



CHAPTER 6

ハイ アベイラビリティ

この章では、ハイ アベイラビリティに関する問題を識別して解決する方法について説明します。

この章は、次の内容で構成されています。

- 「[ハイ アベイラビリティについて](#)」 (P.6-1)
- 「[ハイ アベイラビリティでの問題](#)」 (P.6-3)
- 「[ハイ アベイラビリティのトラブルシューティング コマンド](#)」 (P.6-5)

ハイ アベイラビリティについて

ハイ アベイラビリティ (HA) の目的は、システム内で発生したハードウェアの障害およびソフトウェアのエラーの両方による影響を一定範囲内に抑えることです。Cisco NX-OS オペレーティング システムは、ネットワーク レベル、システム レベル、およびサービス レベルでのハイ アベイラビリティに向けて設計されています。

次の Cisco NX-OS 機能が、障害発生時のトラフィックの中断を防止または最小限に留めます。

- 冗長性：ソフトウェア アーキテクチャのあらゆる面での冗長性。
- プロセスの分離：1 つのプロセス内で発生したエラーで他のプロセスが中断されるのを防ぐためのソフトウェア コンポーネント間の分離。
- 再起動性：ほとんどのシステム機能およびサービスが分離されているため、エラーが発生しても、他のサービスは実行され続けている中で独立して再起動が可能。さらに、ほとんどのシステム サービスはステートフルな再起動を実行するため、その他のサービスに対して透過的に稼働を再開できます。
- スーパーバイザ ステートフル スイッチオーバー：アクティブ/スタンバイ デュアル スーパーバイザ設定。状態と設定が 2 つの仮想スーパーバイザ モジュール (VSM) の間で一貫して同期された状態が保たれ、VSM に障害が発生してもシームレスでステートフルなスイッチオーバーが提供されます。

Cisco Nexus 1000V システムは、次のコンポーネントで構成されます。

- 仮想サーバ内で実行される仮想イーサネット モジュール (VEM)。これらは、VSM 内でモジュールとして表されます。
- リモート管理コンポーネント。たとえば、VMware vCenter Server。
- 仮想マシン (VM) 内で実行される 1 つまたは 2 つの VSM。

システムレベルのハイ アベイラビリティ

Cisco Nexus 1000V では、HA ペアとして実行される冗長 VSM 仮想マシン（プライマリとセカンダリ）がサポートされています。デュアル VSM は、同時にはどちらか片方の VSM しかアクティブになれず、他方はスタンバイ バックアップとして機能する、アクティブ/スタンバイ キャパシティで稼働します。状態と設定が、2 つの VSM 間で一貫して同期された状態が保たれ、アクティブな VSM に障害が発生したときにはステートフルなスイッチオーバーが提供されます。

シングルまたはデュアル スーパーバイザ

Cisco Nexus 1000V システムは、次のコンポーネントで構成されます。

- 仮想サーバ内で実行される仮想イーサネット モジュール（VEM）。これらは、VSM 内でモジュールとして表されます。
- リモート管理コンポーネント。たとえば、VMware vCenter Server。
- 仮想マシン（VM）内で実行される 1 つまたは 2 つの仮想スーパーバイザ モジュール（VSM）。

| シングル VSM 運用 | デュアル VSM 運用 |
|---|--|
| <ul style="list-style-type: none"> • ステートレス：サービスはスタートアップ コンフィギュレーションから起動されます。 • ステートフル：サービスは以前の状態で再開されます。 | <ul style="list-style-type: none"> • 1 つのアクティブな VSM と 1 つのスタンバイ VSM。 • アクティブな VSM は、すべてのシステム アプリケーションを実行し、システムを制御します。 • スタンバイ VSM では、アプリケーションはスタンバイ モードで開始され、初期化されます。また、「実行準備完了」のランタイム コンテキストを維持するために、アクティブな VSM と同期化され、最新の状態に保たれます。 • スwitchオーバーでは、スタンバイ VSM がアクティブ VSM の役目を引き継ぎます。 |

ネットワークレベル ハイ アベイラビリティ

ネットワーク レベルでの Cisco Nexus 1000V HA には、ポート チャネルと Link Aggregation Control Protocol（LACP）が含まれます。ポート チャネルは、物理リンクをまとめて 1 つのチャネル グループに入れ、最大 8 つの物理リンクの帯域幅を集約した単一の論理リンクを作ります。ポート チャネル内のメンバー ポートに障害が発生すると、障害が発生したリンクで伝送されていたトラフィックはポート チャネル内のその他のメンバー ポートに切り替わります。

