

Cluster admin help

Servicios a reiniciar

Cada vez que se cambia la configuración del cluster han de propagarse los cambios a todos los nodos y reiniciarse los servicios. Para *detritus*,

```
$ systemctl stop slurmd.service
$ systemctl stop slurmctld.service
$ systemctl start slurmctld.service
$ systemctl start slurmd.service
```

y para los nodos

```
$ systemctl stop slurmd.service
$ systemctl start slurmd.service
```

amos q un cambio es asi:

```
[root@detritus ~]# vim /etc/slurm/slurm.conf
[root@detritus ~]# systemctl stop slurmd.service
[root@detritus ~]# systemctl stop slurmctld.service
[root@detritus ~]# for x in $(seq 3);do ssh brick0${x} systemctl stop
slurmd.service; done
[root@detritus ~]# for x in $(seq 3);do scp /etc/slurm/slurm.conf
brick0${x}:/etc/slurm/slurm.conf; done
slurm.conf
100% 1575    1.5KB/s   00:00
slurm.conf
100% 1575    1.5KB/s   00:00
slurm.conf
100% 1575    1.5KB/s   00:00
[root@detritus ~]# systemctl start slurmctld.service
[root@detritus ~]# systemctl start slurmd.service
[root@detritus ~]# for x in $(seq 3);do ssh brick0${x} systemctl start
slurmd.service; done
[root@detritus ~]# sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
devel*    up        infinite   4     idle brick[01-03],detritus
cuda      up        infinite   2     idle brick01,detritus
[root@detritus ~]#
```

Rearmando despues del desastre

Digamos que hay un problema gordo y el cluster queda maso menos asi,

```
[root@detritus /]# sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
devel*    up    infinite   1  down* detritus
devel*    up    infinite   1  drain brick01
devel*    up    infinite   2   idle brick[02-03]
cuda      up    infinite   1  down* detritus
cuda      up    infinite   1  drain brick01
fast      up    infinite   1  drain brick01
fast      up    infinite   2   idle brick[02-03]
```

Vamos a arreglarlo con *scontrol*. Primero lo que esta *down*,

```
[root@detritus /]# scontrol show node detritus
NodeName=detritus CoresPerSocket=1
  CPUAlloc=0 CPUErr=0 CPUTot=32 CPULoad=N/A
  AvailableFeatures=(null)
  ActiveFeatures=(null)
  Gres=gpu:tesla:1
  NodeAddr=172.26.2.33 NodeHostName=detritus Version=(null)
  RealMemory=1 AllocMem=0 FreeMem=N/A Sockets=32 Boards=1
  State=DOWN* ThreadsPerCore=1 TmpDisk=0 Weight=3 Owner=N/A MCS_label=N/A
  BootTime=None SlurmdStartTime=None
  CapWatts=n/a
  CurrentWatts=0 LowestJoules=0 ConsumedJoules=0
  ExtSensorsJoules=n/s ExtSensorsWatts=0 ExtSensorsTemp=n/s
  Reason=Not responding [brecia@2019-12-02T10:41:54]
[root@detritus /]# scontrol update NodeName=detritus State=RESUME
[root@detritus /]# scontrol show node detritus
NodeName=detritus CoresPerSocket=1
  CPUAlloc=0 CPUErr=0 CPUTot=32 CPULoad=N/A
  AvailableFeatures=(null)
  ActiveFeatures=(null)
  Gres=gpu:tesla:1
  NodeAddr=172.26.2.33 NodeHostName=detritus Version=(null)
  RealMemory=1 AllocMem=0 FreeMem=N/A Sockets=32 Boards=1
  State=IDLE* ThreadsPerCore=1 TmpDisk=0 Weight=3 Owner=N/A MCS_label=N/A
  BootTime=None SlurmdStartTime=None
  CapWatts=n/a
  CurrentWatts=0 LowestJoules=0 ConsumedJoules=0
  ExtSensorsJoules=n/s ExtSensorsWatts=0 ExtSensorsTemp=n/s
```

```
[root@detritus /]# sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
devel*    up    infinite   1  idle* detritus
devel*    up    infinite   1  drain brick01
devel*    up    infinite   2   idle brick[02-03]
cuda      up    infinite   1  idle* detritus
cuda      up    infinite   1  drain brick01
fast      up    infinite   1  drain brick01
```

```
fast          up    infinite      2    idle brick[02-03]
```

Ahora lo que esta *drain*, voy a ponerlo *down* primero para matar todos los procesos que haya y despues lo levanto.

```
[root@detritus /]# scontrol show node brick01
NodeName=brick01 Arch=x86_64 CoresPerSocket=1
  CPUAlloc=0 CPUErr=0 CPUTot=64 CPULoad=0.01
  AvailableFeatures=(null)
  ActiveFeatures=(null)
  Gres=gpu:tesla:2
  NodeAddr=172.26.2.41 NodeHostName=brick01 Version=16.05
  OS=Linux RealMemory=1 AllocMem=0 FreeMem=255242 Sockets=64 Boards=1
  State=IDLE+DRAIN ThreadsPerCore=1 TmpDisk=0 Weight=1 Owner=N/A
MCS_label=N/A
  BootTime=2019-11-30T10:18:59 SlurmdStartTime=2019-11-30T10:18:41
  CapWatts=n/a
  CurrentWatts=0 LowestJoules=0 ConsumedJoules=0
  ExtSensorsJoules=n/s ExtSensorsWatts=0 ExtSensorsTemp=n/s
  Reason=Duplicate jobid [breacia@2019-12-02T11:31:47]
[root@detritus /]# scontrol update NodeName=brick01 State=DOWN
Reason="undraining"
[root@detritus /]# scontrol update NodeName=brick01 State=RESUME
[root@detritus /]# scontrol show node brick01
NodeName=brick01 Arch=x86_64 CoresPerSocket=1
  CPUAlloc=0 CPUErr=0 CPUTot=64 CPULoad=0.01
  AvailableFeatures=(null)
  ActiveFeatures=(null)
  Gres=gpu:tesla:2
  NodeAddr=172.26.2.41 NodeHostName=brick01 Version=16.05
  OS=Linux RealMemory=1 AllocMem=0 FreeMem=255243 Sockets=64 Boards=1
  State=IDLE ThreadsPerCore=1 TmpDisk=0 Weight=1 Owner=N/A MCS_label=N/A
  BootTime=2019-11-30T10:18:59 SlurmdStartTime=2019-11-30T10:18:41
  CapWatts=n/a
  CurrentWatts=0 LowestJoules=0 ConsumedJoules=0
  ExtSensorsJoules=n/s ExtSensorsWatts=0 ExtSensorsTemp=n/s

[root@detritus /]# sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
devel*    up    infinite   1    idle* detritus
devel*    up    infinite   3    idle  brick[01-03]
cuda      up    infinite   1    idle* detritus
cuda      up    infinite   1    idle  brick01
fast      up    infinite   3    idle  brick[01-03]
```

From:

<http://detritus.fundacioace.com/wiki/> - **Detritus Wiki**

Permanent link:

<http://detritus.fundacioace.com/wiki/doku.php?id=cluster:restart>

Last update: **2020/08/15 09:34**

