

# Risoluzione dei problemi di elevata disponibilità di CNDP Cluster Manager e ripristino

## Sommario

[Introduzione](#)

[Prerequisiti](#)

[Requisiti](#)

[Componenti usati](#)

[Premesse](#)

[Che cos'è SMI?](#)

[Che cos'è SMI-BM o CNDP?](#)

[Che cos'è SMI Cluster Manager?](#)

[Che cos'è DRBD?](#)

[Problema](#)

[Procedura per la manutenzione](#)

## Introduzione

In questo documento viene descritta la procedura per la risoluzione dei problemi di alta disponibilità (HA, High Availability) in Gestione cluster Cloud Native Deployment Platform (CNDP).

## Prerequisiti

### Requisiti

Cisco raccomanda la conoscenza dei seguenti argomenti:

- Cisco Subscriber Microservices Infrastructure (SMI)
- Architettura 5G CNDP o SMI-Bare-Metal (BM)
- Dispositivo a blocchi replicato distribuito (DRBD)

### Componenti usati

Le informazioni fornite in questo documento si basano sulle seguenti versioni software e hardware:

- SMI 2020.02.2.35
- Kubernetes v1.21.0
- DRBD 8.9.10

Le informazioni discusse in questo documento fanno riferimento a dispositivi usati in uno specifico ambiente di emulazione. Su tutti i dispositivi menzionati nel documento la configurazione è stata ripristinata ai valori predefiniti. Se la rete è operativa, valutare attentamente eventuali conseguenze derivanti dall'uso dei comandi.

# Premesse

## Che cos'è SMI?

Cisco SMI è uno stack a più livelli di tecnologie e standard cloud che abilitano applicazioni basate su microservizi provenienti dalle unità aziendali Cisco Mobility, Cable e Broadband Network Gateway (BNG) - tutte con funzioni di gestione degli abbonati e requisiti simili per i datastore.

Attributi:

- Stack di cloud di livello (tecnologie e standard) per fornire distribuzioni top-to-bottom e supportare anche l'infrastruttura cloud degli utenti correnti.
- Ambiente di esecuzione comune condiviso da tutte le applicazioni per funzioni non applicative (storage dei dati, installazione, configurazione, telemetria, allarme). Ciò consente un'interazione e un'esperienza coerenti per tutti i punti di contatto e i punti di integrazione degli utenti.
- Le applicazioni e gli ambienti di esecuzione comuni vengono implementati in contenitori di microservizi e collegati con una rete Mesh di servizio intelligente.
- API esposta per distribuzione, configurazione e gestione, per consentire l'automazione.

## Che cos'è SMI-BM o CNDP?

Cisco SMI-BM o CNDP è una piattaforma bare-metal gestita che fornisce l'infrastruttura per distribuire le funzionalità di rete virtuale (VNF, Virtual Network Functions) e le funzionalità cloud-native (CNF, Cloud-Native Functions), che consentono alle business unit Cisco di mobilità, cavi e BNG.

Attributi:

- BCM che elimina il sovraccarico relativo a Virtualized Infrastructure Manager (VIM).
- Prestazioni migliorate: Altri core per l'applicazione Esecuzione più rapida delle applicazioni
- Workflow di installazione automatizzato, integrato con Network Services Orchestrator (NSO) Core Function Packet (CFP)
- Stack curato per l'installazione di funzioni di rete Cisco 5G (NF)
- Guida semplificata all'ordine e all'installazione

## Che cos'è SMI Cluster Manager?

Un cluster manager è un cluster a 2 nodi `keepalived` cluster utilizzato come punto iniziale per la distribuzione del cluster control plane e user plane. Esegue un cluster Kubernetes a nodo singolo e un set di POD che sono responsabili della configurazione dell'intero cluster. Solo il cluster manager principale è attivo e il cluster secondario subentra solo in caso di errore o viene disattivato manualmente per motivi di manutenzione.

## Che cos'è DRBD?

DRBD viene utilizzato per aumentare la disponibilità dei dati. Si tratta di un componente software open source basato su Linux che facilita la sostituzione dei sistemi di storage condivisi con un mirroring in rete. In breve, si potrebbe dire che si tratta di un "mirroring Raid 1 basato su rete per i

dati".

