

FlexPod Nexus 5000 en comportamiento de vPC durante la interrupción

Contenido

[Introducción](#)

[Topología](#)

[Pruebas](#)

[Pérdida de enlace de datos](#)

[Actualización o recarga disruptiva](#)

[Enlace de peer vPC inactivo](#)

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

[Problemas conocidos con ISSU](#)

[Pérdida del link de keepalive de peer](#)

[Desactivación de la función vPC](#)

[Conclusión](#)

[Conversaciones relacionadas de la comunidad de soporte de Cisco](#)

Introducción

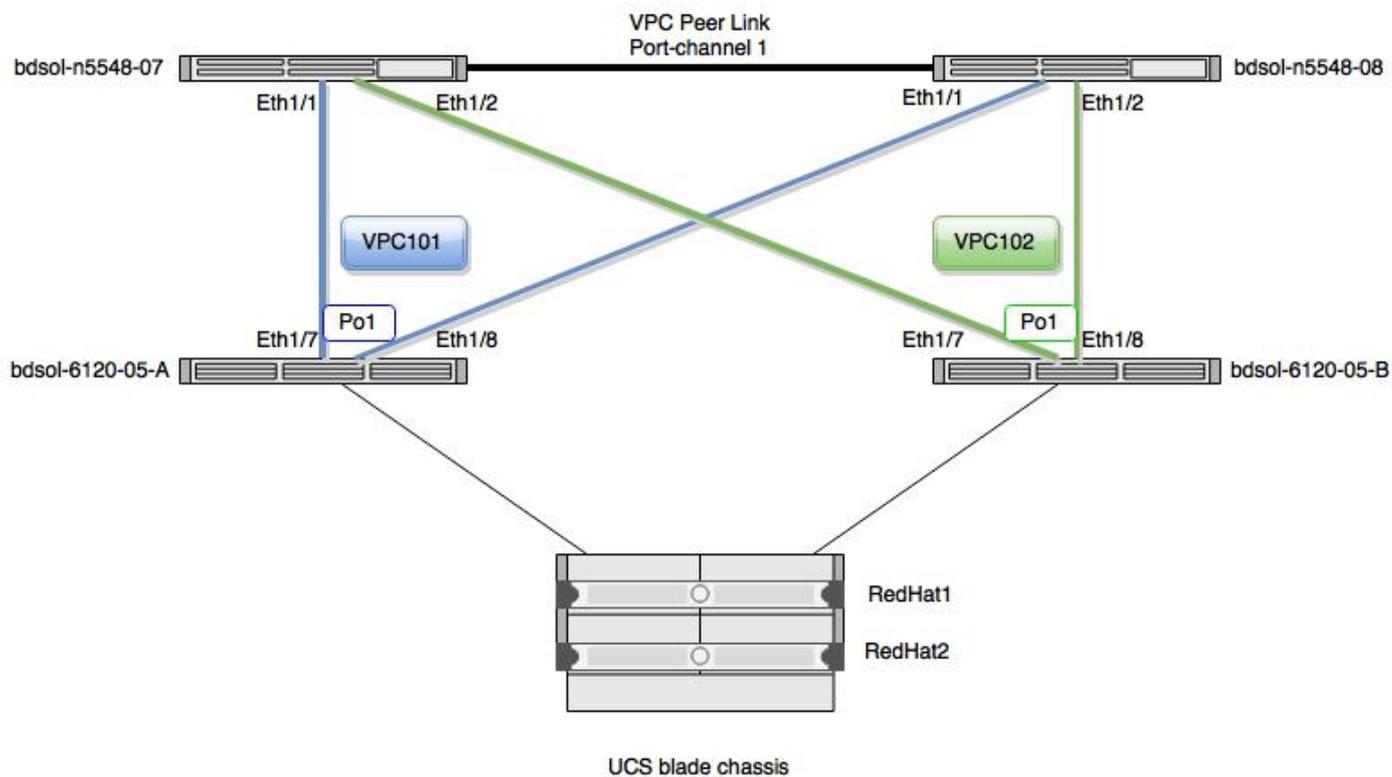
Este documento describirá el comportamiento de la red en reacción a las diferentes interrupciones, concentrándose en Virtual Port-Channel (vPC).

Una interrupción típica sería una recarga, pérdida de link o pérdida de conectividad.

El objetivo de este documento es demostrar la pérdida de paquetes durante escenarios comunes.

Topología

Durante la prueba, a menos que se indique lo contrario a continuación de la topología.



Las líneas verdes y azules indican un canal de puerto vPC de cada Fabric Interconnects a ambos switches Nexus.

