

CNDP Cluster ManagerのHA問題のトラブルシューティングと回復

内容

[概要](#)

[前提条件](#)

[要件](#)

[使用するコンポーネント](#)

[背景説明](#)

[SMIとは何ですか。](#)

[SMI-BMまたはCNDPとは何ですか。](#)

[SMIクラスタ・マネージャについて教えてください。](#)

[DRBDとは](#)

[問題](#)

[メンテナンスの手順](#)

概要

このドキュメントでは、Cloud Native Deployment Platform(CNDP)クラスタマネージャでハイアベイラビリティ(HA)の問題をトラブルシューティングする手順について説明します。

前提条件

要件

次の項目に関する知識があることが推奨されます。

- Cisco Subscriber Microservices Infrastructure(SMI)
- 5G CNDPまたはSMI-Bare-metal(BM)アーキテクチャ
- Distributed Replicated Block Device(DRBD)

使用するコンポーネント

このドキュメントの情報は、次のソフトウェアとハードウェアのバージョンに基づいています。

- SMI 2020.02.2.35
- Kubernetes v1.21.0
- DRBD 8.9.10

このドキュメントの情報は、特定のラボ環境にあるデバイスに基づいて作成されました。このドキュメントで使用するすべてのデバイスは、初期(デフォルト)設定の状態から起動しています。本稼働中のネットワークでは、各コマンドによって起こる可能性がある影響を十分確認してください。

背景説明

SMIとは何ですか。

Cisco SMIは、シスコのモビリティ、ケーブル、およびブロードバンドネットワークゲートウェイ(BNG)ビジネスユニットからのマイクロサービスベースのアプリケーションを可能にする、クラウドテクノロジーと標準の階層化されたスタックです。これらはすべて、同様の加入者管理機能と同様のデータストア要件を備えています。

属性:

- レイヤCloud Stack (テクノロジーと標準) : トップツーツボトムの導入を実現し、現在のユーザクラウドインフラストラクチャにも対応します。
- アプリケーション以外の機能 (データストレージ、導入、設定、テレメトリ、アラーム) に対して、すべてのアプリケーションで共有される共通実行環境。これにより、すべてのユーザタッチポイントと統合ポイントで一貫したインタラクションとエクスペリエンスが提供されます。
- アプリケーションとCommon Execution Environmentsは、マイクロサービスコンテナに導入され、インテリジェントサービスメッシュに接続されます。
- 自動化を可能にする、導入、設定、および管理のための公開API。

SMI-BMまたはCNDPとは何ですか。

Cisco SMI-BMまたはCNDPは、Virtual Network Functions(VNF)およびCloud-Native Functions(CNF)を導入するためのインフラストラクチャを提供するベアメタルプラットフォームで、シスコのモビリティ、ケーブル、およびBNGビジネスユニットを実現します。

属性:

- Virtualized Infrastructure Manager(VIM)関連のオーバーヘッドを排除するBM。
- パフォーマンスの向上 : アプリケーションのコア数の増加アプリケーション実行の高速化
- Network Services Orchestrator(NSO) Core Function Packet(CFP)と統合された自動導入ワークフロー
- Cisco 5Gネットワーク機能(NF)を導入するためのキューレーションスタック
- シンプルな注文および導入ガイド

SMIクラスタ・マネージャについて教えてください。

クラスタマネージャは2ノードです `keepalived` クラスタは、コントロールプレーンとユーザプレーンクラスタの両方の導入の最初のポイントとして使用されます。単一ノードのKubernetesクラスタと、クラスタ全体のセットアップを担当する一連のPODを実行します。プライマリクラスタマネージャだけがアクティブで、セカンダリは障害発生時またはメンテナンスのために手動でダウンした場合にのみ引き継ぎます。

DRBDとは

DRBDは、データの可用性を高めるために使用されます。Linuxベースのオープン・ソース・ソフトウェア・コンポーネントであり、ネットワーク・ミラーによる共有ストレージ・システムの交換を容易にします。つまり、これは「データ用のネットワークベースのRaid 1ミラー」と言えます。

す。

問題

クラスタマネージャは、Distributed Replicated Block Device(DRBD)を使用する2ノードクラスタでホストされます。keepalived. HAが壊れ、スプリットブレイン状態に入る可能性もあります。この手順は、破損したクラスタを回復するのに役立ちます。クラスタマネージャHAの望ましい状態は、cluster manager1 (CM1) はプライマリであり、cluster manager2 (CM2) セカンダリです。ここで、CM1はスプリットブレインの被害者です。

メンテナンスの手順

クラスタマネージャにログインし、DRBDクラスタのステータスを確認します。

```
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%
```

このシナリオでは、CM2はプライマリで、クラスタはスタンドアロンモードです。CM1は現在セカンダリであり、wait for connection 変わります。

クラスタの正しい状態を次に示します。

```
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-2からCM-1にCM VIPを移動し、CM-1をプライマリにします(Cluster Manager VIPは 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リソース (ネットワーク上で共有) を特定します。

```
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
```

リソースの詳細については、DRBD設定ファイルを確認します。

```
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;
```

次に、DRBDリカバリを実行します。

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
```

開始 **keepalived** 両方のCMでプロセスを実行します。VRRPを使用して **keepalived** 接続されたプライマリリソースに基づいて、CM-1をプライマリとして選択します /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とCM-2のDRBDステータスをチェックします。これで、正しいクラスタ状態に変換する必要があります。

```
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 はプライマリノードにのみマウントされます。

```
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
```

翻訳について

シスコは世界中のユーザにそれぞれの言語でサポート コンテンツを提供するために、機械と人による翻訳を組み合わせて、本ドキュメントを翻訳しています。ただし、最高度の機械翻訳であっても、専門家による翻訳のような正確性は確保されません。シスコは、これら翻訳の正確性について法的責任を負いません。原典である英語版（リンクからアクセス可能）もあわせて参照することを推奨します。