

瞭解vPC選舉過程

目錄

[簡介](#)

[必要條件](#)

[需求](#)

[採用元件](#)

[虛擬PortChannel技術](#)

[vPC角色](#)

[vPC角色優先順序](#)

[無中斷vPC角色更改](#)

[vPC對等鏈路斷開時vPC系統的行為](#)

[vPC主粘滯位](#)

[vPC延遲恢復](#)

[vPC延遲恢復介面Vlan](#)

[使用4000 SVI擴展設定時的vPC延遲恢復](#)

[vPC選舉流程](#)

[vPC恢復方案](#)

[與錯誤設定粘滯位相關的網路中斷示例](#)

簡介

本檔案介紹Nexus系列交換器上的虛擬PortChannel (vPC)角色選舉程式。

必要條件

需求

思科建議您瞭解以下主題：

- nexus系列交換機上的vPC
- 生成樹通訊協定(STP)

採用元件

本文檔中的資訊基於Nexus 9000系列交換機平台。

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除 (預設) 的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

虛擬PortChannel技術

虛擬PortChannel (vPC)允許物理上連線到兩台不同Cisco交換機的鏈路顯示為連線到第三台裝置的單個PortChannel。第三個裝置可以是交換器、伺服器或任何其他支援IEEE 802.3ad PortChannels的網路裝置。vPC也允許建立跨兩個交換器的第2層PortChannels。目前，vPC在Cisco Nexus 9000、7000、5000和3000系列平台（使用或不使用Cisco Nexus 2000系列交換矩陣擴展器）上實施。

 注意：Cisco NX-OS軟體vPC和Cisco Catalyst虛擬交換系統(VSS)具有類似的技術。對於Cisco EtherChannel技術，術語Multi-chassis EtherChannel (MCEC)可交換是指其中一種技術。

vPC角色

雖然兩台vPC交換機都顯示為通向下游裝置的單個交換機，但其中兩台vPC交換機明確定義了vPC角色：vPC主要和vPC輔助。

vPC角色為非搶佔式，這意味著裝置可以配置為vPC主裝置，但作為vPC輔助對等裝置運行。這種情況可能會發生：

1. 當原始主裝置發生故障時，輔助vPC裝置會成為新的主裝置。
2. 當系統自行恢復時，以前的主裝置現在為輔助裝置，反之亦然。

vPC角色定義了兩個vPC對等裝置中哪一個處理網橋協定資料單元(BPDU)並響應地址解析協定(ARP)請求。vPC角色還定義了一組要由vPC主裝置和vPC輔助裝置執行的操作，以響應vPC對等鏈路斷開情況。

vPC角色優先順序

您也可以使用vPC域模式命令中使用角色優先順序來影響vPC選舉過程。值範圍為1至65536，預設值為32768。值越低，表示此交換機成為主vPC的機會就越高。

更改vPC對等裝置的優先順序時，可能導致網路中的介面打開和關閉。如果您要再次配置角色優先順序以使一個vPC裝置成為主裝置，請用較低的優先順序值配置主vPC裝置的角色優先順序，用較高的值配置輔助vPC裝置的角色優先順序。然後，關閉兩台裝置上的vPC對等鏈路並輸入shutdown命令，最後重新啟用兩台裝置上的埠通道，並輸入no shutdown命令。

無中斷vPC角色更改

vPC無中斷角色更改功能提供了一個架構，可在不影響流量流的情況下，在vPC對等裝置之間切換vPC角色。vPC角色交換基於vPC域下裝置的角色優先順序值完成。當執行vpc role preempt命令時，具有較低角色優先順序的vPC對等裝置被選為主vPC裝置。

