

# 在中断期间，vPC中的FlexPod Nexus 5k行为

## 目录

[简介](#)

[拓扑](#)

[测试](#)

[数据链路丢失](#)

[中断升级或重新加载](#)

[vPC对等链路断开](#)

[服务中软件升级\(ISSU\)](#)

[ISSU的已知问题](#)

[对等保持连接链路丢失](#)

[禁用vPC功能](#)

[结论](#)

[相关的思科支持社区讨论](#)

## 简介

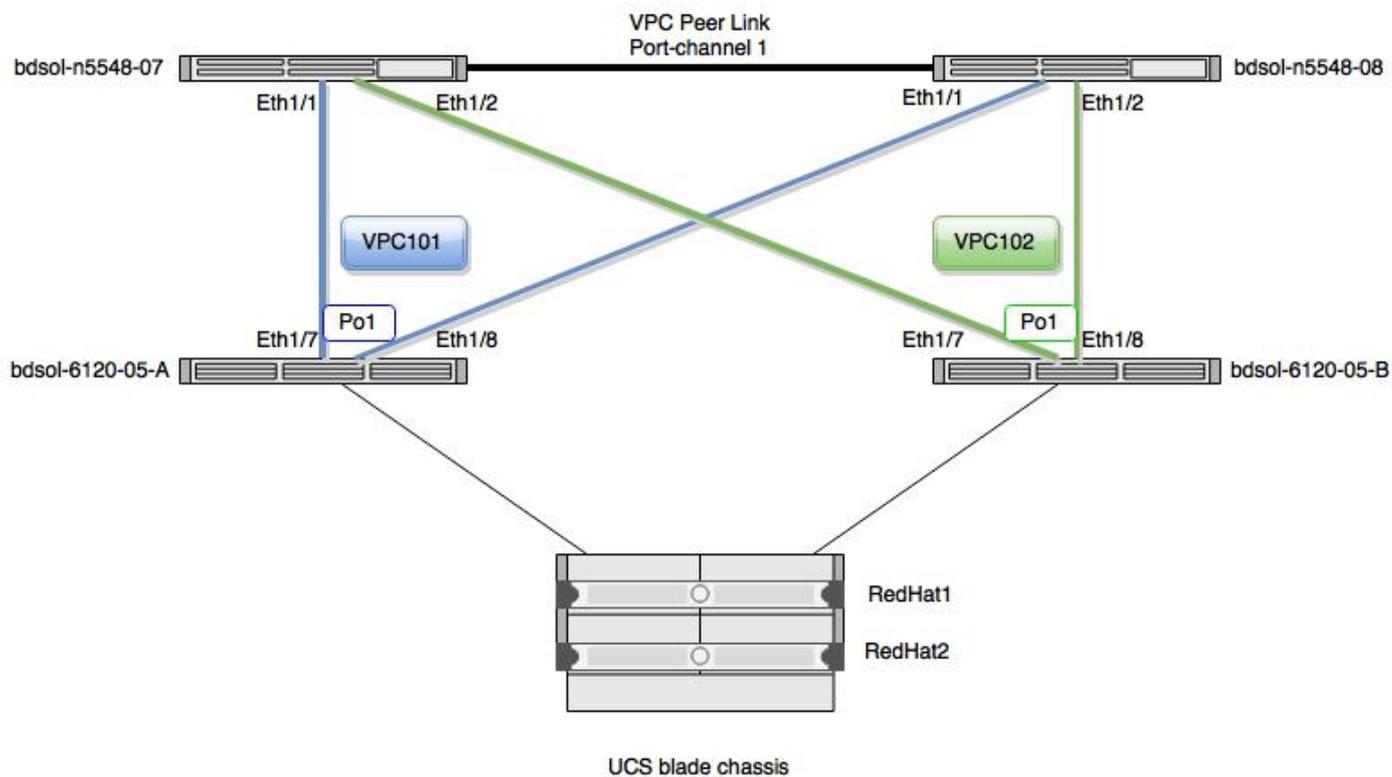
本文档将概述网络针对不同中断的行为，重点介绍虚拟端口通道(vPC)。

典型的中断是重新加载、链路丢失或连接丢失。

本文档旨在演示常见场景中的丢包情况。

## 拓扑

在测试期间，除非另有说明，否则使用以下拓扑。



绿线和蓝线表示从每个交换矩阵互联到两个Nexus交换机的vPC端口通道。

带外管理网络未概述。

它是FlexPod部署中通常建议的简化拓扑，例如：

[http://www.cisco.com/c/en/us/td/docs/unified\\_computing/ucs/UCS\\_CVDs/flexpod\\_esxi51\\_ucsm2.html](http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpod_esxi51_ucsm2.html)

