

中断時の vPC の動作における FlexPod Nexus 5k

内容

[概要](#)

[トポロジ](#)

[テスト](#)

[データ リンク損失](#)

[アップグレードまたはリロードの中断](#)

[vPC ピア リンクがダウンする](#)

[In-Service Software Upgrade \(ISSU \)](#)

[ISSU での既知の問題](#)

[ピア キープアライブ リンクの損失](#)

[vPC 機能を無効にする](#)

[結論](#)

[関連するシスコ サポート コミュニティ ディスカッション](#)

概要

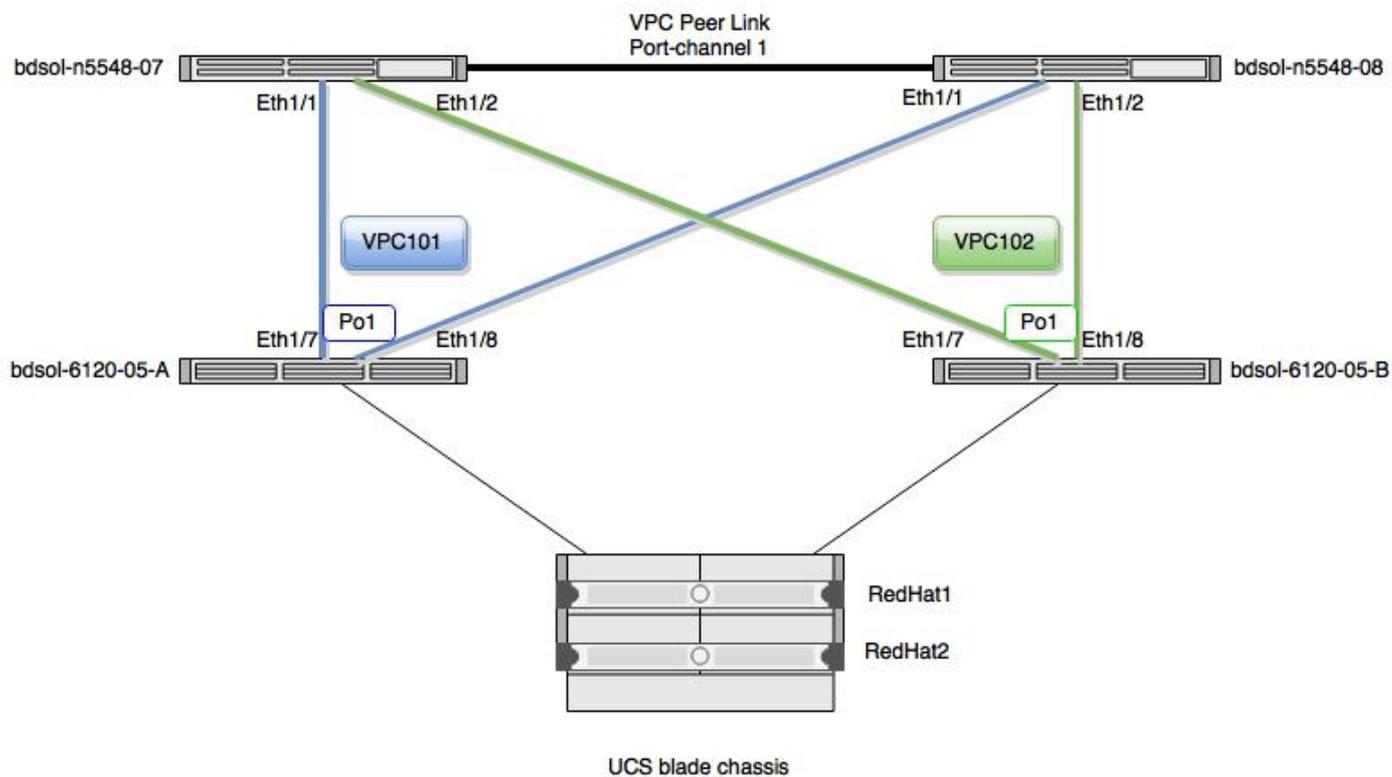
このドキュメントでは、仮想ポート チャンネル (vPC) を中心として、さまざまな中断に対する反応でのネットワークのオンライン動作を示します。