## Problema

Gestione cluster è ospitato in un cluster a 2 nodi con DRBD (Distributed Replicated Block Device) e **keepalived**. L'HA può rompersi e può anche entrare in uno stato di separazione del cervello. Questa procedura consente di ripristinare il cluster interrotto. Lo stato desiderato di HA di Gestione cluster è che **cluster manager1 (CM1)** è principale e **cluster manager2 (CM2)** è secondario. Qui, CM1 è la vittima dello split-brain.

## Procedura per la manutenzione

Accedere a Gestione cluster e verificare lo stato del cluster DRBD.

```
cloud-user@pod-name-cm-1:~$ drbd-overview
0:data/0 WfConnection Secondary/Unknown UpToDate/DUnknown
```

```
cloud-user@pod-name-cm-2:~$ drbd-overview
0:data/0 StandAlone Primary/Unknown UpToDate/DUnknown /mnt/stateful_partition ext4 568G 147G
392G 28%
```

In questo scenario, CM2 è primario e il cluster è in modalità standalone. CM1 è attualmente secondario e in **wait for connection state**.

Di seguito è riportato lo stato corretto del cluster:

```
cloud-user@pod-name-deployer-cm-1:~$ drbd-overview
0:data/0 Connected Primary/Secondary UpToDate/UpToDate /mnt/stateful_partition ext4 568G 364G
176G 68%
cloud-user@pod-name-deployer-cm-2:~$ drbd-overview
0:data/0 Connected Secondary/Primary UpToDate/UpToDate Move the CM VIP to CM-1 from CM-2 and
make CM-1 as primary -
```

Spostare l'indirizzo VIP CM in CM-1 da CM-2 e impostare CM-1 come principale (l'indirizzo VIP di Gestione cluster è **10.x.x.65**)

On CM-2 issue below command

```
cloud-user@pod-name-cm-2:~$sudo systemctl restart keepalived
```

On CM-1 issue below command (Make sure that the VIP is now switched over to CM-1)

```
cloud-user@pod-name-cm-1:~$ip a s | grep 10.x.x
inet 10.x.x.70/26 brd 10.x.x.127 scope global vlan1xx ----> here is the server IP
inet 10.x.x.65/26 scope global secondary vlan1xx. ----> here is the VIP
```

Identificare la risorsa DRBD (condivisa in rete):

```
cloud-user@pod-name-deployer-cm-1:/$ cat /etc/fstab
#/mnt/stateful_partition/data /data none defaults,bind 0 0 ----> /data is the resource
#/mnt/stateful_partition/home /home none defaults,bind 0 0
```

```
cloud-user@pod-name-deployer-cm-1:/$ df -h
Filesystem Size Used Avail Use% Mounted on
udev 189G 0 189G 0% /dev
tmpfs 38G 22M 38G 1% /run
```

```
/dev/sda1 9.8G 3.5G 5.9G 37% /
tmpfs 189G 0 189G 0% /dev/shm
tmpfs 5.0M 0 5.0M 0% /run/lock
tmpfs 189G 0 189G 0% /sys/fs/cgroup
/dev/sda4 71G 1.5G 66G 3% /tmp
/dev/sda3 71G 11G 57G 16% /var/log
/dev/drbd0 568G 365G 175G 68% /mnt/stateful_partition -->/dev/drbd0 is the device name
tmpfs 38G 0 38G 0% /run/user/1000
```

```
cloud-user@pod-name-deployer-cm-1:/$ lsblk
NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
sda 8:0 0 744.1G 0 disk
sda1 8:1 0 10G 0 part /
sda2 8:2 0 10G 0 part
sda3 8:3 0 72.2G 0 part /var/log
sda4 8:4 0 72.2G 0 part /tmp
sda5 8:5 0 577.6G 0 part
drbd0 147:0 0 577.5G 0 disk /mnt/stateful_partition ---> /dev/sda5 is used to create drbd0
```

Controllare il file di configurazione DRBD per i dettagli della risorsa:

```
cloud-user@pod-name-deployer-cm-1:/$ cat /etc/drbd.d/data.res
resource data {
protocol C; --->Synchronous replication protocol. Local write operations on the primary node
are considered completed only after both the local and the remote disk write have been
confirmed. As a result, loss of a single node is guaranteed not to lead to any data loss

....
....

device /dev/drbd0;
disk /dev/disk/by-partlabel/smi-state; --> This translates to /dev/sda5
meta-disk internal;
floating 10.192.1.2:7789;
floating 10.192.1.3:7789;
```

