

CNDP 클러스터 관리자 HA 문제 해결 및 복구

목차

[소개](#)

[사전 요구 사항](#)

[요구 사항](#)

[사용되는 구성 요소](#)

[배경 정보](#)

[SMI란 무엇입니까?](#)

[SMI-BM 또는 CNDP란?](#)

[SMI Cluster Manager란?](#)

[DRBD란 무엇입니까?](#)

[문제](#)

[유지 보수 절차](#)

소개

이 문서에서는 CNDP(Cloud Native Deployment Platform) 클러스터 관리자에서 HA(고가용성) 문제를 해결하는 절차에 대해 설명합니다.

사전 요구 사항

요구 사항

다음 주제에 대한 지식을 보유하고 있으면 유용합니다.

- Cisco Subscriber SMI(Microservices Infrastructure)
- 5G CNDP 또는 SMI-Bare-Metal(BM) 아키텍처
- DRBD(분산 복제 블록 장치)

사용되는 구성 요소

이 문서의 정보는 다음 소프트웨어 및 하드웨어 버전을 기반으로 합니다.

- SMI 2020.02.2.35
- 쿠버네티스 v1.21.0
- DRBD 8.9.10

이 문서의 정보는 특정 랩 환경의 디바이스를 토대로 작성되었습니다. 이 문서에 사용된 모든 디바이스는 초기화된(기본) 컨피그레이션으로 시작되었습니다. 현재 네트워크가 작동 중인 경우 모든 명령의 잠재적인 영향을 미리 숙지하시기 바랍니다.

배경 정보

SMI란 무엇입니까?

Cisco SMI는 Cisco Mobility, Cable 및 BNG(Broadband Network Gateway) 비즈니스 유닛에서 마이크로서비스 기반 애플리케이션을 활성화하는 클라우드 기술 및 표준의 레이어 스택으로서, 모두 유사한 가입자 관리 기능 및 유사한 데이터 저장소 요구 사항을 갖습니다.

특성:

- 레이어 클라우드 스택(기술 및 표준)을 통해 하향식 구축을 제공하고 현재 사용자 클라우드 인프라를 수용합니다.
- 모든 애플리케이션이 공유하는 공통 실행 환경에서 비애플리케이션 기능(데이터 스토리지, 구축, 구성, 텔레메트리, 알람)을 지원합니다. 따라서 모든 사용자 접점과 통합 지점에 일관된 상호 작용 및 환경이 제공됩니다.
- 애플리케이션 및 공통 실행 환경은 마이크로서비스 컨테이너에 구축되며 Intelligent Service Mesh로 연결됩니다.
- 자동화를 지원하기 위한 구축, 구성 및 관리를 위한 노출된 API

SMI-BM 또는 CNDP란?

Cisco SMI-BM 또는 CNDP는 Cisco 모빌리티, 케이블 및 BNG 비즈니스 유닛을 활성화하는 VNF(Virtual Network Functions) 및 CNF(Cloud-Native Functions)를 구축할 수 있는 인프라를 제공하는 큐레이션된 베어 메탈 플랫폼입니다.

특성:

- VIM(Virtualized Infrastructure Manager) 관련 오버헤드를 제거하는 BM
- 향상된 성능: 애플리케이션에 더 많은 코어 지원 애플리케이션 실행 속도 향상
- NSO(Network Services Orchestrator) CFP(Core Function Packet)와 통합된 자동화된 구축 워크플로
- Cisco 5G NF(Network Functions)를 구축하기 위한 Curated Stack
- 간소화된 주문 및 구축 설명서

SMI Cluster Manager란?

클러스터 관리자는 2노드 `keepalived` 클러스터가 컨트롤 플레인 및 사용자 플레인 클러스터 구축의 초기 지점으로 사용됩니다. 단일 노드 Kubernetes 클러스터 및 전체 클러스터 설정을 담당하는 POD 집합을 실행합니다. 기본 클러스터 관리자만 활성 상태이며, 보조 관리자는 장애가 발생하거나 유지 보수를 위해 수동으로 작동 중지된 경우에만 인계받습니다.

DRBD란 무엇입니까?

DRBD는 데이터의 가용성을 높이기 위해 사용됩니다. Linux 기반의 오픈 소스 소프트웨어 구성요소로서 네트워크 미러로 공유 스토리지 시스템을 쉽게 교체할 수 있습니다. 간단히 말해, 이는 "데이터를 위한 네트워크 기반 RAID 1 미러"라고 할 수 있습니다.

문제

클러스터 관리자는 DRBD(Distributed Replicated Block Device)가 있는 2노드 클러스터에서 호스팅되며 `keepalived`. HA는 깨질 수 있고 또한 분할 뇌 상태로 들어갈 수 있습니다. 이 절차는 손상된 클러

스터를 복구하는 데 도움이 됩니다. 클러스터 관리자 HA의 원하는 상태는 **cluster manager1 (CM1)** 기본 및 **cluster manager2 (CM2)** 보조 여기서 CM1은 뇌가 분열된 희생자입니다.

유지 보수 절차

클러스터 관리자에 로그인하여 DRBD 클러스터의 상태를 확인합니다.

```
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%
```

이 시나리오에서 CM2는 기본 클러스터이며 클러스터는 독립형 모드입니다. CM1은 현재 보조 상태이며 **wait for connection** 상태.

다음은 클러스터의 올바른 상태입니다.

```
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-2에서 CM-1로 CM VIP를 이동하고 CM-1을 기본으로 설정(클러스터 관리자 VIP는 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
```

DRBD 리소스(네트워크를 통해 공유됨)를 식별합니다.

```
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
```

DRBD 컨피그레이션 파일에서 리소스 세부 정보를 확인합니다.

```
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;
```

이제 DRBD 복구를 수행합니다.

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
```

시작하기 **keepalived** 두 CM에서 모두 처리합니다. VRRP를 통해 **keepalived** 연결된 기본 리소스를 기반으로 CM-1을 기본 리소스로 선택합니다. /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
```

CM-1 및 CM-2의 DRBD 상태를 확인합니다. 지금까지는 올바른 클러스터 상태로 변환해야 합니다.

```
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 기본 노드에만 마운트됩니다.

```
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
```

이 번역에 관하여

Cisco는 전 세계 사용자에게 다양한 언어로 지원 콘텐츠를 제공하기 위해 기계 번역 기술과 수작업 번역을 병행하여 이 문서를 번역했습니다. 아무리 품질이 높은 기계 번역이라도 전문 번역가의 번역 결과물만큼 정확하지는 않습니다. Cisco Systems, Inc.는 이 같은 번역에 대해 어떠한 책임도 지지 않으며 항상 원본 영문 문서(링크 제공됨)를 참조할 것을 권장합니다.