典型的な中断としては、リロード、リンク損失、または接続損失があります。

このドキュメントの目的は、一般的なシナリオでのパケット損失を説明することです。

トポロジ

別途記載がない限り、テスト中、次のトポロジが使用されます。



緑と青の線は、各ファブリック インターコネクトから両方の Nexus スイッチへの vPC ポート チャネルを示しています。

アウト オブ バンド管理ネットワークについての概要は示されていません。

これは、以下のような FlexPod の配備で一般に推奨されるシンプルなトポロジです。

http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpod_esxi51_ucsm2.html

使用するコンポーネント

2 つの Nexus 5548P スイッチ。

2.2(4b) ソフトウェアを実行している 2 つの Unified Computing System (UCS) 6120 ファブリック インターコネクト。

1 つの 5108 UCS シャーシ。

2.2(4) ソフトウェアが実行されている VIC 1240 アダプタを備えた 2 つの B200M3 ブレード。

接続テストを実行および検証するため、2 つのブレードが取り付けられ、RedHat Enterprise Linux 7.1 オペレーティングシステムがインストールされています。

設定

vPC とポート チャネル設定の両方でデフォルトが使用されています。

```
feature vpc
```

```
vpc domain 75
```

```
role priority 3000
peer-keepalive destination 10.48.43.79 source 10.48.43.78
delay restore 150
peer-gateway
```

```
interface port-channel1
description vPC Peer-Link
switchport mode trunk
spanning-tree port type network
vpc peer-link
```

UCS ファブリック インターコネク ト (FI) に至る vPC の例 (この例では bdsol-6120-05--A)

```
interface port-channel101
description bdsol-6120-05-A
switchport mode trunk
spanning-tree port type edge trunk
vpc 101
```

テスト

次のテストが実行されます。

- データ リンク損失
- アップグレード中断

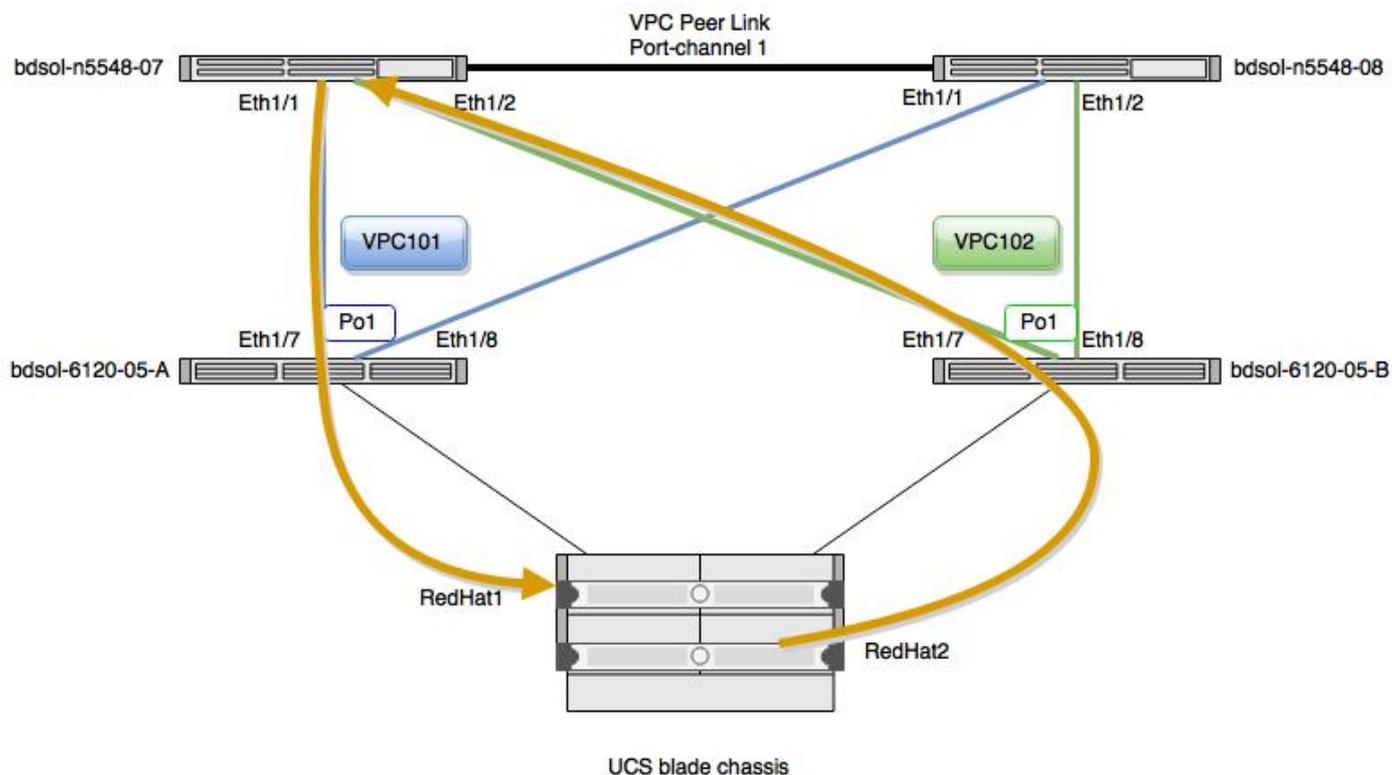
In-Service Software Upgrade (ISSU)

