

FlexPod Nexus 5k nel comportamento vPC durante l'interruzione

Sommario

[Introduzione](#)

[Topologia](#)

[Test](#)

[Perdita di collegamento dati](#)

[Upgrade o ricaricamento con interruzioni](#)

[Collegamento peer vPC non attivo](#)

[Aggiornamento software in servizio \(ISSU\)](#)

[Problemi noti relativi a ISSU](#)

[Perdita del collegamento peer keepalive](#)

[Disattivazione funzionalità vPC](#)

[Conclusioni](#)

[Discussioni correlate nella Cisco Support Community](#)

Introduzione

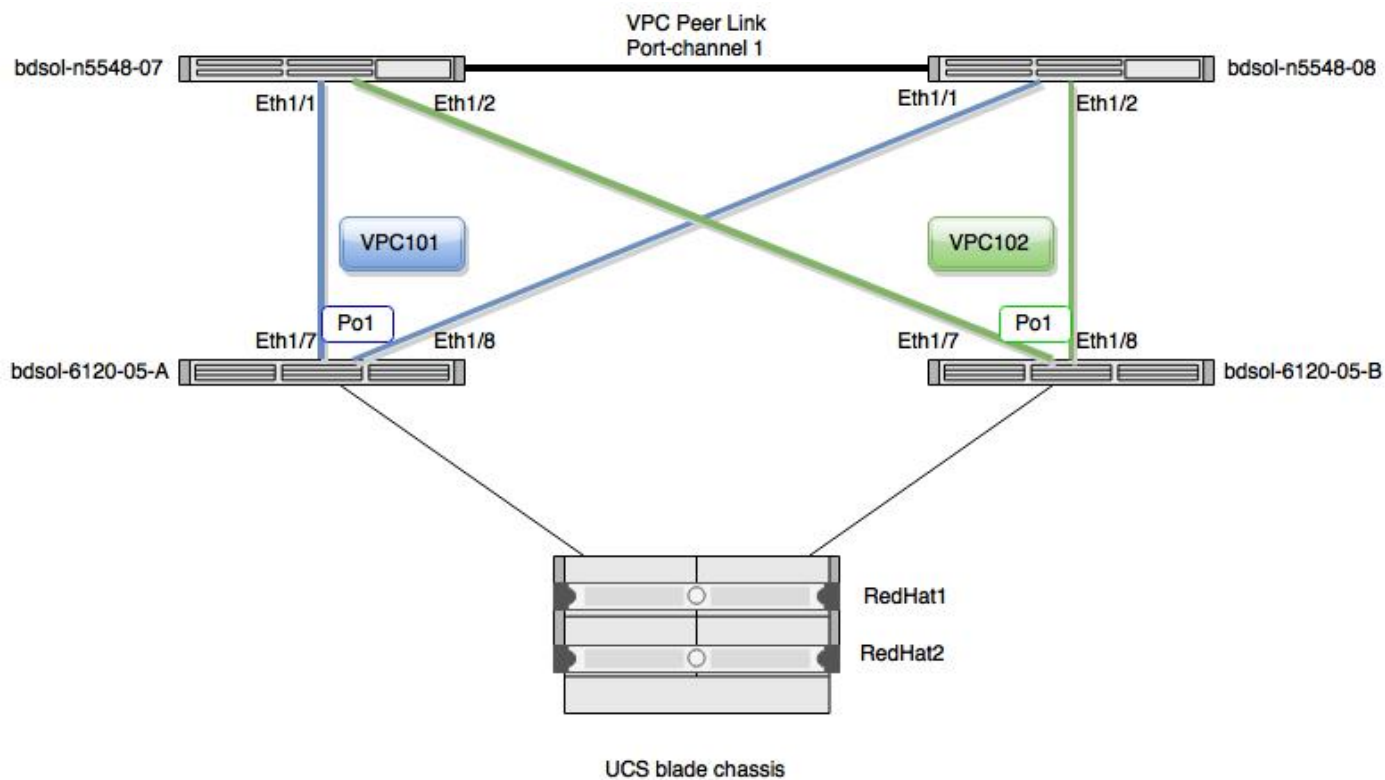
Questo documento delinea il comportamento della rete in reazione a diversi guasti, concentrandosi su Virtual Port-Channel (vPC).

Un'interruzione tipica potrebbe essere un ricaricamento, una perdita di collegamento o una perdita di connettività.

Lo scopo di questo documento è quello di dimostrare la perdita di pacchetti durante scenari comuni.