## 使用的组件

两台Nexus 5548P交换机。

运行2.2(4b)软件的两个统一计算系统(UCS)6120交换矩阵互联。

一个5108 UCS机箱。

两个B200M3刀片，带运行2.2(4)软件的VIC 1240适配器。

要执行和验证连接测试，已安装两个刀片，并安装了RedHat Enterprise Linux 7.1操作系统。

## 配置.

vPC和端口通道配置均使用默认配置。

```
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
```

本例中，vPC示例导向UCS交换矩阵互联(FI)bdsol-6120-05 - A

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

## 测试

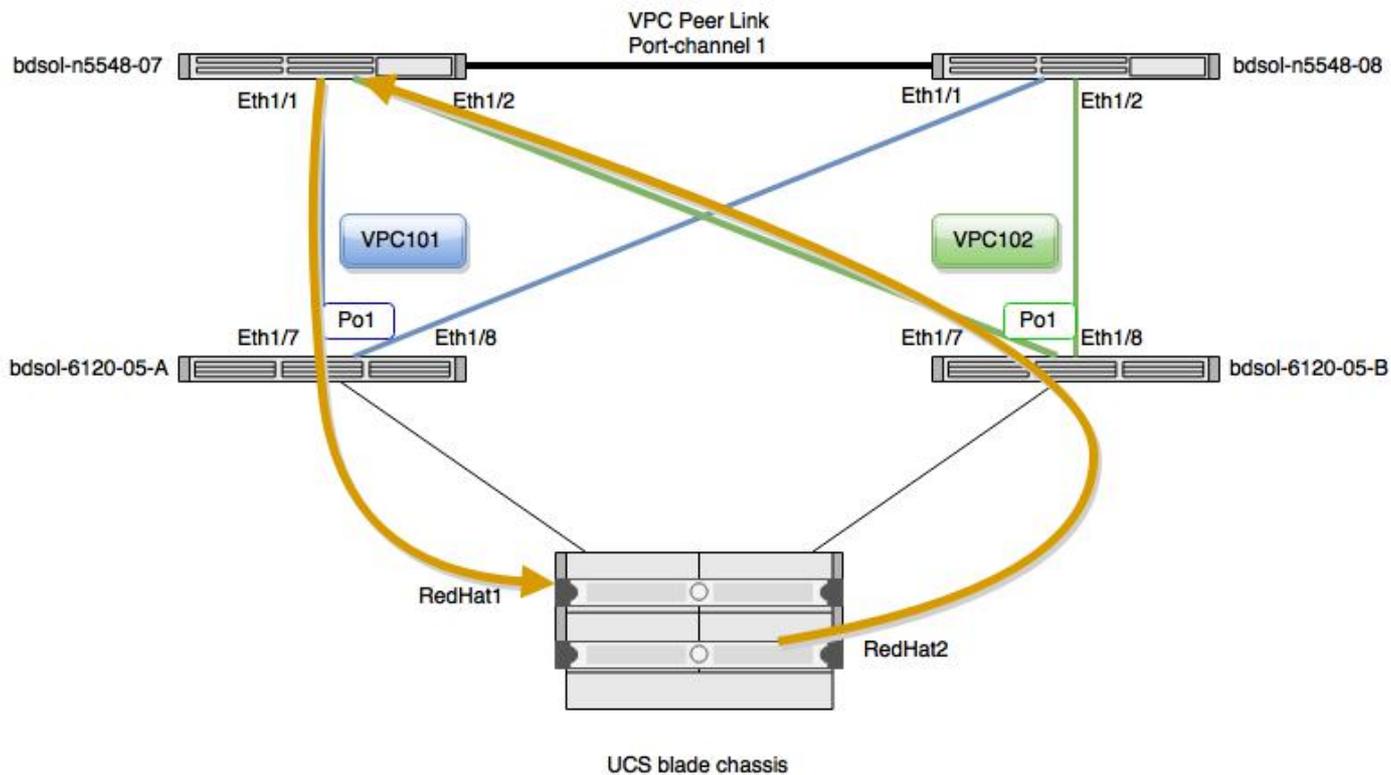
将执行以下测试。

- 数据链路丢失。
- 中断性升级
- 服务中软件升级(ISSU)
- 对等保持连接链路丢失 — 如果出现此拓扑/配置，则为mgmt0接口。
- 丢失对等端口通道 — 此配置中的端口通道1。
- 禁用vPC功能

### 基本流量。

单个iperf3会话用于生成每秒6.5千兆的测试TCP流量，以验证转换期间的帧丢失。

RedHat2固定到交换矩阵互联B，而RedHat1固定到交换矩阵互联A — 这会导致需要通过交换部分的流量。



Iperf3参数：

- 服务器：`iperf3 -s -i 1`
- 客户端：`iperf3 -c 10.37.9.131 -t 0 -i 1 -w 1M -V`

选择上述参数是为了允许高流量速率且容易发现数据包丢失。

TCP窗口被钳制以避免数据突发。允许iperf运行未钳位会导致沿路径的入口缓冲区偶尔发生丢弃——具体取决于QoS配置。以上参数允许6-7 Gbps的持续速率，而不会丢帧。