No se describe la red de administración fuera de banda.

Se trata de una topología simplificada que se recomienda habitualmente en implementaciones de FlexPod, como por ejemplo en:

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

Componentes Utilizados

Dos switches Nexus 5548P.

Dos Fabric Interconnect Unified Computing System (UCS) 6120 con software 2.2(4b).

Un chasis UCS 5108.

Dos blades B200M3 con adaptador VIC 1240 que ejecutan el software 2.2(4).

Para realizar y verificar las pruebas de conectividad se instalaron dos blades y se instaló el sistema operativo RedHat Enterprise Linux 7.1.

Configuración.

Tanto la configuración de vPC como de portchannel están utilizando el valor predeterminado.

```
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
```

Ejemplo de vPC que conduce a UCS Fabric Interconnect (FI) en este caso, bdsol-6120-05: A

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

Pruebas

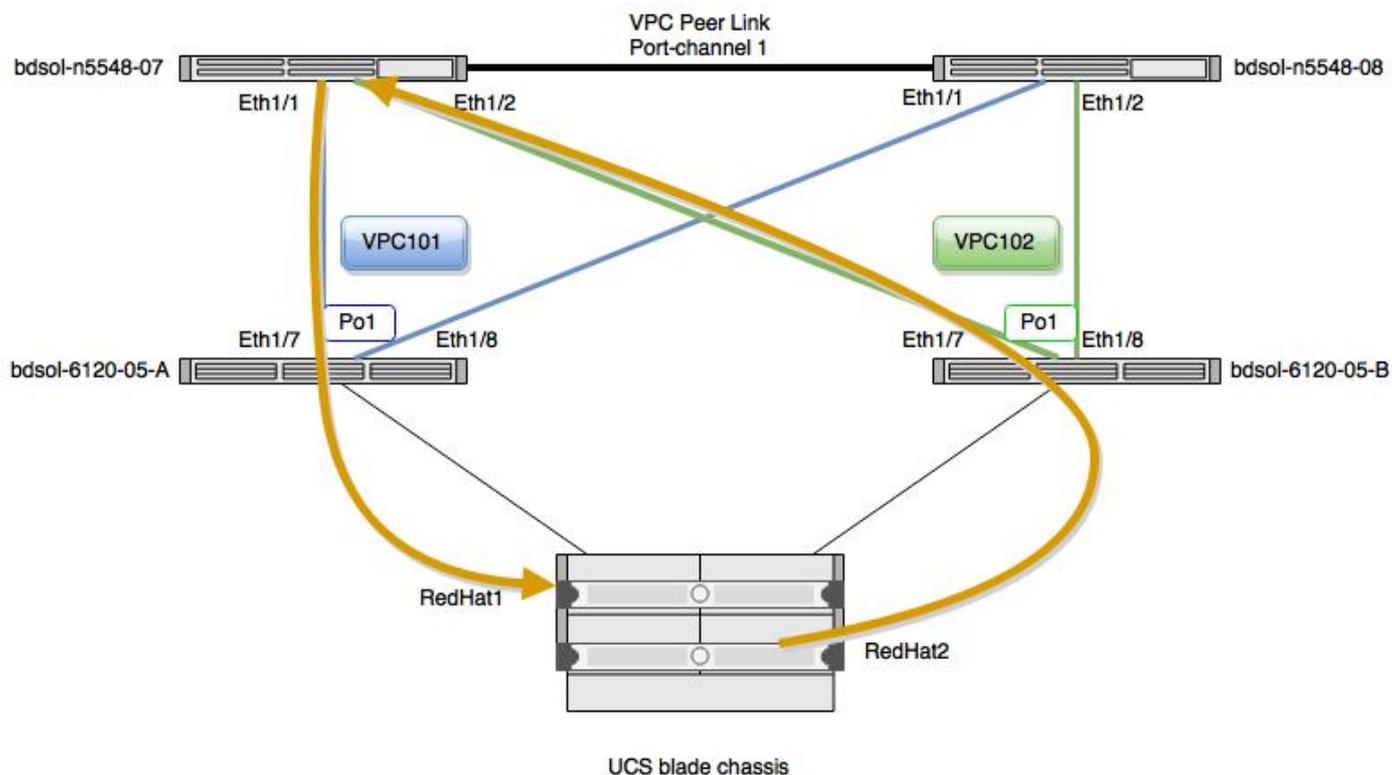
Se realizará la siguiente prueba.