さらに、LACP では最大 16 個のインターフェイスが 1 つのポート チャネルに入るように設定できます。最大 8 つのインターフェイスをアクティブにすることができ、最大 8 つのインターフェイスをスタンバイ状態に入れることができます。

ポート チャネルと LACP のその他の情報については、『Cisco Nexus 1000V Layer 2 Switching Configuration Guide, Release 4.0』を参照してください。

ハイアベイラビリティでの問題

| 症状 | 考えられる原因 | 解決策 |
|-------------------------------------|--|---|
| アクティブな VSM からスタンバイ VSM が見えない。 | <p>ロールが正しく設定されていない。</p> <ul style="list-style-type: none"> • show system redundancy status コマンドを使用して、2つの VSM のロールをチェックします。 | <ol style="list-style-type: none"> 1. ロールがそれぞれプライマリ ロールとセカンダリ ロールになっていることを確認します。 2. 必要に応じて、system redundancy role コマンドを使用して、状況を修正します。 3. ロールを変更した場合は、設定を保存します。 |
| | <p>ネットワーク接続の問題。</p> <ul style="list-style-type: none"> • アップストリーム スイッチおよび仮想スイッチでの VSM 間の VLAN 接続の制御と管理をチェックします。 | <p>ネットワークの問題が存在する場合は、次の手順を実行します。</p> <ol style="list-style-type: none"> 1. vSphere Client から、VSM (スタンバイ モードになっているはず) をシャットダウンします。 2. vSphere Client から、ネットワーク接続を復元させたあとに、スタンバイ VSM を UP にします。 |
| アクティブ VSM がスタンバイ VSM と完全に同期化されていない。 | <p>VSM 間のバージョンの不一致。</p> <ul style="list-style-type: none"> • show version コマンドを使用して、プライマリ VSM とセカンダリ VSM が同じイメージバージョンを使用しているかチェックします。 | <p>アクティブ VSM とスタンバイ VSM のソフトウェアバージョンが異なる場合は、セカンダリ VSM をプライマリ VSM で使用されているのと同じバージョンで再インストールします。</p> |
| | <p>gsync プロセス中に致命的なエラーが発生した。</p> <ul style="list-style-type: none"> • show system internal log sysmgr gsyncctrl コマンドを使用して gsyncctrl ログをチェックし、致命的なエラーを探します。 | <p>reload module module-number コマンドを使用して、スタンバイ VSM を再ロードします。module-number は、スタンバイ VSM のモジュール番号です。</p> |

| 症状 | 考えられる原因 | 解決策 |
|------------------------|--|---|
| スタンバイ VSM が周期的に再起動される。 | <p>VSM が管理インターフェイスを通じてしか接続を持っていない。</p> <ul style="list-style-type: none"> VSM が管理インターフェイスを通じては通信できるけれども、制御インターフェイスを通じては通信できない場合、アクティブ VSM がその状況を検出し、2 つの VSM が HA モードになって同期が取れなくなるのを防ぐために、スタンバイ VSM をリセットします。 show system internal redundancy info コマンドの出力をチェックして、<i>degraded_mode</i> フラグが <i>true</i> に設定されていることを確認します。 | <p>プライマリ VSM とセカンダリ VSM の間の制御 VLAN 接続をチェックします。</p> |
| | <p>VSM のバージョンが異なる。</p> <p>debug system internal sysmgr all コマンドを入力して、バージョンの不一致を示す <i>active_verctrl</i> エントリを探します。次のような出力が表示されます。</p> <pre>2009 May 5 08:34:15.721920 sysmgr: active_verctrl: Stdby running diff version- force download the standby sup.</pre> | <p>スタンバイ VSM を分離してから、起動します。</p> <p>show version コマンドを使用して、両方の VSM のソフトウェアバージョンをチェックします。</p> <p>アクティブ VSM と同じバージョンのイメージをスタンバイ VSM にインストールします。</p> |