为了验证我们可以检查接口上的累积流量速率。

```
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
```

以上输出显示7 Gbps的流量进入接口Ethernet 1/2，离开接口Ethernet 1/1。

## 数据链路丢失

此测试用于测试当属于vPC的链路关闭时数据的行为。

本示例将使用Ethernet 1/1（数据流量的输出接口），它将使用命令行关闭。

```
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
```

在这种情况下，在6.5 Gbps流泛洪中，仅丢失一个数据包。

[ 4]	2025.01-2026.01	sec	800	MBytes	6.71	Gbits/sec	0	272	KBytes
[ 4]	2026.01-2027.01	sec	800	MBytes	6.71	Gbits/sec	0	272	KBytes
[ 4]	2027.01-2028.01	sec	801	MBytes	6.72	Gbits/sec	0	272	KBytes
[ 4]	2028.01-2029.01	sec	798	MBytes	6.69	Gbits/sec	0	272	KBytes
[ 4]	2029.01-2030.01	sec	799	MBytes	6.70	Gbits/sec	0	272	KBytes
[ 4]	2030.01-2031.01	sec	799	MBytes	6.71	Gbits/sec	0	272	KBytes
[ 4]	2031.01-2032.01	sec	799	MBytes	6.70	Gbits/sec	0	272	KBytes
[ 4]	2032.01-2033.01	sec	799	MBytes	6.70	Gbits/sec	0	272	KBytes
[ 4]	2033.01-2034.01	sec	800	MBytes	6.71	Gbits/sec	0	272	KBytes
[ 4]	2034.01-2035.01	sec	799	MBytes	6.70	Gbits/sec	0	272	KBytes
[ 4]	2035.01-2036.01	sec	799	MBytes	6.70	Gbits/sec	0	272	KBytes
[ 4]	2036.01-2037.01	sec	634	MBytes	5.32	Gbits/sec	1	272	KBytes
[ 4]	2037.01-2038.01	sec	792	MBytes	6.65	Gbits/sec	0	272	KBytes
[ 4]	2038.01-2039.01	sec	795	MBytes	6.67	Gbits/sec	0	272	KBytes

流量在UCS上端口通道中的剩余链路之间几乎立即平衡，在本例中，使用UCS FI B的Ethernet 1/8 ( 唯一剩余的 ) 端口升级到Nexus 5548 B，从那里将使用Ethernet 1/1传输到UCS FI A。

```
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
```

## 中断升级或重新加载

通过执行bdsol-n5548-07 ( 主vPC ) 的中断升级，可以模拟数据和控制平面的组合中断。

预计会丢失流量。

在功能上，此测试与重新加载vPC对等体相同。

```
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.  
在10秒后，数据包丢失。

在此期间，仅丢失55个数据包（从6.6 Gbps的流中）。

```
[ 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
```

如果iperf3立即重新启动，则操作员可以验证流量确实已切换到bdsol-n5548-08。

```
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
流量速率显示低于6Gbps，因为速率计数器的平均值为30秒。
```

## vPC对等链路断开

在本示例中，vPC对等链路关闭，由配置更改触发。

此时，流量由bdsol-n5548-07（代理vPC辅助）处理。

事件的顺序。

端口通道1关闭。

2015年7月10日15:00:25 bdsol-n5548-07 %ETHPORT-5-IF\_DOWN\_CFG\_CHANGE:接口port-channel1关闭（配置更改）

由于bdsol-n5548-07是辅助，它将暂停其vPC，因为它无法保证无损拓扑：

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)
```

在此期间，iperf3丢失了一部分流量 — 90个数据包。

```
[ 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
```

但恢复得相当快。

由于vPC在bdsol-n5548-07上挂起，所有流量都由bdsol-n5548-08处理

```
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
```

同样，由于计算了负载平均值，该速率不会立即显示6.5千兆位/秒。

从vPC链路恢复为关闭状态。

当vPC对等链路恢复活动状态时，链路之间的流量可能会重新平衡，并且由于拓扑更改而导致的短时数据包丢失可能是预期的。

在本实验中，测试1数据包丢失。

## 服务中软件升级(ISSU)

在本测试中，执行ISSU升级以验证流量中断。

本测试期间的vPC角色如下：

bdsol-n5548-07 — 主

bdsol-n5548-08 — 辅助。

要执行ISSU定义的条件，必须满足。

为了查找有关用于检查这些条件并执行ISSU的命令的信息，使用了以下指南：

[http://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus5500/sw/upgrade/705\\_N1\\_1/n5500\\_upgrade\\_downgrade\\_700.html#pgfId-727913](http://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus5500/sw/upgrade/705_N1_1/n5500_upgrade_downgrade_700.html#pgfId-727913)。