- ピア キープアライブ リンクの損失 - このトポロジ/設定では mgmt0 インターフェイス。
- ピア ポートチャネルの損失 - この設定のポート チャネル 1。
- vPC 機能を無効にする

基本的なトラフィック フロー。

単一 iperf3 セッションを使用することにより、移行時にフレーム損失を確認するための 6.5 ギガビット/秒のテスト TCP トラフィックが生成されます。

RedHat2 はファブリック インターコネク ト B に固定されており、RedHat1 はファブリック インターコネク ト A に固定されています。その結果、スイッチング部分を通る必要のあるトラフィックが発生します。



Iperf3 のパラメータ :

- [Server] : `iperf3 -s -i 1`
- クライアント : `iperf3 -c 10.37.9.131 -t 0 -i 1 -w 1M -V`

高速トラフィックを可能にし、パケット損失を特定しやすくするため、上記のパラメータが選択されています。

TCP ウィンドウは、iperf でよく発生するデータ バーストを回避するために固定されています。iperf による固定解除の実行を許可すると、QoS の設定によっては、パス上の入力バッファでデータが時々失われることがあります。上記のパラメータは、フレーム損失のない 6 ~ 7 Gbps の持続レートを有効にします。

これを確認するため、インターフェイス上のトラフィックの累積率を確認できます。

```
bdsol-n5548-07# show interface ethernet 1/1-2 | i rate
30 seconds input rate 5612504 bits/sec, 9473 packets/sec
30 seconds output rate 7037817832 bits/sec, 578016 packets/sec
input rate 5.60 Mbps, 9.38 Kpps; output rate 7.01 Gbps, 576.10 Kpps
30 seconds input rate 7037805336 bits/sec, 578001 packets/sec
30 seconds output rate 5626064 bits/sec, 9489 packets/sec
input rate 7.01 Gbps, 575.71 Kpps; output rate 6.56 Mbps, 9.79 Kpps
```

上記の出力では、インターフェイス Ethernet 1/2 には 7 Gbps のトラフィックが到着し、インターフェイス Ethernet 1/1 では出力されていることが示されています。

データ リンク損失

このテストは、vPC を構成するリンクがシャットダウンされると、データがどのように動作するかをテストするように指定されています。

この例では、データトラフィック用の出力インターフェイスである Ethernet 1/1 が使用され、コマンドラインを使用してシャットダウンされます。

```
bdsol-n5548-07# conf t
Enter configuration commands, one per line. End with CNTL/Z.
bdsol-n5548-07(config)# int et1/1
bdsol-n5548-07(config-if)# shut
```

この場合、大量の 6.5 Gbps ストリームのうち、1つのパケットのみが失われています。

[41	2025.01-2026.01	sec	800	MBytes	6.71	Gbits/sec	0	272	KBytes
[41	2026.01-2027.01	sec	800	MBytes	6.71	Gbits/sec	0	272	KBytes
[41	2027.01-2028.01	sec	801	MBytes	6.72	Gbits/sec	0	272	KBytes
[41	2028.01-2029.01	sec	798	MBytes	6.69	Gbits/sec	0	272	KBytes
[41	2029.01-2030.01	sec	799	MBytes	6.70	Gbits/sec	0	272	KBytes
[41	2030.01-2031.01	sec	799	MBytes	6.71	Gbits/sec	0	272	KBytes
[41	2031.01-2032.01	sec	799	MBytes	6.70	Gbits/sec	0	272	KBytes
[41	2032.01-2033.01	sec	799	MBytes	6.70	Gbits/sec	0	272	KBytes
[41	2033.01-2034.01	sec	800	MBytes	6.71	Gbits/sec	0	272	KBytes
[41	2034.01-2035.01	sec	799	MBytes	6.70	Gbits/sec	0	272	KBytes
[41	2035.01-2036.01	sec	799	MBytes	6.70	Gbits/sec	0	272	KBytes
[41	2036.01-2037.01	sec	634	MBytes	5.32	Gbits/sec	1	272	KBytes
[41	2037.01-2038.01	sec	792	MBytes	6.65	Gbits/sec	0	272	KBytes
[41	2038.01-2039.01	sec	795	MBytes	6.67	Gbits/sec	0	272	KBytes