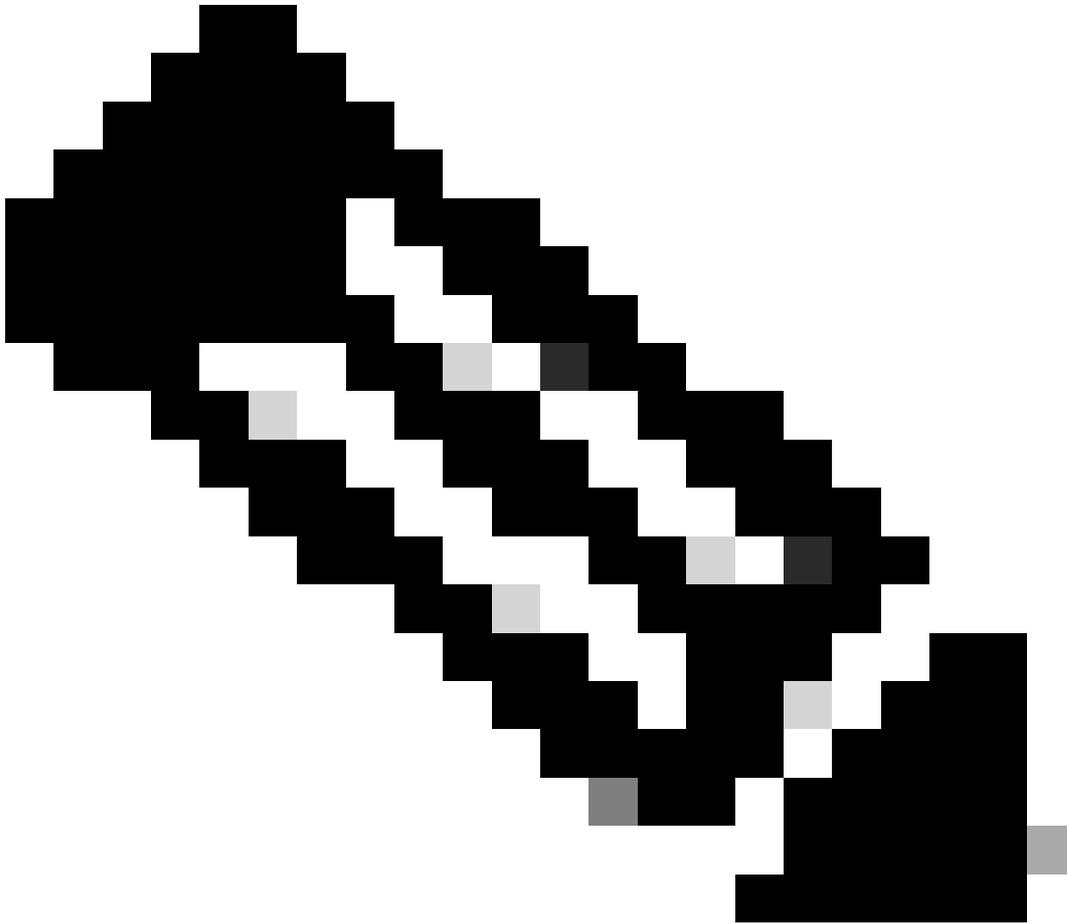
有關更多詳細資訊，請參閱[無中斷vPC角色更改的使用情形方案](#)。

vPC對等鏈路斷開時vPC系統的行為

當vPC對等鏈路發生故障且vPC對等保持連線鏈路仍處於接通狀態時，vPC輔助對等裝置將執行以下操作：

1. 掛起其vPC成員埠。
2. 關閉與vPC VLAN關聯的SVI。

vPC的這種保護行為將所有從南向北的流量重定向到vPC主裝置。



注意：當vPC對等鏈路關閉時，兩個vPC對等裝置不再能夠彼此同步，因此設計的保護機制導致其中一個對等裝置（出現次對等裝置時）與資料路徑隔離。

vPC主粘滯位

vPC主粘滯位是一種程式設計保護機制，引入此機制是為了避免在意外重新載入主交換機時不必要的角色更改（可能導致網路中斷）。vPC主粘滯位允許活動交換機在死交換機重新恢復活動狀態或隔離的交換機整合回VPC域時保留其主角色。

