

# Dépannage des problèmes de haute disponibilité de CNDP Cluster Manager et récupération

## Table des matières

[Introduction](#)

[Conditions préalables](#)

[Exigences](#)

[Composants utilisés](#)

[Informations générales](#)

[Qu'est-ce que SMI ?](#)

[Qu'est-ce que SMI-BM ou CNDP ?](#)

[Qu'est-ce qu'un SMI Cluster Manager ?](#)

[Qu'est-ce que DRBD ?](#)

[Problème](#)

[Procédure de maintenance](#)

## Introduction

Ce document décrit la procédure de dépannage des problèmes de haute disponibilité dans le gestionnaire de cluster CNDP (Cloud Native Deployment Platform).

## Conditions préalables

### Exigences

Cisco vous recommande de prendre connaissance des rubriques suivantes :

- Infrastructure Cisco Subscriber Microservices (SMI)
- Architecture 5G CNDP ou SMI-Bare-metal (BM)
- DRBD (Distributed Replicated Block Device)

### Composants utilisés

Les informations contenues dans ce document sont basées sur les versions de matériel et de logiciel suivantes :

- SMI 2020.02.2.35
- Kubernetes v1.21.0
- DRBD 8.9.10

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Si votre réseau est en ligne, assurez-vous de bien comprendre l'incidence possible des commandes.

# Informations générales

## Qu'est-ce que SMI ?

Cisco SMI est une pile en couches de technologies et de normes cloud qui permettent l'utilisation d'applications basées sur les microservices à partir des unités commerciales Cisco Mobility, Cable et Broadband Network Gateway (BNG), qui ont toutes des fonctions de gestion des abonnés similaires et des exigences similaires en matière de data store.

Attributs :

- La pile de nuages de couches (technologies et normes) pour fournir des déploiements de haut en bas et pour s'adapter à l'infrastructure de nuage utilisateur actuelle.
- Environnement d'exécution commun partagé par toutes les applications pour les fonctions autres que les applications (stockage de données, déploiement, configuration, télémétrie, alarme). Cela garantit une interaction et une expérience cohérentes pour tous les points de contact et d'intégration des utilisateurs.
- Les applications et les environnements d'exécution communs sont déployés dans des conteneurs de microservices et connectés à un maillage de service intelligent.
- API exposée pour le déploiement, la configuration et la gestion, afin de permettre l'automatisation.

## Qu'est-ce que SMI-BM ou CNDP ?

Cisco SMI-BM ou CNDP est une plate-forme sans système d'exploitation organisée qui fournit l'infrastructure nécessaire au déploiement des fonctions de réseau virtuel (VNF) et des fonctions natives du cloud (CNF), qui permettent aux unités commerciales Cisco Mobility, Cable et BNG.

Attributs :

- BM qui élimine les frais généraux liés à Virtualized Infrastructure Manager (VIM).
- Performances améliorées : Plus de coeurs pour l'application Exécution plus rapide des applications
- Workflow de déploiement automatisé, intégré à Network Services Orchestrator (NSO) Core Function Packet (CFP)
- Pile organisée pour déployer les fonctions réseau Cisco 5G (NF)
- Guide de commande et de déploiement simplifié

## Qu'est-ce qu'un SMI Cluster Manager ?

Un gestionnaire de cluster est un système à deux noeuds **keepalived** cluster utilisé comme point initial pour le déploiement du cluster du plan de contrôle et du plan utilisateur. Il exécute un cluster Kubernetes à noeud unique et un ensemble de POD qui sont responsables de la configuration complète du cluster. Seul le gestionnaire de cluster principal est actif et le secondaire prend le relais uniquement en cas de panne ou arrêté manuellement pour maintenance.

## Qu'est-ce que DRBD ?

Le DRBD est utilisé pour accroître la disponibilité des données. Il s'agit d'un composant logiciel open source basé sur Linux qui facilite le remplacement des systèmes de stockage partagés par un miroir en réseau. En bref, vous pouvez dire qu'il s'agit d'un « miroir Raid 1 basé sur le réseau pour les données ».

## Problème

Le gestionnaire de cluster est hébergé dans un cluster à 2 noeuds avec DRBD (Distributed Replicated Block Device) et `keepalived`. La HA peut se briser et elle peut également passer à un état de scission cérébrale. Cette procédure permet de récupérer le cluster endommagé. L'état souhaité de la haute disponibilité du gestionnaire de cluster est le suivant : `cluster manager1 (CM1)` est primaire et `cluster manager2 (CM2)` est secondaire. Ici, CM1 est la victime du dédoublement du cerveau.

## Procédure de maintenance

Connectez-vous au gestionnaire de cluster et vérifiez l'état du 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%
```

Dans ce scénario, CM2 est principal et le cluster est en mode autonome. CM1 est actuellement secondaire et `wait for connection` province.

Voici l'état correct du 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 -
```

Déplacez le VIP CM vers CM-1 de CM-2 et faites de CM-1 le principal (le VIP du gestionnaire de cluster est `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 vlanlxx ----> here is the server IP
inet 10.x.x.65/26 scope global secondary vlanlxx. ----> here is the VIP
```

Identifiez la ressource DRBD (partagée sur le réseau) :

```
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
```

Consultez le fichier de configuration DRBD pour obtenir les détails sur les ressources :

```
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;
```

Effectuez maintenant une restauration 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
```

Début **keepalived** sur les deux MC. VRRP avec l'aide de **keepalived** sélectionne CM-1 comme principal, en fonction de la ressource principale connectée /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
```

Vérifiez l'état DRBD sur CM-1 et CM-2. Il doit maintenant être transformé en l'état de cluster correct.

```
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 est monté uniquement sur le noeud principal.

```
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
```

À propos de cette traduction

Cisco a traduit ce document en traduction automatisée vérifiée par une personne dans le cadre d'un service mondial permettant à nos utilisateurs d'obtenir le contenu d'assistance dans leur propre langue.

Il convient cependant de noter que même la meilleure traduction automatisée ne sera pas aussi précise que celle fournie par un traducteur professionnel.