- Pérdida de link de datos.
- Actualización disruptiva
- In-Service Software Upgrade (ISSU)
- Pérdida del link keepalive de peer - interfaz mgmt0 en caso de esta topología/configuración.
- Pérdida del portchannel del peer - Port-channel 1 en esta configuración.
- Inhabilitación de la función vPC

Flujo de tráfico básico.

Una sola sesión iperf3 se utiliza para generar 6,5 gigabits por segundo del tráfico TCP de prueba para verificar la pérdida de tramas durante las transiciones.

RedHat2 está anclado a Fabric Interconnect B mientras que RedHat1 está anclado a Fabric Interconnect A - esto produce tráfico que necesita cruzar la parte de conmutación.



Parámetros Iperf3:

- Servidor: `iperf3 -s -i 1`
- Cliente `iperf3 -c 10.37.9.131 -t 0 -i 1 -w 1M -V`

Los parámetros anteriores se seleccionaron para permitir una alta velocidad de tráfico y una pérdida de paquetes fácil de detectar.

La ventana TCP está atascada para evitar ráfagas de datos que el iperf conoce. Permitir que iperf se ejecute sin atornillar podría ocasionar caídas ocasionales en los búferes de ingreso a lo largo de la trayectoria, dependiendo de la configuración de QoS. Los parámetros anteriores permiten una velocidad sostenida de 6-7 Gbps sin pérdida de trama.

Para verificar podemos verificar la velocidad acumulada del tráfico en las 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
```

El resultado anterior muestra 7 Gbps de tráfico que ingresa en la interfaz Ethernet 1/2 y sale en la interfaz Ethernet 1/1.

Pérdida de enlace de datos

Esta prueba se designa para probar cómo se comportarán los datos si se apaga un link que forma parte de vPC.

En este ejemplo se utilizará Ethernet 1/1, la interfaz de salida para el tráfico de datos; se cerrará mediante la línea de comandos.

```
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
```

En este caso, solo se perdió un paquete, debido a la inundación de flujo de 6,5 Gbps.

[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

El tráfico se equilibra casi inmediatamente entre los enlaces restantes en portchannel en UCS, en este caso usando el puerto Ethernet 1/8 (el único restante) de UCS FI B que llega al Nexus 5548 B, desde ahí se transportará a UCS FI A mediante 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
```

Actualización o recarga disruptiva

Se puede emular una interrupción combinada de los datos y del plano de control mediante una actualización disruptiva del bdsol-n5548-07 (vPC principal).

Se espera una pérdida de tráfico.

Funcionalmente, esta prueba es la misma que la recarga de un par vPC.

```
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.

Después de los 10 segundos mencionados se produce la pérdida de paquetes.

Durante ese tiempo, sólo se pierden 55 paquetes (de la secuencia de 6,6 Gbps).

```
[ 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
```

Si el iperf3 se reinició inmediatamente, el operador puede verificar que el tráfico realmente se conmutó a 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
```

La velocidad de tráfico se muestra por debajo de los 6 Gbps, ya que el contador de velocidad se promedia durante 30 segundos.

Enlace de peer vPC inactivo

En este ejemplo, el link de par vPC se desactiva, activado por un cambio de configuración.

En ese momento, el tráfico es manejado por bdsol-n5548-07, actuando vPC secundario.

Secuencia de eventos.

El canal de puerto 1 deja de funcionar.

10 de julio de 2015 15:00:25 bdsol-n5548-07 %ETHPORT-5-IF_DOWN_CFG_CHANGE: La interfaz port-channel1 está inactiva(cambio de configuración)

Dado que bdsol-n5548-07 actúa como secundario, suspenderá sus vPC, ya que no puede garantizar la topología de descarga:

```
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)
```

Durante este tiempo, iperf3 perdió una parte del tráfico: 90 paquetes.

```
[ 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
```

Pero fue capaz de recuperarse bastante rápido.

Dado que los vPC están suspendidos en bdsol-n5548-07, todo el tráfico es manejado por 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
```

De nuevo, la velocidad no muestra 6,5 gigabits por segundo inmediatamente debido a que se calcula el promedio de carga.

Recuperación del enlace vPC inactivo.

Cuando el link de par vPC vuelve a funcionar, es posible que se vuelva a equilibrar el tráfico entre los enlaces y que se espere una pérdida de paquetes de corta duración debido al cambio de topología.

En el caso de esta prueba de laboratorio se perdió 1 paquete.

In-Service Software Upgrade (ISSU)

En esta prueba se realizó una actualización de ISSU para verificar la interrupción del tráfico.

Las funciones de vPC durante esta prueba son las siguientes:

bdsol-n5548-07 - primario

bdsol-n5548-08 - secundaria.

Para realizar un criterio definido por ISSU se deben cumplir.

Para encontrar información sobre los comandos usados para verificar estos criterios y realizar un ISSU se utilizó la siguiente guía:

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

Después de realizar un ISSU primero en el primario y después en el par vPC secundario, no se han perdido paquetes.

Esto se debe al hecho de que ISSU no interrumpe todas las funciones del plano de datos y que solo el tráfico del plano de control se vería afectado.

