

Fehlerbehebung bei HA-Problemen mit dem CNDP-Cluster Manager und Wiederherstellung

Inhalt

[Einleitung](#)

[Voraussetzungen](#)

[Anforderungen](#)

[Verwendete Komponenten](#)

[Hintergrundinformationen](#)

[Was ist SMI?](#)

[Was ist SMI-BM oder CNDP?](#)

[Was ist ein SMI Cluster Manager?](#)

[Was ist DRBD?](#)

[Problem](#)

[Verfahren für die Wartung](#)

Einleitung

In diesem Dokument wird das Verfahren zur Behebung von HA-Problemen (High Availability) im Cloud Native Deployment Platform (CNDP)-Cluster-Manager beschrieben.

Voraussetzungen

Anforderungen

Cisco empfiehlt, dass Sie über Kenntnisse in folgenden Bereichen verfügen:

- Cisco Subscriber Microservices Infrastructure (SMI)
- 5G CNDP- oder SMI-Bare-Metal-Architektur (BM)
- Distributed Replicate Block Device (DRBD)

Verwendete Komponenten

Die Informationen in diesem Dokument basierend auf folgenden Software- und Hardware-Versionen:

- SMI 2020,02,2,35
- Kubernetes v1.21.0
- DRBD 8,9,10

Die Informationen in diesem Dokument beziehen sich auf Geräte in einer speziell eingerichteten Testumgebung. Alle Geräte, die in diesem Dokument benutzt wurden, begannen mit einer gelöschten (Nichterfüllungs) Konfiguration. Wenn Ihr Netzwerk in Betrieb ist, stellen Sie sicher, dass Sie die möglichen Auswirkungen aller Befehle verstehen.

Hintergrundinformationen

Was ist SMI?

Cisco SMI ist ein mehrschichtiges Gefüge aus Cloud-Technologien und -Standards, das auf Mikroservices basierende Anwendungen aus den Geschäftsbereichen Cisco Mobility, Cable und Broadband Network Gateway (BNG) ermöglicht. Alle diese Lösungen verfügen über ähnliche Funktionen zur Teilnehmerverwaltung und ähnliche Anforderungen an den Datenspeicher.

Attribute:

- Layer-Cloud-Stack (Technologien und Standards) zur Bereitstellung von Top-to-Bottom-Bereitstellungen und zur Anpassung an die aktuelle Cloud-Infrastruktur der Benutzer
- Gemeinsame Ausführungsumgebung, die von allen Anwendungen für Funktionen genutzt wird, die keine Anwendungen sind (Datenspeicherung, Bereitstellung, Konfiguration, Telemetrie, Alarm). Dies sorgt für eine konsistente Interaktion und Benutzererfahrung für alle Berührungspunkte und Integrationspunkte der Benutzer.
- Anwendungen und allgemeine Ausführungsumgebungen werden in Mikroservicecontainern bereitgestellt und mit einem Intelligent Service Mesh verbunden.
- Verfügbare API für Bereitstellung, Konfiguration und Management zur Automatisierung

Was ist SMI-BM oder CNDP?

Cisco SMI-BM oder CNDP ist eine kuratierte Bare-Metal-Plattform, die die Infrastruktur für die Bereitstellung von Virtual Network Functions (VNF) und Cloud-Native Functions (CNFs) bereitstellt, die die Geschäftsbereiche Cisco Mobility, Cable und BNG unterstützen.

Attribute:

- IBM, das den Overhead im Zusammenhang mit dem Virtualized Infrastructure Manager (VIM) eliminiert.
- Verbesserte Leistung: Mehr Cores für die AnwendungSchnellere Anwendungsausführung
- Automatisierter Bereitstellungs-Workflow, integriert in Network Services Orchestrator (NSO) Core Function Packet (CFP)
- Geplanter Stack für die Bereitstellung von Cisco 5G-Netzwerkfunktionen (NFs)
- Leitfaden für vereinfachte Bestellung und Bereitstellung

Was ist ein SMI Cluster Manager?

Ein Cluster-Manager ist ein 2-Node- `keepalived` Cluster, der als Ausgangspunkt für die Cluster-Bereitstellung auf Kontroll- und Benutzerebene verwendet wird. Es führt einen Kubernetes-Cluster mit einem Knoten und eine Reihe von PODs aus, die für die gesamte Cluster-Einrichtung verantwortlich sind. Nur der primäre Cluster-Manager ist aktiv, und der sekundäre übernimmt nur bei einem Ausfall oder wird aus Wartungsgründen manuell heruntergefahren.

Was ist DRBD?

DRBD wird eingesetzt, um die Verfügbarkeit von Daten zu erhöhen. Es handelt sich um eine Linux-basierte Open-Source-Softwarekomponente, die den Austausch von gemeinsam genutzten

Speichersystemen durch eine Netzwerkspiegelung erleichtert. Kurz gesagt, Sie können sagen, dies ist ein "Netzwerk-basiertes Raid 1-Mirror für die Daten".

Problem

Der Cluster-Manager wird in einem Cluster mit zwei Knoten und Distributed Replicate Block Device (DRBD) gehostet und `keepalived`. Die HA kann abbrechen und auch in einen Split-Brain-Zustand übergehen. Mit diesem Verfahren können Sie den defekten Cluster wiederherstellen. Der gewünschte Zustand des Cluster-Managers HA ist folgender: `cluster manager1 (CM1)` ist primär und `cluster manager2 (CM2)` ist sekundär. Hier ist CM1 das Split-Brain-Opfer.

Verfahren für die Wartung

Melden Sie sich beim Cluster-Manager an, und überprüfen Sie den Status des DRBD-Clusters.

```
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 diesem Szenario ist CM2 primär und der Cluster befindet sich im Standalone-Modus. CM1 ist derzeit sekundär und in `wait for connection status`.

Der richtige Status des Clusters:

```
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 -
```

CM-VIP von CM-2 auf CM-1 verschieben und CM-1 als primär festlegen (Cluster-Manager-VIP ist `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
```

Identifizieren Sie die DRBD-Ressource (über das Netzwerk freigegeben):

```
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
```

In der DRBD-Konfigurationsdatei finden Sie die Ressourcendetails:

```
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;
```

Führen Sie jetzt die DRBD-Wiederherstellung durch:

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
```

Start **keepalived** auf beiden CMs. VRRP mithilfe von **keepalived** wählt CM-1 als primär, basierend auf der verbundenen Primärressource **/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
```

Überprüfen Sie den DRBD-Status von CM-1 und CM-2. Sie muss jetzt in den richtigen Cluster-Status umgewandelt werden.

```
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 ist nur auf dem Primärknoten gemountet.

```
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
```

Informationen zu dieser Übersetzung

Cisco hat dieses Dokument maschinell übersetzen und von einem menschlichen Übersetzer editieren und korrigieren lassen, um unseren Benutzern auf der ganzen Welt Support-Inhalte in ihrer eigenen Sprache zu bieten. Bitte beachten Sie, dass selbst die beste maschinelle Übersetzung nicht so genau ist wie eine von einem professionellen Übersetzer angefertigte. Cisco Systems, Inc. übernimmt keine Haftung für die Richtigkeit dieser Übersetzungen und empfiehlt, immer das englische Originaldokument (siehe bereitgestellter Link) heranzuziehen.