# Contents

Introduction

This document will outline behavior of the network in reaction to different disruptions, concentrating on Virtual Port-Channel (vPC).

A typical disruption would be - a reload, link loss, or connecivity loss.

The aim of this document is to demonstrate packet loss during common scenarios.

Topology

During testing, unless otherwise stated following topology is used.
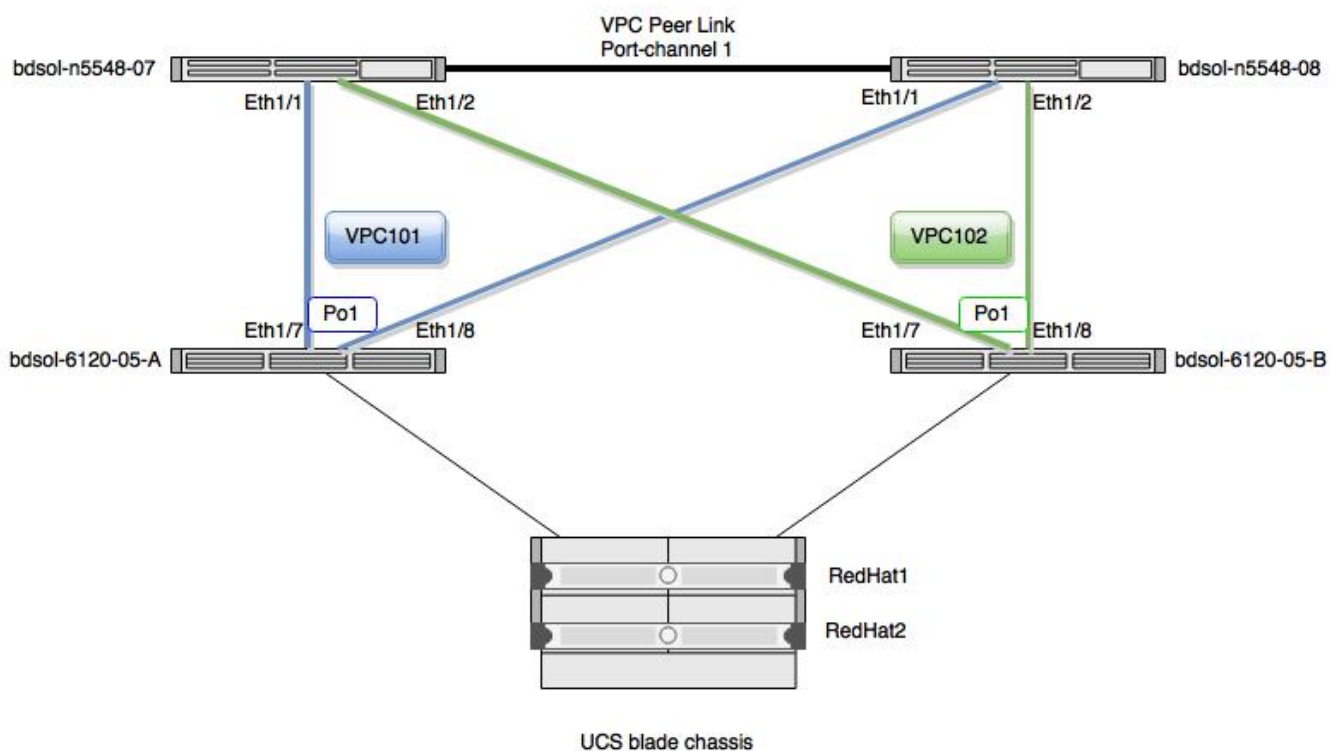
Green and blue lines indicate a vPC port channel from each of the Fabric Interconnects to both Nexus switches.

Not outlined is the out of band management network.

It is a simplified topology commonly recommended in FlexPod deployments as seen for example in:

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

Components used

Two Nexus 5548P switches.

Two Unified Computing System (UCS) 6120 Fabric Interconnect running 2.2(4b) software.

One 5108 UCS chassis.

Two B200M3 blades with VIC 1240 adapter running 2.2(4) software.

