

Módulo 7

Gerenciamento de Clusters (Sistemas Distribuídos)

M7-I -Introdução à configuração e gerenciamento de Clusters

Caderno de atividades

Informações gerais:

Usuários cadastrados nas máquinas virtuais:

Usuários	Senha
root	123456
verao	verao2020
user01	123456
user02	123456

Diretório exportado via NFS para os nós: /scratch

Diretório home dos usuários do cluster: /scratch/\$USER

Comando executado pelo superusuário na máquina login
login# comando parametros

Comando executado pelo superusuário em todas as máquinas (login, no01 e no02)
todos# comando parametros

Comando executado pelo superusuário nos nós computacionais (no01 e no02)
nodes# comando parametros

Comando executado por um usuário comum
login\$ comando parametros

Lab 1. Instalando os pré-requisitos para a instalação do SLURM:

1.1. MUNGE

O munge pode ser instalado a partir do código fonte, disponível em <https://dun.github.io/munge/> ou do repositório epel-release do CentOS. Ele deverá ser instalado em todos os nós do cluster e a chave munge.key deve ser idêntica em todas as máquinas.

Instalação em todos os nós:

```
# clush -w login,no0[1-2] yum install -y munge munge-libs munge-devel
```

Criando a chave

```
login# dd if=/dev/urandom bs=1 count=1024 > /etc/munge/munge.key
login# chown munge:munge /etc/munge/munge.key
login# chmod 0600 /etc/munge/munge.key
```

Outra opção:

```
login# /usr/sbin/create-munge-key -r /dev/random
login# chown munge: /etc/munge/munge.key
login# chmod 0600 /etc/munge/munge.key
```

Copiar a chave para todos os nodes

```
login# clush -w no0[1-2] -c /etc/munge/munge.key
login# clush -w no0[1-2] chown munge:munge /etc/munge/munge.key
```

Iniciar o processo do munge em todo o cluster

```
login# clush -w login,no0[1-2] systemctl enable munge
login# clush -w login,no0[1-2] systemctl start munge
login# clush -w login,no0[1-2] systemctl status munge
```

Testando:

```
login# munge -n | unmunge
login# munge -n | ssh no01 unmunge
login# munge -n | ssh no02 unmunge
```

O resultado deve apresentar o status Success (0)

```
root@login ~]# munge -n | unmunge
STATUS:          Success (0)
ENCODE_HOST:     login.lncc.local (192.168.0.1)
...
```

1.2. Banco de dados para Accounting (MySQL/MariaDB)

Será necessário adicionar o repositório oficial do projeto MariaDB pois a versão disponível no repositório do Centos 7 não possui todas as libs necessárias para a compilação do SLURM.

```
todos# curl -sS https://downloads.mariadb.com/MariaDB/mariadb_repo_setup | bash
```

Instalação do MariaDB no Service Node

```
login# yum install -y MariaDB-client MariaDB-shared MariaDB-devel \
MariaDB-server MariaDB-backup
```

Instalação das libs nos computer nodes

```
nodes# yum install -y MariaDB-client MariaDB-shared MariaDB-devel
```

Ajustar configurações do banco de dados para os valores recomendados

Criar o arquivo /etc/my.cnf.d/innodb.cnf com o conteúdo abaixo:

```
login# vi /etc/my.cnf.d/innodb.cnf
[mysqld]
innodb_buffer_pool_size=1024M
innodb_log_file_size=64M
innodb_lock_wait_timeout=900
```

Iniciar o serviço e configurar a inicialização automática durante o boot

```
login# systemctl start mariadb
login# systemctl enable mariadb
```

Ajustes de segurança no mysql

```
login# mysql_secure_installation
Senha de root: 123456
Switch to unix_socket authentication: Yes
Change the root password? No
Remove anonymous users? Yes
Disallow root login remotely? Yes
Remove test database and access to it? Yes
Reload privilege tables now? Yes
```

Criação do database e usuário do slurm no mysql

Acesse o terminal com o usuário root e conecte-se ao mysql, em seguida, siga os passos abaixo para criar o database e o usuários utilizados pelo slurm.

```
# mysql
mysql> create database slurm_acct_db;
mysql> create user 'slurm'@'localhost' identified by 'slurm_pass';
mysql> grant all on slurm_acct_db.* TO 'slurm'@'localhost';
mysql> FLUSH PRIVILEGES;
mysql> exit;
```

Testando o acesso:

```
login# mysql -u slurm -D slurm_acct_db -p
```

1.3. Dependências do S.O.

Além do MUNGE e do MySQL diversas bibliotecas são necessárias para a compilação e instalação do SLURM. Utilize o gerenciador de pacotes yum para instalar todas as dependências listadas abaixo:

```
todos# yum install -y openssl openssl-devel pam-devel numactl numactl-devel
hwloc hwloc-devel lua lua-devel readline-devel rrdtool-devel ncurses-devel
man2html libibmad libibumad perl-ExtUtils-MakeMaker wget
```