トラフィックは、UCS 上のポートチャネル内の残りのリンクの間ではほぼ即座にバランス調整されます。この場合は、Nexus 5548 B を宛先とする UCS FI B の Ethernet 1/8 の (残っている唯一の) ポートを使用して、そこから Ethernet 1/1 を使用して UCS FI A に転送されます。

```
bdsol-n5548-08# show interface ethernet 1/1-2 | i rate
30 seconds input rate 5575896 bits/sec, 9413 packets/sec
30 seconds output rate 6995947064 bits/sec, 574567 packets/sec
input rate 2.21 Mbps, 3.70 Kpps; output rate 2.78 Gbps, 227.99 Kpps
30 seconds input rate 6995940736 bits/sec, 574562 packets/sec
30 seconds output rate 5581920 bits/sec, 9418 packets/sec
input rate 2.78 Gbps, 227.99 Kpps; output rate 2.22 Mbps, 3.71 Kpps
```

アップグレードまたはリロードの中断

データとコントロールプレーン中断の組み合わせは、中断アップグレード bdsol-n5548-07 (プライマリ vPC) を実行することによりエミュレートできます。

トラフィックの損失が予期されます。

このテストは、機能的には、vPC ピアをリロードすることと同じです。

```
bdsol-n5548-07# install all kickstart bootflash:n5000-uk9-kickstart.7.1.0.N1.1a.bin system
bootflash:n5000-uk9.7.1.0.N1.1a.bin
(...)
```

```
Compatibility check is done:
Module bootable Impact Install-type Reason
```

```
-----
1 yes disruptive reset Incompatible image
```

```
(...)
```

```
Switch will be reloaded for disruptive upgrade.
```

Do you want to continue with the installation (y/n)? [n] y

Install is in progress, please wait.

Performing runtime checks.

[#####] 100% -- SUCCESS

Setting boot variables.

[#####] 100% -- SUCCESS

Performing configuration copy.

[#####] 100% -- SUCCESS

Finishing the upgrade, switch will reboot in 10 seconds.

10 秒後に、そのパケット損失が発生します。

その間の損失は 55 個のパケットのみです (6.6 Gbps ストリームのうち)。

```
[ 4] 3571.01-3572.01 sec 800 MBytes 6.71 Gbits/sec 0 272 KBytes
[ 4] 3572.01-3573.01 sec 801 MBytes 6.72 Gbits/sec 0 272 KBytes
[ 4] 3573.01-3574.01 sec 800 MBytes 6.71 Gbits/sec 0 272 KBytes
[ 4] 3574.01-3575.01 sec 638 MBytes 5.35 Gbits/sec 55 211 KBytes
[ 4] 3575.01-3576.01 sec 805 MBytes 6.75 Gbits/sec 0 242 KBytes
[ 4] 3576.01-3577.01 sec 801 MBytes 6.72 Gbits/sec 0 260 KBytes
[ 4] 3577.01-3578.01 sec 801 MBytes 6.72 Gbits/sec 0 269 KBytes
[ 4] 3578.01-3579.01 sec 799 MBytes 6.70 Gbits/sec 0 272 KBytes
[ 4] 3579.01-3580.01 sec 797 MBytes 6.68 Gbits/sec 0 272 KBytes
[ 4] 3580.01-3581.01 sec 800 MBytes 6.71 Gbits/sec 0 272 KBytes
^C[ 4] 3581.01-3581.83 sec 656 MBytes 6.71 Gbits/sec 0 272 KBytes
-----
Test Complete. Summary Results:
[ ID] Interval          Transfer          Bandwidth          Retr
[ 4]  0.00-3581.83 sec 3002125194048 bits 6.71 Gbits/sec    55
[ 4]  0.00-3581.83 sec  0.00 Bytes 0.00 bits/sec
CPU Utilization: local/sender 34.6% (0.5%u/34.1%s), remote/receiver 0.0% (0.0%u/0.0%s)
iperf3: interrupt - the client has terminated
```

