

Resolución de problemas de HA del administrador de clústeres CNDP y recuperación

Contenido

[Introducción](#)

[Prerequisites](#)

[Requirements](#)

[Componentes Utilizados](#)

[Antecedentes](#)

[¿Qué es SMI?](#)

[¿Qué es SMI-BM o CNDP?](#)

[¿Qué es un administrador de clústeres SMI?](#)

[¿Qué es DRBD?](#)

[Problema](#)

[Procedimiento de mantenimiento](#)

Introducción

En este documento se describe el procedimiento para solucionar problemas de alta disponibilidad (HA) en el administrador de clústeres de la plataforma de implementación nativa en la nube (CNDP).

Prerequisites

Requirements

Cisco recomienda que tenga conocimiento sobre estos temas:

- Cisco Subscriber Microservices Infrastructure (SMI)
- Arquitectura 5G CNDP o SMI-Bare-metal (BM)
- Dispositivo de bloque replicado distribuido (DRBD)

Componentes Utilizados

La información que contiene este documento se basa en las siguientes versiones de software y hardware.

- SMI 2020.2.2.35
- Kubernetes v1.21.0
- DRBD 8.9.10

La información que contiene este documento se creó a partir de los dispositivos en un ambiente de laboratorio específico. Todos los dispositivos que se utilizan en este documento se pusieron en funcionamiento con una configuración verificada (predeterminada). Si tiene una red en vivo, asegúrese de entender el posible impacto de cualquier comando.

Antecedentes

¿Qué es SMI?

Cisco SMI es una pila por niveles de tecnologías y estándares de nube que permiten el uso de aplicaciones basadas en microservicios de las unidades empresariales Cisco Mobility, Cable y Broadband Network Gateway (BNG), todas las cuales tienen funciones de gestión de suscriptores y requisitos de almacén de datos similares.

Atributos:

- Layer Cloud Stack (tecnologías y estándares) para proporcionar implementaciones de arriba a abajo y también dar cabida a la infraestructura de nube de usuario actual.
- Entorno de ejecución común compartido por todas las aplicaciones para funciones que no son de aplicación (almacenamiento de datos, implementación, configuración, telemetría, alarma). Esto proporciona una interacción y experiencia uniformes para todos los puntos de contacto y puntos de integración del usuario.
- Las aplicaciones y los entornos de ejecución comunes se implementan en contenedores de microservicios y se conectan con una malla de servicio inteligente.
- API expuesta para implementación, configuración y gestión, con el fin de habilitar la automatización.

¿Qué es SMI-BM o CNDP?

Cisco SMI-BM o CNDP es una plataforma sin software específico con funciones de curaduría que proporciona la infraestructura necesaria para implementar funciones de red virtual (VNF) y funciones nativas de la nube (CNF), que habilitan las unidades empresariales de Cisco Mobility, Cable y BNG.

Atributos:

- IBM que elimina la sobrecarga relacionada con Virtualized Infrastructure Manager (VIM).
- Rendimiento mejorado: Más núcleos para la aplicación Ejecución de aplicaciones más rápida
- Flujo de trabajo de implementación automatizado, integrado con el paquete de funciones principales (CFP) de Network Services Orchestrator (NSO)
- Pila revisada para implementar las funciones de red (NF) 5G de Cisco
- Guía de implementación y pedido simplificada

¿Qué es un administrador de clústeres SMI?

Un administrador de clústeres es de 2 nodos `keepalived` clúster utilizado como punto inicial para la implementación de clústeres tanto del plano de control como del plano de usuario. Ejecuta un clúster de Kubernetes de un solo nodo y un conjunto de POD que son responsables de toda la configuración del clúster. Sólo el administrador de clústeres principal está activo y el secundario asume el control sólo en caso de error o se desactiva manualmente para mantenimiento.

¿Qué es DRBD?

DRBD se utiliza para aumentar la disponibilidad de los datos. Se trata de un componente de software de código abierto basado en Linux que facilita la sustitución de los sistemas de almacenamiento compartido por un espejo en red. En resumen, puede decir que se trata de un "reflejo Raid 1 basado en red para los datos".

Problema

El administrador de clústeres está alojado en un clúster de 2 nodos con dispositivo de bloque replicado distribuido (DRBD) y `keepalived`. La AH puede romperse y también puede entrar en un estado de cerebro partido. Este procedimiento ayuda a recuperar el clúster dañado. El estado deseado del HA del administrador de clústeres es que `cluster manager1 (CM1)` es principal y `cluster manager2 (CM2)` es secundario. Aquí, CM1 es la víctima del cerebro partido.

Procedimiento de mantenimiento

Inicie sesión en el administrador de clústeres y compruebe el estado del clúster 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%
```

En esta situación, CM2 es principal y el clúster está en modo independiente. CM1 es secundario actualmente y está en `wait for connection` estado.

Este es el estado correcto del clúster:

```
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 -
```

Mueva el CM VIP a CM-1 desde CM-2 y convierta CM-1 en principal (el VIP del administrador de clústeres es `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**

Identifique el recurso DRBD (compartido a través de la red):

```
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
```

Verifique el archivo de configuración DRBD para obtener los detalles del recurso:

```
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;
```

Ahora, realice la recuperación 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
```

Inicio **keepalived** en ambos CM. VRRP con la ayuda de **keepalived** selecciona CM-1 como principal, basándose en el recurso principal conectado **/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
```

Compruebe el estado DRBD en CM-1 y CM-2. Debe ser transformado al estado correcto del cluster por ahora.

```
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 está montado sólo en el nodo 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
```

Acerca de esta traducción

Cisco ha traducido este documento combinando la traducción automática y los recursos humanos a fin de ofrecer a nuestros usuarios en todo el mundo contenido en su propio idioma.

Tenga en cuenta que incluso la mejor traducción automática podría no ser tan precisa como la proporcionada por un traductor profesional.

Cisco Systems, Inc. no asume ninguna responsabilidad por la precisión de estas traducciones y recomienda remitirse siempre al documento original escrito en inglés (insertar vínculo URL).