2. Instalação do Slurm

Criar grupo slurm
todos# groupadd -g 991 slurm

Criar usuário slurm
todos# useradd -m -c " SLURM workload manager " -d \
/var/lib/slurm -u 991 -g slurm -s /bin/bash slurm

Baixar o fonte do slurm no diretório /scratch/app:
login# cd /scratch/app
login# wget <https://download.schedmd.com/slurm/slurm-19.05.5.tar.bz2>

Criar os pacotes
login# rpmbuild -ta slurm-19.05.5.tar.bz2
login# mkdir /scratch/app/slurm-rpms
login# cp /root/rpmbuild/RPMS/x86_64/slurm-* /scratch/app/slurm-rpms

2.1. Instalando o Login/Service node

login# cd /scratch/app/slurm-rpms
login# yum --nogpgcheck localinstall -y \
slurm-19.05.5-1.el7.x86_64.rpm \
slurm-perlapi-19.05.5-1.el7.x86_64.rpm \
slurm-slurmctld-19.05.5-1.el7.x86_64.rpm \
slurm-slurmdbd-19.05.5-1.el7.x86_64.rpm \
slurm-example-configs-19.05.5-1.el7.x86_64.rpm \
slurm-devel-19.05.5-1.el7.x86_64.rpm \
slurm-contrib-19.05.5-1.el7.x86_64.rpm

2.2. Instalando os computer nodes

nodes# cd /scratch/app/slurm-rpms
nodes# yum --nogpgcheck localinstall -y
slurm-19.05.5-1.el7.x86_64.rpm \
slurm-perlapi-19.05.5-1.el7.x86_64.rpm \
slurm-slurmd-19.05.5-1.el7.x86_64.rpm \
slurm-pam_slurm-19.05.5-1.el7.x86_64.rpm

2.3. Configurando o slurmdbd

Copiar o exemplo do arquivo de configuração e editá-lo conforme abaixo:

login# cp /etc/slurm/slurmdbd.conf.example /etc/slurm/slurmdbd.conf
login# vi /etc/slurm/slurmdbd.conf

```
AuthType=auth/munge  
DbdAddr=localhost  
DbdHost=localhost  
SlurmUser=slurm  
DebugLevel=verbose  
LogFile=/var/log/slurm/slurmdbd.log  
PidFile=/var/run/slurmdbd.pid  
StorageType=accounting_storage/mysql  
StorageHost=localhost  
StoragePass= slurm_pass  
StorageUser=slurm  
StorageLoc=slurm_acct_db
```

Criar o diretório de log e spool e alterar o dono e o grupo

login# mkdir -p /var/log/slurm/ /var/spool/slurm/d /var/spool/slurm/ctld
login# chown -R slurm:slurm /var/log/slurm /var/spool/slurm

Iniciar o serviço do slurmdbd e verificar os logs

```
login# systemctl start slurmdbd
```

O log do slurmdbd deve apresentar uma saída semelhante a descrita abaixo:

```
[data] Accounting storage MYSQL plugin loaded  
[data] slurmdbd version 19.05.5 started
```

2.4. Configurando o daemon slurmd

Todos os nós do cluster devem ter o mesmo arquivo de configuração do slurmd. Uma sugestão para organizar essa configuração é manter o arquivo slurmd.conf no mesmo nó que roda o daemon do slurmctl – no nosso caso, a máquina login.lncc.local. Toda alteração na configuração do slurm deverá ser feita nesse arquivo e copiada para os demais nós logo em seguida.

Copiar o arquivo de exemplo do slurm e editar conforme abaixo:

```
login# cp /etc/slurm/slurm.conf.example /etc/slurm/slurm.conf  
login# vi /etc/slurm/slurm.conf
```

```
ClusterName=verao20  
ControlMachine=login  
SlurmctldHost=login  
SlurmdUser=slurm  
AuthType=auth/munge  
StateSaveLocation=/var/spool/slurm/ctld  
SlurmdSpoolDir=/var/spool/slurm/d
```

```
# SCHEDULING  
SchedulerType=sched/backfill
```

```
#LOGGING  
SlurmctldDebug=info  
SlurmctldLogFile=/var/log/slurm/slurmctld.log  
SlurmdDebug=info  
SlurmdLogFile=/var/log/slurm/slurmd.log  
JobCompType=jobcomp/none
```

```
# COMPUTE NODES  
NodeName=no0[1-2] CPUs=1 State=UNKNOWN
```

```
PartitionName=teste Nodes=no01 Default=Yes MaxTime=INFINITE State=UP
```

Os demais parâmetros do arquivo podem permanecer com as opções default.