切換vPC主粘滯位：

1. 在此場景中，vPC主粘滯位值設定為TRUE：

- 當前的vPC主交換機重新啟動，啟用vPC的交換機將其角色從vPC輔助交換機更改為vPC運行主交換機。如果角色從「vPC操作輔助裝置」更改為「vPC主裝置」，則不會設定粘滯位。
- 當重新載入恢復計時器（預設為240秒）過期時，啟用vPC的交換機將其角色從None establish更改為vPC Primary。

2. 在下列情況下，vPC主粘滯位值設定為FALSE：

- 啟用vPC的交換機將重新啟動（預設情況下，粘滯位設定為FALSE）。
- vPC角色優先順序已更改或重新輸入。

vPC粘滯主位在vPC管理器軟體元件結構下報告，並且可以使用此NX-OS exec模式命令進行檢查。

```
Campus_N7K2-VPC# show system internal vpcm info global | include ignore-case sticky
Sticky Primary: TRUE
Campus_N7K2-VPC#
```

vPC延遲恢復

vPC對等裝置重新載入並恢復運行後，路由協定需要一段時間才能重新收斂。恢復的vPC支路可以黑洞（從接入到聚合/核心）路由流量，直到重新建立上行鏈路第3層連線。

vPC延遲恢復功能可延遲恢復的vPC對等裝置上的vPC分支啟動。vPC延遲恢復允許第3層路由協定在允許vPC分支上的任何流量之前收斂。這可以在恢復階段實現更平穩的恢復以及零資料包丟失（流量仍然會在活動的vPC對等裝置上轉移）。預設情況下此功能已啟用，vPC恢復預設計時器為30秒。計時器可以調整為特定第3層融合基線，從1秒到3600秒不等。

vPC延遲恢復介面Vlan

要延遲恢復的vPC對等裝置上的VLAN介面打開，請使用delay restore命令的interfaces-vlan選項。預設情況下此功能已啟用，vPC恢復預設計時器為10秒。

使用4000 SVI擴展設定時的vPC延遲恢復

引入了新的命令delay restore interface-VLAN batch <1-4094>，以配置路徑器，使其一次啟動一批200個SVI中的VLAN或網橋域介面。可以將vPC delay restore timer命令delay restore <Timeout value>配置為大於所有已配置批處理計時器的總和值。這樣做是為了在所有SVI都完全接通之後才啟用VPC支路，以避免流量出現任何黑洞。