Topologia

Durante il test, se non diversamente specificato nella topologia.



Le linee verdi e blu indicano un canale di porta vPC da ciascuna interconnessione fabric a entrambi gli switch Nexus.

Non è descritta la rete di gestione fuori banda.

Si tratta di una topologia semplificata comunemente consigliata nelle implementazioni FlexPod, come mostrato ad esempio in:

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

Componenti usati

Due switch Nexus 5548P.

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

Uno chassis UCS 5108.

Due blade B200M3 con adattatore VIC 1240 e software 2.2(4).

Per eseguire e verificare i test di connettività sono stati installati due blade e il sistema operativo RedHat Enterprise Linux 7.1 è installato.

Configurazione.

Per la configurazione di vPC e portchannel viene utilizzata l'impostazione predefinita.

```
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
```

Esempio di vPC che porta a UCS Fabric Interconnect (FI) in questo caso bdsol-6120-05—A

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

Test

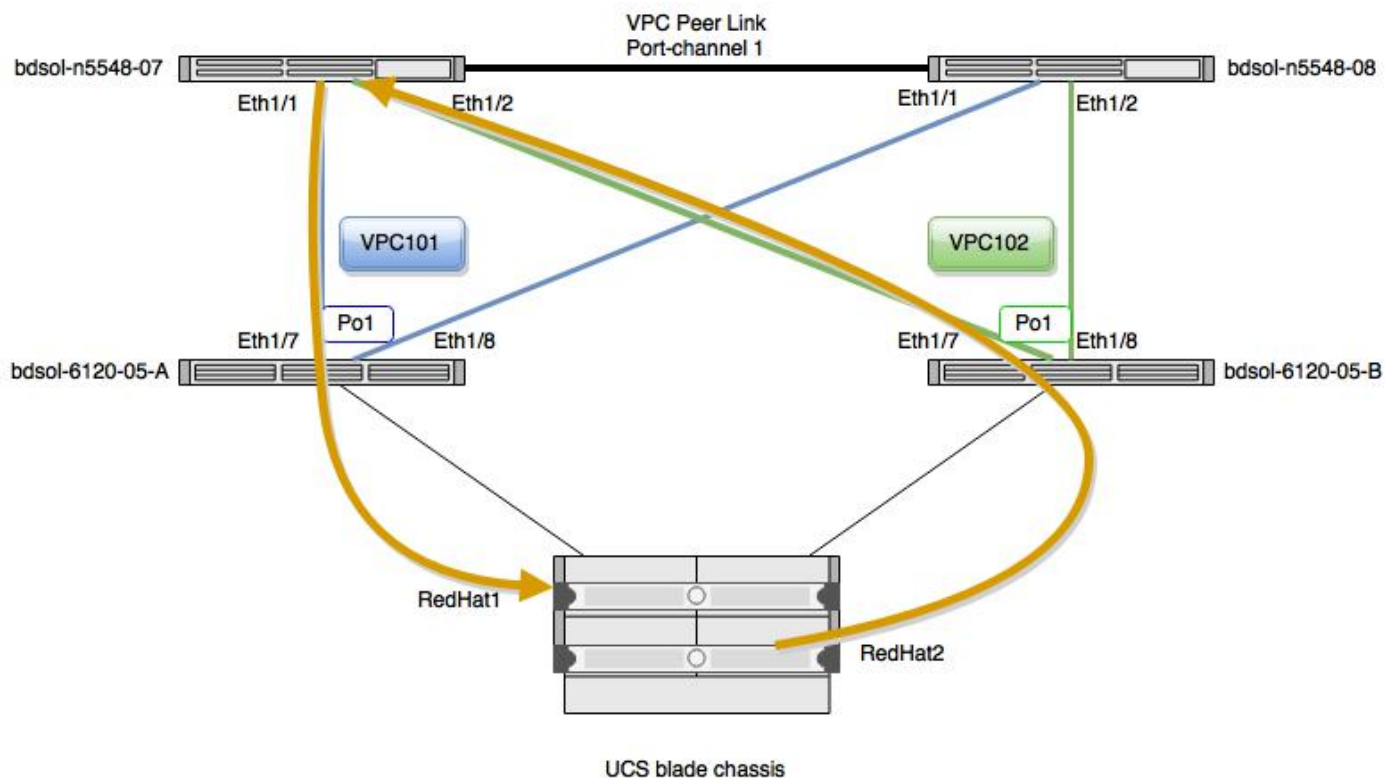
Verrà eseguito il seguente test.