iperf3 が再起動したらすぐにオペレータは、トラフィックが実際に bdsol-n5548-08 に切り替わっていることを確認できます。

```
bdsol-n5548-08# show interface ethernet 1/1-2 | i rate
30 seconds input rate 5601392 bits/sec, 9455 packets/sec
30 seconds output rate 7015307760 bits/sec, 576159 packets/sec
input rate 2.25 Mbps, 3.77 Kpps; output rate 2.81 Gbps, 231.14 Kpps
30 seconds input rate 7015303696 bits/sec, 576152 packets/sec
30 seconds output rate 5605280 bits/sec, 9462 packets/sec
input rate 2.81 Gbps, 231.14 Kpps; output rate 2.25 Mbps, 3.77 Kpps
```

速度カウンタは 30 秒間についての平均が取られるため、トラフィック速度は 6 Gbps 未満を示しています。

vPC ピア リンクがダウンする

この例では、設定変更をきっかけとして vPC ピア リンクがダウンします。

その時点でトラフィックは、vPC セカンダリとして動作する bdsol-n5548-07 によって処理されています。

一連のイベント。

port-channel 1 がダウンします。

```
2015 Jul 10 15:00:25 bdsol-n5548-07 %ETHPORT-5-IF_DOWN_CFG_CHANGE:Interface port-channel1 is down(Config change)
```

bdsol-n5548-07 がセカンダリとして動作しており、ループレス トポロジを保証できないため、vPC が中断されます。

```
2015 Jul 10 15:00:28 bdsol-n5548-07 %VPC-2-VPC_SUSP_ALL_VPC: Peer-link going down, suspending all vPCs on secondary
2015 Jul 10 15:00:28 bdsol-n5548-07 %ETHPORT-5-IF_DOWN_INITIALIZING: Interface port-channel928 is down (Initializing)
2015 Jul 10 15:00:28 bdsol-n5548-07 %ETHPORT-5-IF_DOWN_INITIALIZING: Interface port-channel102 is down (Initializing)
2015 Jul 10 15:00:28 bdsol-n5548-07 %ETHPORT-5-IF_DOWN_INITIALIZING: Interface port-channel101 is down (Initializing)
```

この間に iperf3 はトラフィックの一部、90 パケットを失いました。

```
[ 41] 5871.01-5872.01 sec 798 MBytes 6.70 Gbits/sec 0 272 KBytes
[ 41] 5872.01-5873.01 sec 798 MBytes 6.70 Gbits/sec 0 272 KBytes
[ 41] 5873.01-5874.01 sec 801 MBytes 6.72 Gbits/sec 0 272 KBytes
[ 41] 5874.01-5875.01 sec 801 MBytes 6.72 Gbits/sec 0 272 KBytes
[ 41] 5875.01-5876.01 sec 800 MBytes 6.71 Gbits/sec 0 272 KBytes
[ 41] 5876.01-5877.01 sec 796 MBytes 6.68 Gbits/sec 0 272 KBytes
[ 41] 5877.01-5878.01 sec 796 MBytes 6.68 Gbits/sec 0 272 KBytes
[ 41] 5878.01-5879.01 sec 599 MBytes 5.03 Gbits/sec 90 272 KBytes
[ 41] 5879.01-5880.01 sec 800 MBytes 6.71 Gbits/sec 0 272 KBytes
[ 41] 5880.01-5881.01 sec 799 MBytes 6.70 Gbits/sec 0 272 KBytes
[ 41] 5881.01-5882.01 sec 798 MBytes 6.70 Gbits/sec 0 272 KBytes
[ 41] 5882.01-5883.01 sec 798 MBytes 6.69 Gbits/sec 0 272 KBytes
[ 41] 5883.01-5884.01 sec 801 MBytes 6.72 Gbits/sec 0 272 KBytes
[ 41] 5884.01-5885.01 sec 800 MBytes 6.71 Gbits/sec 0 272 KBytes
```