To perform and verify connectivity tests two blades were installed and RedHat Enterprise Linux 7.1 operating system is installed.

Configuration.

Both the vPC and portchannel configuration is using default.

Example vPC leading to UCS Fabirc Interconnect (FI) in this case bdsol-6120-05--A

Tests

Following test will be performed.

- Data link loss.

- Disruptive upgrade

- In-Service Software Upgrade (ISSU)

- Loss of peer keepalive link - mgmt0 interface in case of this topology/configuration.

- Loss of peer portchannel - Port-channel 1 in this configuration.

- Disabling vPC feature

Basic traffic flow.

A single iperf3 session is used to generate 6.5 gigabits per second of test TCP traffic to verify frame loss during transitions.

RedHat2 is pinned to Fabric Interconnect B while RedHat1 is pinned to fabric interconnect A - this

results in traffic which needs to cross the switching portion.

UCS blade chassis
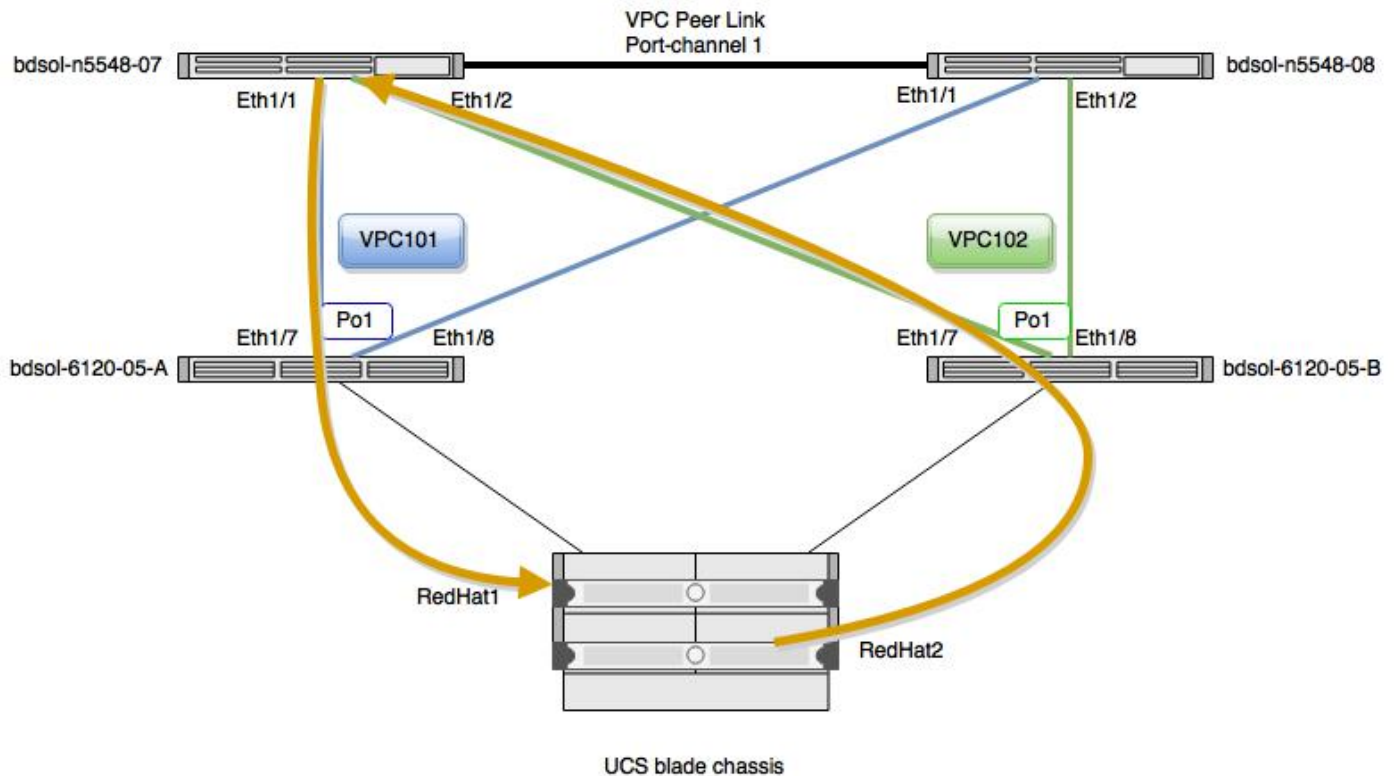
Iperf3 parameters:

- Server: *iperf3 -s -i 1*
- Client *iperf3 -c 10.37.9.131 -t 0 -i 1 -w 1M -V*

The above parameters were picked to allow high rate of traffic and easy to spot packet loss.

TCP window is clamped to avoid data bursts iperf is know for. Allowing iperf to run unclamped could result in occasional drops in ingress buffers along the path - depending on QoS configuration. The above parameters allow for a sustained rate of 6-7 Gbps without frame loss.

To verify we can check cumulative rate of traffic on interfaces.

```
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
```

The above output shows 7 Gbps of traffic entering on interface Ethernet 1/2 and leaving on interface Ethernet 1/1.

Data link loss

This test is designated to test how data will behave if a link which is part of vPC is shut down.

This example will use Ethernet 1/1, the output interface for data traffic, it will be shut down using command line.

```
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
```

In this case only a single packet was lost, out of flood of 6.5 Gbps stream.

Traffic is almost immediately balanced among remaining links in portchannel on UCS, in this case using UCS FI B's Ethernet 1/8 (the only remaining) port going up to Nexus 5548 B, from there it will be transported to UCS FI A using Ethernet 1/1.

```
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
```

Disruptive upgrade or reload

A combined data and control plane disruption can be emulated by performing a disruptive upgrade the bdsol-n5548-07 (primary vPC).

Traffic loss is expected.

Functionally this test is the same as reloading a vPC peer.

```
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 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 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 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.
```
After the 10 seconds mentioned packet loss occurs.

During that time only 55 packets are lost (out of the stream of 6.6 Gbps).

If the iperf3 was restarted immediately the operator can verify that traffic indeed switched over to
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
```
The traffic rate is showing below 6Gbps since the rate counter is averaged over 30 seconds.

vPC peer link going down

In this example the vPC peer link goes down, triggered by a configuration change.

At that time traffic is handled by bdsol-n5548-07, acting vPC secondary.

The sequence of events.

Port-channel 1 goes down.

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

Since bdsol-n5548-07 is acting secondary it will suspend its vPCs since it cannot guarantee loopless topology:

```
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)
```
During this time iperf3 lost a portion of traffic - 90 packets.

```
[  4] 5871.01-5872.01 sec    798 MBytes   6.70 Gbits/sec    0    272 KBytes
[  4] 5872.01-5873.01 sec    798 MBytes   6.70 Gbits/sec    0    272 KBytes
[  4] 5873.01-5874.01 sec    801 MBytes   6.72 Gbits/sec    0    272 KBytes
[  4] 5874.01-5875.01 sec    801 MBytes   6.72 Gbits/sec    0    272 KBytes
[  4] 5875.01-5876.01 sec    800 MBytes   6.71 Gbits/sec    0    272 KBytes
[  4] 5876.01-5877.01 sec    796 MBytes   6.68 Gbits/sec    0    272 KBytes
[  4] 5877.01-5878.01 sec    796 MBytes   6.68 Gbits/sec    0    272 KBytes
[  4] 5878.01-5879.01 sec    599 MBytes   5.03 Gbits/sec    90   272 KBytes
[  4] 5879.01-5880.01 sec    800 MBytes   6.71 Gbits/sec    0    272 KBytes
[  4] 5880.01-5881.01 sec    799 MBytes   6.70 Gbits/sec    0    272 KBytes
[  4] 5881.01-5882.01 sec    798 MBytes   6.70 Gbits/sec    0    272 KBytes
[  4] 5882.01-5883.01 sec    798 MBytes   6.69 Gbits/sec    0    272 KBytes
[  4] 5883.01-5884.01 sec    801 MBytes   6.72 Gbits/sec    0    272 KBytes
[  4] 5884.01-5885.01 sec    800 MBytes   6.71 Gbits/sec    0    272 KBytes
```

But was able to recover pretty fast.

Since vPCs are suspended on bdsol-n5548-07, all traffic is handled by bdsol-n5548-08