- Perdita del collegamento dati.
- Upgrade con interruzioni
- Aggiornamento software in servizio (ISSU)
- Perdita del collegamento keepalive peer - interfaccia mgmt0 in caso di topologia/configurazione.
- Perdita del portchannel peer - Port-channel 1 in questa configurazione.
- Disattivazione funzione vPC

Flusso del traffico di base.

Una singola sessione iperf3 viene utilizzata per generare 6,5 gigabit al secondo di traffico TCP di prova per verificare la perdita di frame durante le transizioni.

RedHat2 è bloccato sulla Fabric Interconnect B mentre RedHat1 è bloccato sull'interconnessione A del fabric: questo genera traffico che deve attraversare la porzione di commutazione.



Parametri Iperf3:

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

I suddetti parametri sono stati scelti per consentire un'alta velocità di traffico e una facile individuazione delle perdite dei pacchetti.

La finestra TCP è bloccata per evitare burst di dati per cui iperf è noto. Consentendo l'iperf di eseguire lo sblocco, si potrebbero verificare occasionali cadute nei buffer in entrata lungo il percorso, a seconda della configurazione QoS. I suddetti parametri consentono una velocità sostenuta di 6-7 Gbps senza perdita di frame.

Per verificare, è possibile controllare la velocità cumulativa del traffico sulle interfacce.

```
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
```

L'output precedente mostra 7 Gbps di traffico in entrata sull'interfaccia Ethernet 1/2 e in uscita sull'interfaccia Ethernet 1/1.

Perdita di collegamento dati

Questo test è progettato per verificare il comportamento dei dati se un collegamento che fa parte di vPC viene arrestato.

In questo esempio viene utilizzato Ethernet 1/1, l'interfaccia di output per il traffico di dati, che viene chiusa dalla riga di comando.

```

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

```

In questo caso è stato perso un solo pacchetto, su un flusso di 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

Il traffico viene quasi immediatamente bilanciato tra i collegamenti rimanenti in portchannel su UCS, in questo caso utilizzando la porta Ethernet 1/8 (l'unica restante) di UCS FI B fino a Nexus 5548 B, da dove verrà trasportato a UCS FI A utilizzando 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

```

Upgrade o ricaricamento con interruzioni

È possibile emulare un'interruzione combinata dei dati e del control plane eseguendo un aggiornamento dirompente del bdsol-n5548-07 (vPC primario).

È prevista una perdita di traffico.

Dal punto di vista funzionale, questo test equivale a ricaricare un peer 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.

Dopo i 10 secondi indicati, il pacchetto viene perso.

Durante questo periodo vengono persi solo 55 pacchetti (su un flusso di 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
```

Se l'iperf3 viene riavviato immediatamente, l'operatore può verificare che il traffico sia effettivamente passato 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 velocità del traffico è inferiore a 6 Gb/s poiché il contatore della velocità è in media di oltre 30 secondi.

Collegamento peer vPC non attivo

In questo esempio il collegamento peer vPC diventa inattivo, attivato da una modifica della configurazione.

A quel punto il traffico viene gestito da bdsol-n5548-07, che agisce come vPC secondario.

Sequenza degli eventi.

Il canale porta 1 si blocca.

2015 lug 10 15:00:25 bdsol-n5548-07 %ETHPORT-5-IF_DOWN_CFG_CHANGE: Interface port-channel1 non è attivo (modifica della configurazione)

Poiché bdsol-n5548-07 agisce come secondario, sospenderà i suoi vPC poiché non può garantire una topologia senza loop:

```
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 questo periodo iperf3 perse una parte del traffico - 90 pacchetti.

```
[ 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
```

Ma è stato in grado di riprendersi abbastanza in fretta.

Poiché i vPC sono sospesi su bdsol-n5548-07, tutto il traffico viene gestito da 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
```

Anche in questo caso, la velocità non mostra 6,5 gigabit al secondo immediatamente a causa del calcolo della media di carico.

Ripristino dal collegamento vPC non attivo.

Quando il collegamento peer vPC torna attivo, è possibile che il traffico venga riequilibrato tra i collegamenti e che si verifichi una perdita di pacchetti di breve durata dovuta a modifiche della topologia.

In questo test di laboratorio 1 pacchetto è andato perso.

Aggiornamento software in servizio (ISSU)

In questo test è stato eseguito un aggiornamento di ISSU per verificare l'interruzione del traffico.

I ruoli vPC durante il test sono i seguenti:

bdsol-n5548-07 - primario

bdsol-n5548-08 - secondario.