Eeguire ora il ripristino DRBD:

On CM-2

```
cloud-user@pod-name-cm-2:~$ sudo systemctl stop keepalived ---> stop to avoid VRRP VIP
switchover
cloud-user@pod-name-cm-2:~$ sudo drbdadm disconnect data ---> data is the cluster resource
cloud-user@pod-name-cm-2:~$ sudo drbdadm secondary data ---> Make it secondary manually
cloud-user@pod-name-cm-2:~$ sudo drbdadm connect --discard-my-data data ---> Force discard of
all modifications on the split brain victim
cloud-user@pod-name-cm-2:~$ drbd-overview status
```

On CM-1:

```
cloud-user@pod-name-cm-1:~$ sudo systemctl stop keepalived ---> stop to avoid VRRP VIP
switchover
cloud-user@pod-name-cm-1:~$ sudo drbdadm connect data ---> Data will be connected as primary
cloud-user@pod-name-cm-1:~$ drbd-overview status
```

Inizio **keepalived** su entrambi i CM. VRRP con l'aiuto di **keepalived** seleziona CM-1 come principale, in base alla risorsa principale connessa **/data**:

```
cloud-user@pod-name-cm-1:~$ sudo systemctl start keepalived
cloud-user@pod-name-cm-1:~$ sudo systemctl status keepalived
```

```
cloud-user@pod-name-cm-2:~$ sudo systemctl start keepalived
cloud-user@pod-name-cm-2:~$ sudo systemctl status keepalived
```

Controllare lo stato di DRBD su CM-1 e CM-2. A questo punto, è necessario trasformarlo nello stato corretto del cluster.

```
cloud-user@pod-name-deployer-cm-1:~$ drbd-overview
0:data/0 Connected Primary/Secondary UpToDate/UpToDate /mnt/stateful_partition ext4 568G 364G
176G 68%
```

```
cloud-user@pod-name-deployer-cm-2:~$ drbd-overview
0:data/0 Connected Secondary/Primary UpToDate/UpToDate Move the CM VIP to CM-1 from CM-2 and
make CM-1 as primary
```

/data è montato solo sul nodo primario.

```
cloud-user@pod-name-deployer-cm-1:~$ df -h
Filesystem Size Used Avail Use% Mounted on
udev 189G 0 189G 0% /dev
tmpfs 38G 22M 38G 1% /run
/dev/sda1 9.8G 3.5G 5.9G 37% /
tmpfs 189G 0 189G 0% /dev/shm
tmpfs 5.0M 0 5.0M 0% /run/lock
tmpfs 189G 0 189G 0% /sys/fs/cgroup
/dev/sda4 71G 1.5G 66G 3% /tmp
/dev/sda3 71G 11G 57G 16% /var/log
/dev/drbd0 568G 364G 175G 68% /mnt/stateful_partition
tmpfs 38G 0 38G 0% /run/user/1000
```

```
cloud-user@pod-name-deployer-cm-secondary:~$ df -h
Filesystem Size Used Avail Use% Mounted on
udev 189G 0 189G 0% /dev
tmpfs 38G 2.3M 38G 1% /run
/dev/sda1 9.8G 2.0G 7.3G 22% /
tmpfs 189G 0 189G 0% /dev/shm
tmpfs 5.0M 0 5.0M 0% /run/lock
tmpfs 189G 0 189G 0% /sys/fs/cgroup
/dev/sda3 71G 9.3G 58G 14% /var/log
/dev/sda4 71G 53M 67G 1% /tmp
tmpfs 38G 0 38G 0% /run/user/1000
```

## Informazioni su questa traduzione

Cisco ha tradotto questo documento utilizzando una combinazione di tecnologie automatiche e umane per offrire ai nostri utenti in tutto il mondo contenuti di supporto nella propria lingua. Si noti che anche la migliore traduzione automatica non sarà mai accurata come quella fornita da un traduttore professionista. Cisco Systems, Inc. non si assume alcuna responsabilità per l'accuratezza di queste traduzioni e consiglia di consultare sempre il documento originale in inglese (disponibile al link fornito).