しかし、比較的速く回復することができました。

vPC が bdsol-n5548-07 で中断されているため、すべてのトラフィックは bdsol-n5548-08 で処理されます。

```
bdsol-n5548-08# show int ethernet 1/1-2 | i rate
30 seconds input rate 5623248 bits/sec, 9489 packets/sec
30 seconds output rate 7036030160 bits/sec, 577861 packets/sec
input rate 2.83 Mbps, 4.74 Kpps; output rate 3.54 Gbps, 290.64 Kpps
30 seconds input rate 7036025712 bits/sec, 577854 packets/sec
30 seconds output rate 5627216 bits/sec, 9498 packets/sec
input rate 3.54 Gbps, 290.64 Kpps; output rate 2.83 Mbps, 4.75 Kpps
```

この場合も、負荷平均の計算のため、速度として 6.5 ギガビット/秒が即座には表示されません。

vPC リンク ダウンからの回復

vPC ピア リンクが再びアクティブになると、トラフィックは複数リンクの間でバランスが再調整され、トポロジの変更による短期間のパケット損失が予期されます。

このラボテストの場合、1 パケットが失われました。

In-Service Software Upgrade (ISSU)

このテストでは、トラフィックの中断を確認するため、ISSU アップグレードが実行されました。

このテストでの vPC の役割は、次のとおりです。

bdsol-n5548-07 - プライマリ

bdsol-n5548-08 - セカンダリ

ISSU を実行するには、定義されている基準が満たされている必要があります。

これらの基準をチェックし、ISSU を実行するために使用されるコマンドに関する情報を検索するために、次のガイドが使用されました。

http://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus5500/sw/upgrade/705_N1_1/n5500_upgrade_downgrade_700.html#pgfId-727913.

まずプライマリ vPC ピア上、その後にセカンダリ vPC ピア上で ISSU が実行された後、パケットの損失はありませんでした。

これは、ISSU のすべてのデータプレーン機能は中断されないまま残り、コントロールプレーントラフィックのみに影響したという事実によります。

```
[ 4] 1096.01-1097.01 sec 798 MBytes 6.69 Gbits/sec 0 272 KBytes
[ 4] 1097.01-1098.01 sec 800 MBytes 6.71 Gbits/sec 0 272 KBytes
[ 4] 1098.01-1099.01 sec 798 MBytes 6.69 Gbits/sec 0 272 KBytes
[ 4] 1099.01-1100.01 sec 799 MBytes 6.70 Gbits/sec 0 272 KBytes
^C[ 4] 1100.01-1100.71 sec 563 MBytes 6.72 Gbits/sec 0 272 KBytes
```

Test Complete. Summary Results:

```
[ ID] Interval Transfer Bandwidth Retr
[ 4] 0.00-1100.71 sec 856 GBytes 6.68 Gbits/sec 0 sender
[ 4] 0.00-1100.71 sec 0.00 Bytes 0.00 bits/sec receiver
CPU Utilization: local/sender 35.0% (0.5%u/34.5%s), remote/receiver 0.0% (0.0%u/0.0%s)
iperf3: interrupt - the client has terminated
```

ISSU での既知の問題

レイヤ 3 の機能およびライセンス

ISSU テスト中に、解決する必要がある問題がいくつか検出されました。「show install all impact ...」コマンドでは、ISSU を実行できないという出力が、以下の説明とともに示される場合があります。「Non-disruptive install not supported if L3 was enabled.」テスト環境においてこれは、インストールされているライセンス ファイルで使用中の LAN_BASE_SERVICES_PKG が原因でした。

LAN_BASE_SERVICES_PKG には L3 機能が含まれており、ISSU を実行するには、このパッケージは未使用でなければならず、「clear license LICENSEFILE」コマンドを使用することによって、デバイスからライセンス ファイルをクリアする必要があります。現在、ライセンス ファイルがデバイスで使用中になっている可能性があります。そのようなライセンス ファイルをクリアするには、「show license usage」を使用し、それらのパッケージの機能を無効にすることによって、どのパケットが使用中かをチェックすることが重要です。

非エッジ STP ポート

テスト中に、ノースバウンド ポートチャンネルをシャットダウンする必要もありました。これは、「show spanning-tree issu-impact」の非エッジ、基準 3 チェックに合格せず、中断アップグレードの原因となることがあるからです。このノースバウンド ポートチャンネルは、「show spanning-tree vlan 1」コマンドで、vPC エッジではないものとして、リストされていました。