Per eseguire un'operazione è necessario che siano soddisfatti i criteri definiti dall'ISSU.

Per trovare informazioni sui comandi utilizzati per verificare questi criteri ed eseguire un problema, è stata utilizzata la seguente guida:

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

Dopo aver eseguito un'operazione ISSU prima sul sistema primario e successivamente sul sistema vPC secondario, non è stato perso alcun pacchetto.

Ciò è dovuto al fatto che l'ISSU non interrompe tutte le funzionalità del data plane e influisce solo sul traffico del control plane.

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

Problemi noti relativi a ISSU

Funzionalità e licenze di layer 3.

Durante la verifica dell'ISSU è stato necessario risolvere una serie di problemi. Il messaggio "show install all impact ..." potrebbe fornire un output che non è possibile eseguire con la seguente spiegazione: "Installazione senza interruzioni non supportata se L3 è stato abilitato." Nell'ambiente di test, ciò è dovuto al fatto che LAN_BASE_SERVICES_PKG è in uso nel file di licenza installato.

LAN_BASE_SERVICES_PKG include la funzionalità L3 e per eseguire il comando IOS, questo pacchetto deve essere inutilizzato e il file della licenza deve essere cancellato dal dispositivo usando il comando "clear license LICENSEFILE". È possibile che il file di licenza sia attualmente utilizzato dal dispositivo. Per cancellare un file di licenza di questo tipo, è importante controllare quali pacchetti sono in uso usando il comando "show license usage" e disabilitando le funzionalità di questi pacchetti.

Porte STP non edge

Durante i test è stato necessario arrestare anche il canale della porta in direzione nord, in quanto non ha superato il controllo "show spanning-tree issu-impact" non-edge, Criteria 3, e questo avrebbe portato a un upgrade con interruzioni. Questo canale della porta verso nord non è stato elencato come vPC Edge nel comando "show spanning-tree vlan 1".

Perdita del collegamento peer keepalive

Dopo la perdita del collegamento peer keepalive mgmt0, non è stata registrata alcuna interruzione del traffico. In questa topologia, l'interfaccia di gestione (mgmt0) viene utilizzata come collegamento keepalive, pertanto non influisce sul traffico di dati generato durante il test.

```
[ 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
```

I dispositivi rilevano l'interruzione dell'interfaccia mgmt0 e un errore dei pacchetti keepalive peer, ma poiché il collegamento peer è attivo, la comunicazione tra i siti dati può continuare.

```
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
```

Disattivazione funzionalità vPC

Questo test descrive cosa succede quando il vPC viene disabilitato su uno degli switch durante il trasferimento di dati in tempo reale.

La funzione VPC può essere disabilitata usando il seguente comando nella modalità di configurazione globale:

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

La disattivazione della funzionalità vPC sul peer vPC primario o secondario comporta la perdita immediata della connettività dei dati. Ciò è dovuto alla natura peer-based di vPC. Non appena la funzione viene disabilitata, tutte le funzionalità vPC sullo switch cessano di funzionare, il collegamento peer si interrompe, lo stato vPC keepalive viene sospeso e il canale porta 101 dell'ambiente di test si interrompe. Ciò è evidente nell'output show vPC dello switch peer per cui è ancora abilitata la funzione 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 -
```

L'interruzione del traffico, come in precedenza, è di breve durata.

Nelle condizioni di test sopra descritte, 50-80 pacchetti sono stati persi da una singola sessione.

La rimozione del comando "feature vpc" ha causato anche la rimozione della configurazione vPC dai canali delle porte.

Questa configurazione deve essere letta.

Conclusioni

La funzionalità vPC è progettata per offrire prestazioni di resilienza dividendo il traffico di dati in un canale porta tra più dispositivi.

Questa semplice idea richiede complicate implementazioni del control plane.

I test sopra descritti sono stati concepiti per mostrare interruzioni sia al control-plane che al data-plane che possono verificarsi durante il ciclo di vita della feature.

Come previsto, le interruzioni del data plane sono state rilevate e corrette quasi immediatamente, con la perdita di singoli pacchetti durante i test.

Le interruzioni del control plane testate mostrano che vPC mantiene ancora il tempo di convergenza al di sotto del secondo anche quando il control plane è interessato.

Il test più dirompente eseguito, ovvero l'arresto del collegamento peer vPC, combina potenzialmente sia il guasto dei dati che quello del control plane. È stato comunque dimostrato un rapido tempo di convergenza.