```
[ 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
```

Problemas conocidos con ISSU

Funciones y licencias de capa 3.

Durante las pruebas de ISSU, es necesario resolver varios problemas. "show install all Impact..." puede proporcionar resultados que ISSU no se puede realizar con la siguiente explicación: "No se admite la instalación no disruptiva si se ha activado L3". En el entorno de prueba, esto se debió a que LAN_BASE_SERVICES_PKG se estaba utilizando en el archivo de licencia instalado.

LAN_BASE_SERVICES_PKG incluye la funcionalidad L3 y para realizar el ISSU, este paquete debe estar sin utilizar y el archivo de licencia debe eliminarse del dispositivo mediante el comando "clear license LICENSEFILE". Es posible que el dispositivo esté utilizando el archivo de licencia. Para borrar tal archivo de licencia, es importante verificar qué paquetes están en uso usando "show license usage" y desactivando las funciones de estos paquetes.

Puertos STP no de borde

Durante las pruebas, también fue necesario cerrar el canal de puerto ascendente, ya que no pasó el "show spanning-tree issu-Impact" non-edge, según los Criterios 3, y esto habría llevado a una actualización disruptiva. Este canal de puerto ascendente no figuraba como extremo de vPC en el comando "show spanning-tree vlan 1".

Pérdida del link de keepalive de peer

Después de la pérdida del link del peer keepalive mgmt0, no se registró ninguna interrupción en el tráfico. En esta topología, la interfaz de administración (mgmt0) se utiliza como enlace keepalive, por lo que no afecta al tráfico de datos generado durante las pruebas.

```
[ 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
[ 41] 0.00-76.02 sec 0.00 Bytes 0.00 bits/sec
CPU Utilization: local/sender 35.2% (0.5%u/34.7%u), remote/receiver 0.0% (0.0%u/0.0%u)
iperf3: interrupt - the client has terminated
```

Los dispositivos advierten que la interfaz mgmt0 se desactiva y que las señales de mantenimiento de peer fallan, pero dado que el link de peer está activo, la comunicación del lugar de datos puede continuar.

```
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
```

Desactivación de la función vPC

Esta prueba describirá lo que sucede cuando se inhabilita vPC en uno de los switches durante la transferencia de datos en directo.

La función VPC se puede inhabilitar mediante el siguiente comando en el modo de configuración global:

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

Si se desactiva la función vPC en el par vPC primario o secundario, se produce una pérdida instantánea de la conectividad de datos. Esto se debe a la naturaleza basada en pares de vPC. Tan pronto como se inhabilita la función, toda la funcionalidad de vPC en el switch deja de funcionar, el link de par se desactiva, el estado de keepalive de vPC se suspende y el canal de puerto 101 del entorno de prueba se desactiva. Esto es evidente en la salida show vPC del switch peer que aún tiene activada la función vPC.

```
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 -
```

La interrupción del tráfico, como antes, es de poca duración.

En las condiciones de prueba mencionadas anteriormente, se perdieron entre 50 y 80 paquetes de una sola sesión.

El comando "feature vpc" también provocó que la configuración de vPC se quitara de los canales de puerto.

Esta configuración debe ser leída.

Conclusión

La función vPC pretende aportar un rendimiento de resistencia dividiendo el tráfico de datos en un canal de puerto entre varios dispositivos.

Esta idea sencilla requiere implementaciones complicadas del plano de control.

Las pruebas anteriores tenían por objeto mostrar las interrupciones tanto en el plano de control como en el plano de datos que pueden producirse durante el ciclo de vida de la función.

Como se esperaba, las interrupciones del plano de datos se detectaron y corrigieron casi inmediatamente, con paquetes únicos perdidos en las pruebas.

Las interrupciones en el plano de control que se han probado muestran que vPC mantiene un tiempo de convergencia de menos de un segundo incluso cuando el plano de control se ve afectado.

La prueba más disruptiva realizada (el enlace de par vPC que se está cerrando) potencialmente combina fallos de plano de control y datos. Aún así, se demostró un tiempo de convergencia rápido.