```
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
```
Again, the rate does not show 6.5 gigabits per second immediately due to load average being calculated.

Recovery from vPC link down.

When vPC peer link comes back alive, traffic may be re-balanced between links and a short lived packet loss due to topology change may be expected.

In case of this lab test 1 packet was lost.

In-Service Software Upgrade (ISSU)

In this test an ISSU upgrade was performed in order to verify traffic disruption.

The vPC roles during this test are as follows:
bdsol-n5548-07 - primary

bdsol-n5548-08 - secondary.

To perform an ISSU defined criteria must be met.

In order to find information regarding commands used to check these criteria and perform an ISSU the following guide was used:

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

After performing an ISSU first on the primary and afterwards on the secondary vPC peer no packets have been lost.

This is due to the fact that ISSU all data plane functionality remains undisrupted and only control plane traffic would be affected.

```
[  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
```

Known Issues with ISSU

Layer 3 features and licenses.

During the ISSU testing a number of issues needed to be resolved. The "show install all impact ..." command may provide output that ISSU cannot be performed with the following explanation: "Non-disruptive install not supported if L3 was enabled." In the testing environment this was due to the LAN_BASE_SERVICES_PKG being in use in the installed license file.

LAN_BASE_SERVICES_PKG includes L3 functionality and in order to perform the ISSU this package must be unused and the license file has to be cleared from the device by using the "clear license LICENSEFILE" command. It is possible that the license file is currently in use by the device. In order to clear such a license file it is important to check which packages are in use by using the "show license usage" and disabling the features of these packages.

Non-edge STP ports

During testing it was also necessary to shutdown the northbound port-channel as it did not pass the "show spanning-tree issu-impact" non-edge, Criteria 3, check and this would have lead to a disruptive upgrade. This northbound port-channel was listed not as a vPC Edge in the "show spanning-tree vlan 1" command.

Loss of peer keepalive link

After the loss of the peer keepalive mgmt0 link no disruption in the traffic was recorded. In this topology, the management interface (mgmt0) is used as keepalive link, hence does not impact the data traffic generated during testing.

```
[  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%s), remote/receiver 0.0% (0.0%u/0.0%s)
iperf3: interrupt - the client has terminated
```

The devices notice mgmt0 interface going down, and peer keepalives failing, but since peer link is up data place communication can continue.

```
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
```

Disabling vPC feature

This test will describe what happens when vPC is disabled on one of the switches during live data transfer.

VPC feature can be disabled using the following command in the global configuration mode:

```
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
```

Disabling the vPC feature on either primary or secondary vPC peer leads to instant loss of data connectivity. This is due to the peer based nature of vPC. As soon as the feature is disabled, all vPC functionality on the switch ceases to function, the peer link goes down, vPC keepalive status is Suspended and port-channel 101 of the testing environment goes down. This is evident in the show vPC output of the peer switch which still has vPC feature enabled.

```
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
```

The traffic interruption, as before, is only short lived.

Under above mentioned testing conditions 50-80 packets were lost from a single session.

Remove "feature vpc" command also caused vPC configuration o be removed from port-channels.

This configuration needs to be readded.

Conclusion

vPC feature is intended to bring resiliency performance by splitting data traffic in a port channel among multiple devices.

This simple idea requires complicated control plane implementations.

The above tests were meant to show disruptions to both the control- and data-plane which may occur during life cycle of the feature.

As expected data plane disruptions were detected and corrected almost immediately - with single packets lost in tests.

The control plane disruptions tested show that vPC still maintains sub-second convergence time even when control plane is affected.

The most disruptive test performed - vPC peer link being shut down - potentially combines both data and control plane failure. Still a fast convergence time was demonstrated.