SlurmctldHost=master   #controlAddr=master
        SlurmdSpoolDir=/var/share/slurm
        StateSaveLocation=/var/share/slurm/ctld

 # create directory on master

mkdir -p /var/share/slurm/ctld

 # create file on master

touch /var/log/slurmctld.log

Give permission to ctld file

chown -R slurm:slurm /var/share/slurm

 # on all client

slurmd -C

```
[root@node1 surya1]# slurmd -C
NodeName=node1 CPUs=1 Boards=1 SocketsPerBoard=1 CoresPerSocket=1 ThreadsPerCore=1 RealMemory=1959
UpTime=0-02:40:58
```

copy 1st line

On master:-

vi /etc/slurm/slurm.conf

```
# COMPUTE NODES
#NodeName=linux[1-32] CPUs=1 State=UNKNOWN
NodeName=node2 CPUs=1 Boards=1 SocketsPerBoard=1 CoresPerSocket=1 ThreadsPerCore=1 RealMemory=1959
NodeName=node1 CPUs=1 Boards=1 SocketsPerBoard=1 CoresPerSocket=1 ThreadsPerCore=1 RealMemory=1959
PartitionName=debug Nodes=ALL Default=YES MaxTime=INFINITE State=UP
```

hint:check 2nd last line

copy slurm.conf file on all node

scp /etc/slurm/slurm.conf root@node1:/etc/slurm/slurm.conf

scp /etc/slurm/slurm.conf root@node2:/etc/slurm/slurm.conf

```
[root@master ~]# scp /etc/slurm/slurm.conf root@node1:/etc/slurm/slurm.conf
slurm.conf                                    100% 3229     2.4MB/s   00:00
[root@master ~]# scp /etc/slurm/slurm.conf root@node2:/etc/slurm/slurm.conf
slurm.conf                                    100% 3229     2.6MB/s   00:00
```

change owner of slurm.conf on all node

chown -R slurm:slurm /etc/slurm/slurm.conf

On client:-

 # make directory
mkdir -p /var/share/slurm/d

 # change owner
chown -R slurm:slurm /var/share/slurm/

 # create file
touch /var/log/slurmd.log

 # start munge service
systemctl start munge

 # start slurmd
systemctl start slurmd

 # enable slurmd
systemctl enable slurmd

 # run on server for check their worker nodes active or not
sinfo

```
[root@master ~]# sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
debug*       up   infinite      1  idle* node1
debug*       up   infinite      1  down* node2
```

 # for start forcefully node
scontrol update nodename=’down_node_name’ state=idle

 # provide shell of the node
srun -N1 –pty /bin/sh

 # show node on server
scontrol show node ‘node_name’

 # show partition on server
scontrol show part ‘part_name’

—————X—————X—————X—————X—————

 # Create partition on server
vi /etc/slurm/slurm.conf

PartitionName=debug Nodes=ALL Default=YES
MaxTime=INFINITE State=UP
PartitionName=small Nodes=node1 Default=YES
MaxTime=INFINITE State=UP
PartitionName=medium Nodes=node2 Default=YES
MaxTime=INFINITE State=UP

```
PartitionName=debug Nodes=ALL Default=YES MaxTime=INFINITE State=UP
PartitionName=small Nodes=node1 Default=YES MaxTime=INFINITE State=UP
PartitionName=medium Nodes=node2 Default=YES MaxTime=INFINITE State=UP
```

add line like this in last row

copy slurm.conf file on both nodes

```
[root@master ~]# scp /etc/slurm/slurm.conf node1:/etc/slurm/slurm.conf
slurm.conf                                    100% 3371     3.0MB/s   00:00
[root@master ~]# scp /etc/slurm/slurm.conf node2:/etc/slurm/slurm.conf
Warning: Permanently added the ECDSA host key for IP address '192.168.174.180' to the list
 of known hosts.
slurm.conf                                    100% 3371     2.2MB/s   00:00
[root@master ~]# systemctl restart slurmctld
[root@master ~]# systemctl status slurmctld
```

restart slurmctld on master
systemctl restart slurmctld
systemctl status slurmctld

restart slurmd on nodes
systemctl restart slurmd
systemctl status slurmd

 # run on server for check their worker nodes status
sinfo

```
[root@master ~]# sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
debug        up   infinite      2   idle node[1-2]
small        up   infinite      1   idle node1
medium*      up   infinite      1   idle node2
```

Idle- ready to start for run any job
Up-   node start
Down-node down
Comp-submitted their job