| 症状 | 考えられる原因 | 解決策 |
|-----------------------|--|--|
| 両方の VSM がアクティブモードになる。 | <p>ネットワーク接続の問題。</p> <ul style="list-style-type: none"> アップストリームと仮想スイッチで、VSM 間の制御 VLAN 接続と管理 VLAN 接続を調べます。 VSM がこの2つのインターフェイスのどちらを介しても通信できない場合、その両方がアクティブ VSM になろうとします。 | <p>ネットワークの問題が存在する場合は、次の手順を実行します。</p> <ol style="list-style-type: none"> vSphere Client から、VSM (スタンバイモードになっているはず) をシャットダウンします。 vSphere Client から、ネットワーク接続を復元させたあとに、スタンバイ VSM を UP にします。 |
| | <p>2つの VSM でのドメイン ID が異なる。</p> <p>show system internal redundancy info コマンドを使用して、<i>domain</i> 値を調べます。</p> | <p>必要であれば、ドメイン ID を更新し、それをスタートアップコンフィギュレーションに保存します。</p> <ul style="list-style-type: none"> デュアル VSM システムでドメイン ID をアップグレードする場合は、決まった手順に従う必要があります。 <ul style="list-style-type: none"> 正しいドメイン ID を持つ VSM を分離して、他方の VSM と通信できないようにします。 分離した VSM でドメイン ID を変更し、設定を保存し、VSM の電源を切ります。 分離した VSM を再接続し、電源を入れます。 |

ハイアベイラビリティのトラブルシューティングコマンド

ここでは、ハイアベイラビリティに関する問題のトラブルシューティングに使用できるコマンドを示します。

プロセスログとコアのリストを表示するには、次のコマンドを使用します。

- show cores**

```
n1000V# show cores
VDC No Module-num      Process-name      PID      Core-create-time
-----
1          1          private-vlan     3207     Apr 28 13:29
```

- show processes log [pid pid]**

```
n1000V# show processes log
VDC Process      PID      Normal-exit  Stack  Core  Log-create-time
-----
1 private-vlan   3207          N      Y      N     Tue Apr 28 13:29:48 2009
```

```
n1000V# show processes log pid 3207
```

```
=====
Service: private-vlan
Description: Private VLAN
```

```

Started at Wed Apr 22 18:41:25 2009 (235489 us)
Stopped at Tue Apr 28 13:29:48 2009 (309243 us)
Uptime: 5 days 18 hours 48 minutes 23 seconds

Start type: SRV_OPTION_RESTART_STATELESS (23)
Death reason: SYSMGR_DEATH_REASON_FAILURE_SIGNAL (2) <-- Reason for the process abort
Last heartbeat 46.88 secs ago
System image name: nexus-1000v-mzg.4.0.4.SV1.1.bin
System image version: 4.0(4)SV1(1) S25

PID: 3207
Exit code: signal 6 (core dumped) <-- Indicates that a cores for the process was generated.

CWD: /var/sysmgr/work
...

```

冗長性ステータスをチェックするには、次のコマンドを使用します。

- **show system redundancy status**

```

N1000V# show system redundancy status
Redundancy role
-----
      administrative: primary <-- Configured redundancy role
      operational:    primary <-- Current operational redundancy role

Redundancy mode
-----
      administrative: HA
      operational:    HA

This supervisor (sup-1)
-----
      Redundancy state: Active <-- Redundancy state of this VSM
      Supervisor state: Active
      Internal state:   Active with HA standby

Other supervisor (sup-2)
-----
      Redundancy state: Standby <-- Redundancy state of the other VSM
      Supervisor state: HA standby
      Internal state:   HA standby <-- The standby VSM is in HA mode and in sync

```

システム内部の冗長性ステータスをチェックするには、次のコマンドを使用します。

- **show system internal redundancy info**

```

n1000V# show system internal redundancy info
My CP:
  slot: 0
  domain: 184 <-- Domain id used by this VSM
  role:    primary <-- Redundancy role of this VSM
  status: RDN_ST_AC <-- Indicates redundancy state (RDN_ST) of the this VSM is Active
  (AC)
  state: RDN_DRV_ST_AC_SB
  intr:  enabled
  power_off_reqs: 0
  reset_reqs:    0
Other CP:
  slot: 1
  status: RDN_ST_SB <-- Indicates redundancy state (RDN_ST) of the other VSM is
  Standby (SB)
  active: true