ピア キープアライブ リンクの損失

ピア キープアライブ mgmt0 リンクの損失後は、トラフィックの中断が記録されていません。このトポロジでは、管理インターフェイス (mgmt0) がキープアライブ リンクとして使用されているため、テスト中に生成されるデータトラフィックに影響を与えることはありません。

```
[ 41] 71.01-72.01 sec 793 MBytes 6.65 Gbits/sec 0 272 KBytes
[ 41] 72.01-73.01 sec 794 MBytes 6.66 Gbits/sec 0 272 KBytes
[ 41] 73.01-74.01 sec 791 MBytes 6.63 Gbits/sec 0 272 KBytes
[ 41] 74.01-75.01 sec 793 MBytes 6.65 Gbits/sec 0 272 KBytes
[ 41] 75.01-76.01 sec 793 MBytes 6.65 Gbits/sec 0 272 KBytes
^C [ 41] 76.01-76.02 sec 8.13 MBytes 5.83 Gbits/sec 0 272 KBytes
-----
Test Complete. Summary Results:
[ ID] Interval Transfer Bandwidth Retr
[ 41] 0.00-76.02 sec 58.6 GBytes 6.62 Gbits/sec 0 sender
[ 41] 0.00-76.02 sec 0.00 Bytes 0.00 bits/sec receiver
CPU Utilization: local/sender 35.2% (0.5%/34.7%/s), remote/receiver 0.0% (0.0%/0.0%/s)
iperf3: interrupt - the client has terminated
```

デバイスは、mgmt0 インターフェイスがダウンし、ピア キープアライブに失敗したことを認識しますが、ピア リンクはアップ状態なので、データ場所通信は続行できます。

```
2015 Jul 14 12:11:28 bdsol-n5548-07 %IM-5-IM_INTF_STATE: mgmt0 is DOWN in vdc 1
2015 Jul 14 12:11:32 bdsol-n5548-07 %VPC-2-PEER_KEEP_ALIVE_RECV_FAIL: In domain 75, VPC peer
keep-alive receive has failed
2015 Jul 14 12:12:07 bdsol-n5548-07 %IM-5-IM_INTF_STATE: mgmt0 is UP in vdc 1
```

vPC 機能を無効にする

このテストでは、ライブ データ転送中にスイッチの 1 つで vPC を無効にするとどうなるかについて説明します。

VPC 機能は、グローバル設定モードで次のコマンドを使用することにより、無効にすることができます。

```
bdsol-n5548-07(config)# no feature vpc
```

プライマリまたはセカンダリのいずれかの vPC ピア上で vPC 機能を無効にすると、データ接続が即座に失われます。これは、vPC のピアベースの性質によるものです。機能が無効になると、すぐに、スイッチ上の vPC 機能すべてが機能を停止し、ピア リンクがダウンし、vPC キープアライブ ステータスが [Suspended] になり、テスト環境のポートチャンネル 101 がダウンします。これは、まだ vPC 機能が有効なピア スイッチの show vPC の出力で明らかです。

```
bdsol-n5548-07# show vpc
```

Legend:

(*) - local vPC is down, forwarding via vPC peer-link

```
vPC domain id          : 75
Peer status            : peer link is down
vPC keep-alive status  : Suspended (Destination IP not reachable)
...
```

vPC status

```
-----
id Port Status Consistency Reason Active vlans
-----
```

```
101 Po101 down success success -
```

前と同じように、トラフィックの中断は短期間だけです。

上記のテスト条件で、単一セッションで 50 ~ 80 パケットが失われました。

「remove feature vpc」コマンドも、vPC 設定がポート チャンネルから削除される原因になっていました。

この設定を再度追加する必要があります。

結論

vPC 機能は、ポート チャンネル内のデータ トラフィックを複数のデバイスの間に分割することにより、復元パフォーマンスをもたらすように設計されています。

このシンプルなアイデアでは、複雑なコントロール プレーンの実装が必要です。

上記のテストは、機能のライフサイクルにおいて発生する可能性のあるコントロール プレーンおよびデータ プレーンの両方の中断を示すことを目的としたものでした。

予期されたとおり、データ プレーンの中断は、ほとんど即座に検出され、修正されました。テストでの損失は単一パケットでした。

コントロール プレーン中断のテストは、コントロール プレーンが影響を受けている間も、vPC がまだサブセカンド コンバージェンスの時間を維持することを示していました。

実行された中断テストとして最も破壊的だったのは、vPC ピア リンクのシャットダウンであり、データ プレーンとコントロール プレーンの両方の障害の組み合わせの結果である可能性があります。それでも、高速のコンバージェンス時間が示されました。