Criar o diretório de spool, de log e alterar as permissões

```
nodes# mkdir -p /var/spool/slurm/d /var/spool/slurm/ctld /var/log/slurm  
nodes# chown -R slurm:slurm /var/spool/slurm /var/log/slurm
```

Copiando as informações para os computer nodes

```
login# clush -w no0[1-2] mkdir /etc/slurm  
login# clush -w no0[1-2] -c /etc/slurm/slurm.conf
```

2.5. Iniciando os daemons, habilitando a inicialização automática e verificando o funcionamento:

```
login# systemctl start slurmdbd
login# systemctl enable slurmdbd
login# systemctl status slurmdbd
```

```
login# systemctl start slurmctld
login# systemctl enable slurmctld
login# systemctl status slurmctld
```

```
nodes# systemctl start slurmd
nodes# systemctl enable slurmd
nodes# systemctl status slurmd
```

2.6. Testando o ambiente

Obtendo as informações do hardware dos nós:

```
nodes# slurmd -C
NodeName=no01 CPUs=1 Boards=1 SocketsPerBoard=1 CoresPerSocket=1
ThreadsPerCore=1 RealMemory=991
```

Verificando o status da partição e dos nodes

```
login# sinfo
PARTITION  AVAIL  TIMELIMIT  NODES  STATE   NODELIST
teste*     up      infinite    2      idle   no[01-02]
```

Teste o funcionamento do cluster com o comando srun

Faça o login com o usuário user01 e submeta o comando hostname para a partição teste solicitando dois nós:

O resultado deverá ser:

```
[user01@login ~]$ srun -N 2 -p teste hostname
no01
no02
```

Lab 3 – Habilitando o sistema de accounting

Editar o arquivo `slurm.conf` e alterar as propriedades `AccountingStorageType` e `AccountingStorageEnforce` para que as informações de accounting sejam armazenadas em banco de dados e que sejam respeitadas as restrições de limites de recursos.

```
login# vi /etc/slurm/slurm.conf
AccountingStorageEnforce=limits
AccountingStorageType=accounting_storage/slurmdbd
```

Copiar o arquivo para todos os nós computacionais e reiniciar os serviços `slurmd` e `slurmctld`

```
login# clush -w no0[1-2] -c /etc/slurm/slurm.conf
login# clush -w no0[1-2] systemctl restart slurmd
login# systemctl restart slurmctld
```

Verifique os logs do daemon `slurmctld`. Notou algum erro? Caso afirmativo, indique uma possível solução.

Lab 4 – Criando associations

4.1. Cadastrando o cluster.

Como pode ser observado no Lab 3, ao habilitarmos o controle de limites de recursos devemos ter, pelo menos, a associação do cluster cadastrada no slurm. Diante disso, adicione o nosso cluster ao banco de dados de accounting (o nome do cluster deve ser idêntico ao valor do parâmetro ClusterName do arquivo slurm.conf). Utilize o comando sacctmgr para isso.

Inicie o serviço do slurmctld e verifique se ainda existe algum problema nos arquivos de log.

Acesse o sistema com o usuário user01 e faça a submissão do comando hostname para a partição teste solicitando 1 nó.

```
user01@login~$ srun -N 2 -p teste hostname
```

Comente o resultado do comando.

4.2. Cadastrando um account.

Cadastre um Account chamado “alunos” sem nenhum tipo de restrição. Em seguida cadastre um Account chamado restrito com a seguinte configuração:

- Cluster: verao20
- Número máximo de nós por job: 1
- Número máximo de jobs submetidos: 2
- Tempo máximo de execução dos jobs: 60 segundos

Exemplo:

```
# sacctmgr create account propriedade=valor propriedade=valor ...
```

4.3. Cadastrando usuários.

Adicione os usuários user01 e user02 ao banco de dados de accounting. Defina o account alunos para ambas as contas e configure o account padrão como restrito apenas para a conta user02.

Repita o teste de submissão de job realizado na atividade 4.1 com os dois usuários e comente o resultado.

4.4. Repita o teste da atividade 4.1 com o usuário user02 e especifique a utilização do account alunos durante a submissão do job.

Lab 5 – Consultando as informações de accounting

5.1. Utilize o comando sacct e liste todos os jobs executados pelo usuário user01.

```
# sacct -u user01
```

5.2. Liste todos os jobs submetido para o account “restrito”.

5.3. Liste os jobs que foram executados nas últimas 2 horas.

Referências:

<https://slurm.schedmd.com/documentation.html>

<https://github.com/dun/munge/blob/master/QUICKSTART>

<https://www.slothparadise.com/how-to-install-slurm-on-centos-7-cluster/>

https://wiki.fysik.dtu.dk/niflheim/Slurm_installation#build-slurm-rpms

<https://slurm.schedmd.com/SC16/SlurmOverview.pdf>