```

```

ver_rcvd: true
  degraded_mode: false <-- When true, it indicates that communication through the
control interface is faulty
Redun Device 0: <-- This device maps to the control interface
  name: ha0
  pdev: ad7b6c60
  alarm: false
  mac: 00:50:56:b7:4b:59
  tx_set_ver_req_pkts: 11590
  tx_set_ver_rsp_pkts: 4
  tx_heartbeat_req_pkts: 442571
  tx_heartbeat_rsp_pkts: 6
  rx_set_ver_req_pkts: 4
  rx_set_ver_rsp_pkts: 1
  rx_heartbeat_req_pkts: 6
  rx_heartbeat_rsp_pkts: 442546 <-- Counter should be increasing, as this indicates
that communication between VSM is working properly.
  rx_drops_wrong_domain: 0
  rx_drops_wrong_slot: 0
  rx_drops_short_pkt: 0
  rx_drops_queue_full: 0
  rx_drops_inactive_cp: 0
  rx_drops_bad_src: 0
  rx_drops_not_ready: 0
  rx_unknown_pkts: 0
Redun Device 1: <-- This device maps to the mgmt interface
  name: ha1
  pdev: ad7b6860
  alarm: true
  mac: ff:ff:ff:ff:ff:ff
  tx_set_ver_req_pkts: 11589
  tx_set_ver_rsp_pkts: 0
  tx_heartbeat_req_pkts: 12
  tx_heartbeat_rsp_pkts: 0
  rx_set_ver_req_pkts: 0
  rx_set_ver_rsp_pkts: 0
  rx_heartbeat_req_pkts: 0
  rx_heartbeat_rsp_pkts: 0 <-- When communication between VSM through the control
interface is interrupted but continues through the mgmt interface, the
rx_heartbeat_rsp_pkts will increase.
  rx_drops_wrong_domain: 0
  rx_drops_wrong_slot: 0
  rx_drops_short_pkt: 0
  rx_drops_queue_full: 0
  rx_drops_inactive_cp: 0
  rx_drops_bad_src: 0
  rx_drops_not_ready: 0
  rx_unknown_pkts: 0

```

システム内部の `sysmgr` ステータスをチェックするには、次のコマンドを使用します。

- **show system internal sysmgr state**

```
N1000V# show system internal sysmgr state
```

```

The master System Manager has PID 1988 and UUID 0x1.
Last time System Manager was gracefully shutdown.
The state is SRV_STATE_MASTER_ACTIVE_HOTSTDBY entered at time Tue Apr 28 13:09:13
2009.

```

```
The '-b' option (disable heartbeat) is currently disabled.
```

```
The '-n' (don't use rlimit) option is currently disabled.
```

```
Hap-reset is currently enabled.
```

```
Watchdog checking is currently disabled.
```

```
Watchdog kgdb setting is currently enabled.
```

```
Debugging info:
```

```
The trace mask is 0x00000000, the syslog priority enabled is 3.
The '-d' option is currently disabled.
The statistics generation is currently enabled.
```

```
HA info:
```

```
slotid = 1      supid = 0
cardstate = SYSMGR_CARDSTATE_ACTIVE .
cardstate = SYSMGR_CARDSTATE_ACTIVE (hot switchover is configured enabled).
Configured to use the real platform manager.
Configured to use the real redundancy driver.
Redundancy register: this_sup = RDN_ST_AC, other_sup = RDN_ST_SB.
EOBC device name: eth0.
Remote addresses:  MTS - 0x00000201/3      IP - 127.1.1.2
MSYNC done.
Remote MSYNC not done.
Module online notification received.
Local super-state is: SYSMGR_SUPERSTATE_STABLE
Standby super-state is: SYSMGR_SUPERSTATE_STABLE
Swover Reason : SYSMGR_SUP_REMOVED_SWOVER <-- Reason for the last switchover
Total number of Switchovers: 0 <-- Number of switchovers
>> Duration of the switchover would be listed, if any.
```

```
Statistics:
```

```
Message count:          0
Total latency:          0      Max latency:          0
Total exec:             0      Max exec:          0
```

モジュールを再ロードするには、次のコマンドを使用します。

- **reload module**

```
n1000V# reload module 2
```

このコマンドは、セカンダリ VSM を再ロードします。



(注) モジュールを指定せずに **reload** コマンドを実行すると、システム全体が再ロードされます。

スタンバイ VSM コンソールに接続するには、次のコマンドを使用します。

- **attach module**

スタンバイ VSM コンソールには、外部からはアクセスできません。アクセスするには、アクティブ VSM から **attach module module-number** コマンドを使用します。

```
n1000V# attach module 2
```

このコマンドは、セカンダリ VSM のコンソールに接続します。