示例：4000個Vlan，200個批處理，15秒延遲

延遲還原> (4000/200)x 15

vPC選舉流程

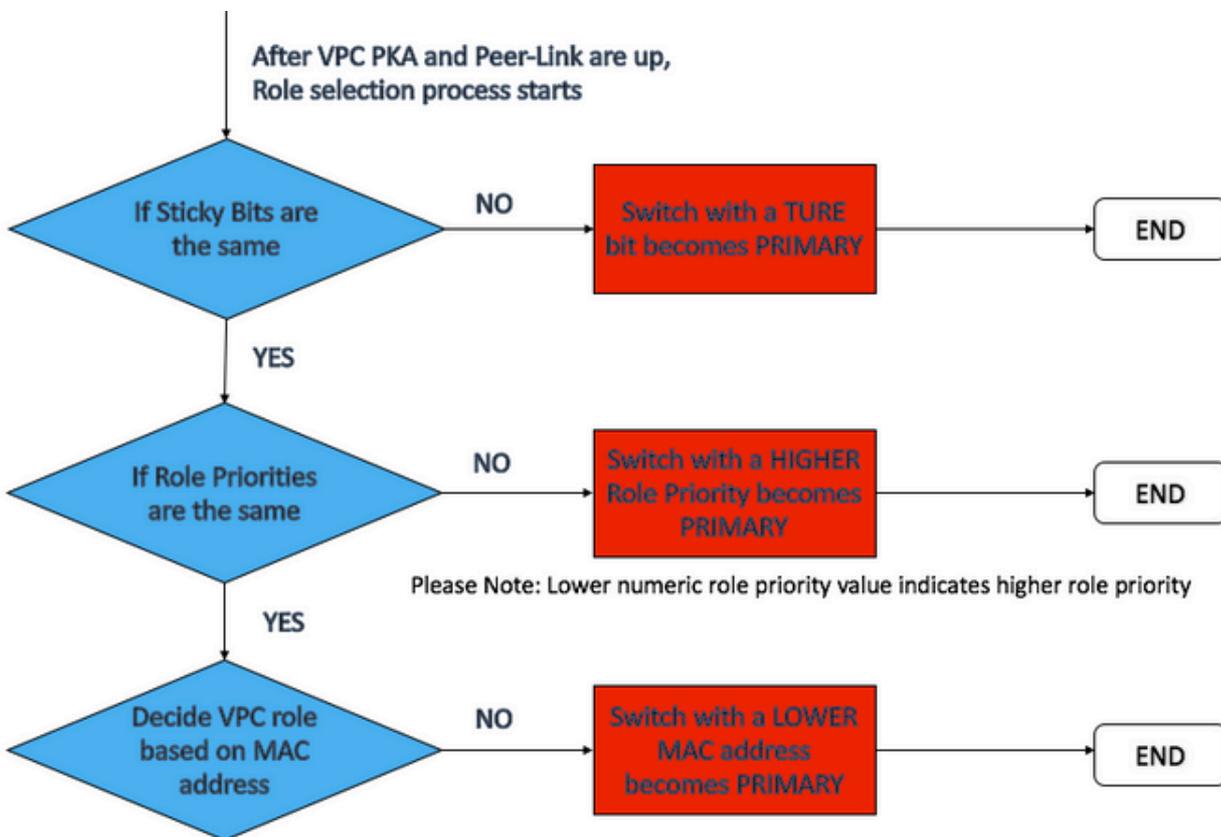
在vPC系統中，一個vPC對等裝置被定義為vPC主裝置，一個被定義為vPC輔助裝置，基於這些引數

並按照以下順序進行定義

1. vPC主粘滯位設定為0或1。
2. 使用者定義的vPC角色優先順序 (Cisco NX-OS軟體使用最小的數值選擇主裝置)。
3. 系統MAC地址值 (Cisco NX-OS軟體使用最小的MAC地址來選擇主要裝置)。

此流程圖 (圖1) 總結了兩台vPC對等裝置在vPC主交換機選舉過程中經過的步驟。

1. 在vPC主選舉過程中，兩個裝置之間的第一個檢查引數是vPC主粘滯位。如果vPC對等裝置獲得此比較，則無論兩個對等裝置具有配置的vPC角色優先順序值或系統MAC地址，該裝置都將成為vPC主裝置。
2. 如果兩台vPC對等交換機具有相同的粘滯位值，則選舉過程將繼續進行下一步，以比較使用者定義的vPC角色優先順序。
3. 如果兩個vPC角色都配置為相同的值，則選舉過程將繼續比較系統MAC地址。



影像 1

如圖所示，當vPC交換機的vPC主粘滯位設定為1 (TRUE條件)，並且其對等體的粘滯位設定為0 (FALSE條件) 時，TRUE端會贏得選舉並承擔vPC主粘滯位的角色。

vPC對等體1粘滯位設定為1	vPC對等體2粘滯位設定為1	vPC主要
False (0)	False (0)	並列
真(1)	False (0)	vPC對等體1
False (0)	真(1)	vPC對等體2
真(1)	真(1)	並列

vPC恢復方案

瞭解vPC選舉過程很重要，而且不能低估此過程，特別是在vPC恢復方案中。

圖2顯示了典型的VPC設定，Nexus-01是VPC主裝置，Nexus-02是VPC次裝置。預設情況下，這兩個介面的Sticky Bits均重新設定為FALSE。