先在主设备上执行ISSU，然后在辅助vPC对等设备上执行ISSU后，没有丢失任何数据包。

这是因为ISSU所有数据平面功能保持不中断，并且仅控制平面流量会受到影响。

```
[ 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
[ 4]  0.00-1100.71 sec  0.00 Bytes  0.00 bits/sec
CPU Utilization: local/sender 35.0% (0.5%u/34.5%u), remote/receiver 0.0% (0.0%u/0.0%u)
iperf3: interrupt - the client has terminated
```

## ISSU的已知问题

### 第3层功能和许可证。

在ISSU测试期间，需要解决许多问题。“show install all impact ...”命令可能提供ISSU无法执行的输出，说明如下：“如果启用了L3，则不支持无中断安装。”在测试环境中，这是由于LAN\_BASE\_SERVICES\_PKG正在安装的许可证文件中使用。

LAN\_BASE\_SERVICES\_PKG包含L3功能，为了执行ISSU，此软件包必须未使用，并且必须使用“clear license LICENSEFILE”命令从设备清除许可证文件。设备当前可能正在使用许可证文件。要清除此类许可证文件，请务必使用“show license usage”并禁用这些包的功能，以检查正在使用的包。

### 非边缘STP端口

在测试期间，还需要关闭北向端口通道，因为它未通过“show spanning-tree issu-impact”非边缘、标准3、检查，这会导致中断性升级。此北向端口通道未在“show spanning-tree vlan 1”命令中列为vPC边缘。

### 对等保持连接链路丢失

在对等保持连接mgmt0链路丢失后，未记录流量中断。在此拓扑中，管理接口(mgmt0)用作保持连接链路，因此不会影响测试期间生成的数据流量。

```
[ 4] 71.01-72.01 sec 793 MBytes 6.65 Gbits/sec 0 272 KBytes
[ 4] 72.01-73.01 sec 794 MBytes 6.66 Gbits/sec 0 272 KBytes
[ 4] 73.01-74.01 sec 791 MBytes 6.63 Gbits/sec 0 272 KBytes
[ 4] 74.01-75.01 sec 793 MBytes 6.65 Gbits/sec 0 272 KBytes
[ 4] 75.01-76.01 sec 793 MBytes 6.65 Gbits/sec 0 272 KBytes
^C[ 4] 76.01-76.02 sec 8.13 MBytes 5.83 Gbits/sec 0 272 KBytes
-----
Test Complete. Summary Results:
[ ID] Interval          Transfer      Bandwidth      Retr
[ 4]  0.00-76.02 sec 58.6 GBytes 6.62 Gbits/sec 0          sender
[ 4]  0.00-76.02 sec 0.00 Bytes 0.00 bits/sec          receiver
CPU Utilization: local/sender 35.2% (0.5%u/34.7%u/s), remote/receiver 0.0% (0.0%u/0.0%u/s)
iperf3: interrupt - the client has terminated
```

设备会注意到mgmt0接口关闭，对等保持连接失败，但是由于对等链路处于打开状态，因此数据位置通信可以继续。

```
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
```

## 禁用vPC功能

本测试将描述在实时数据传输期间，当其中一个交换机上禁用vPC时会发生什么情况。

可在全局配置模式下使用以下命令禁用VPC功能：

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

在主vPC对等体或辅助vPC对等体上禁用vPC功能会导致数据连接立即丢失。这是由于vPC基于对等体的性质。禁用该功能后，交换机上的所有vPC功能即停止运行，对等链路断开，vPC保持连接状态为“暂停”，测试环境的端口通道101断开。这在仍启用vPC功能的对等交换机的show vPC输出中显而易见。

```
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 -
```

与以前一样，流量中断只是短暂的。

在上述测试条件下，单个会话丢失了50-80个数据包。

删除“feature vpc”命令也导致vPC配置从端口通道中删除。

此配置需要重新读取。

## 结论

vPC功能旨在通过在多个设备之间的端口通道中拆分数据流量来实现恢复能力性能。

这个简单的想法需要复杂的控制平面实施。

上述测试旨在显示在功能的生命周期中可能发生的控制平面和数据平面中断。

由于测试中丢失了单个数据包，因此几乎可以立即检测并纠正预期的数据平面中断。

测试的控制平面中断表明，即使控制平面受到影响，vPC仍保持亚秒级收敛时间。

执行的最具破坏性的测试 — vPC对等链路关闭 — 可能会合并数据和控制平面故障。实验表明，该算法收敛速度较快。