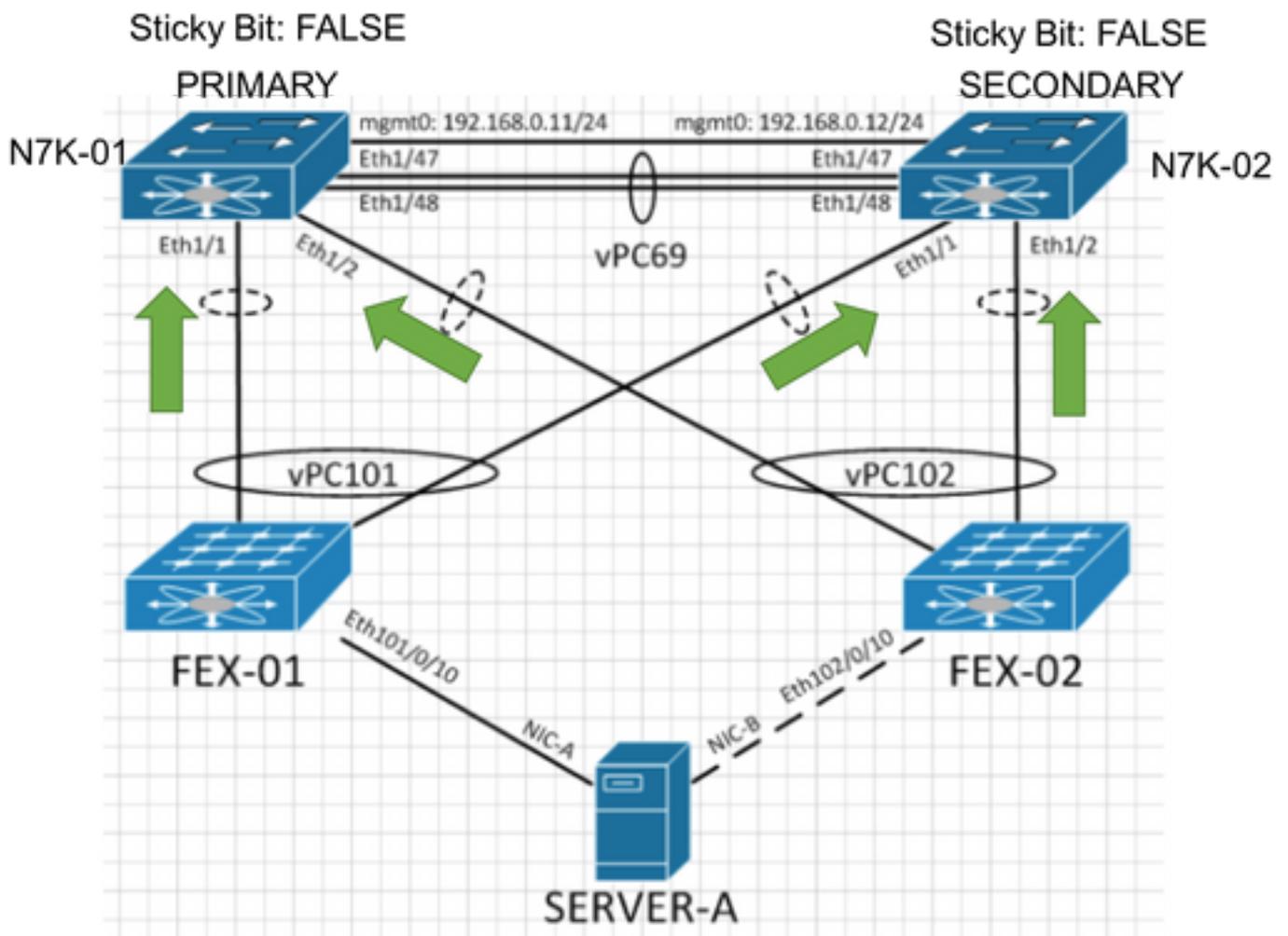


圖2

如圖所示，Nexus-01現在出現斷電，並且已與網路隔離。Nexus-02將自身提升為vPC主裝置並將vPC粘滯位設定為TRUE。

而且Nexus-02現在成為運行主裝置，粘滯位現在設定為TRUE。

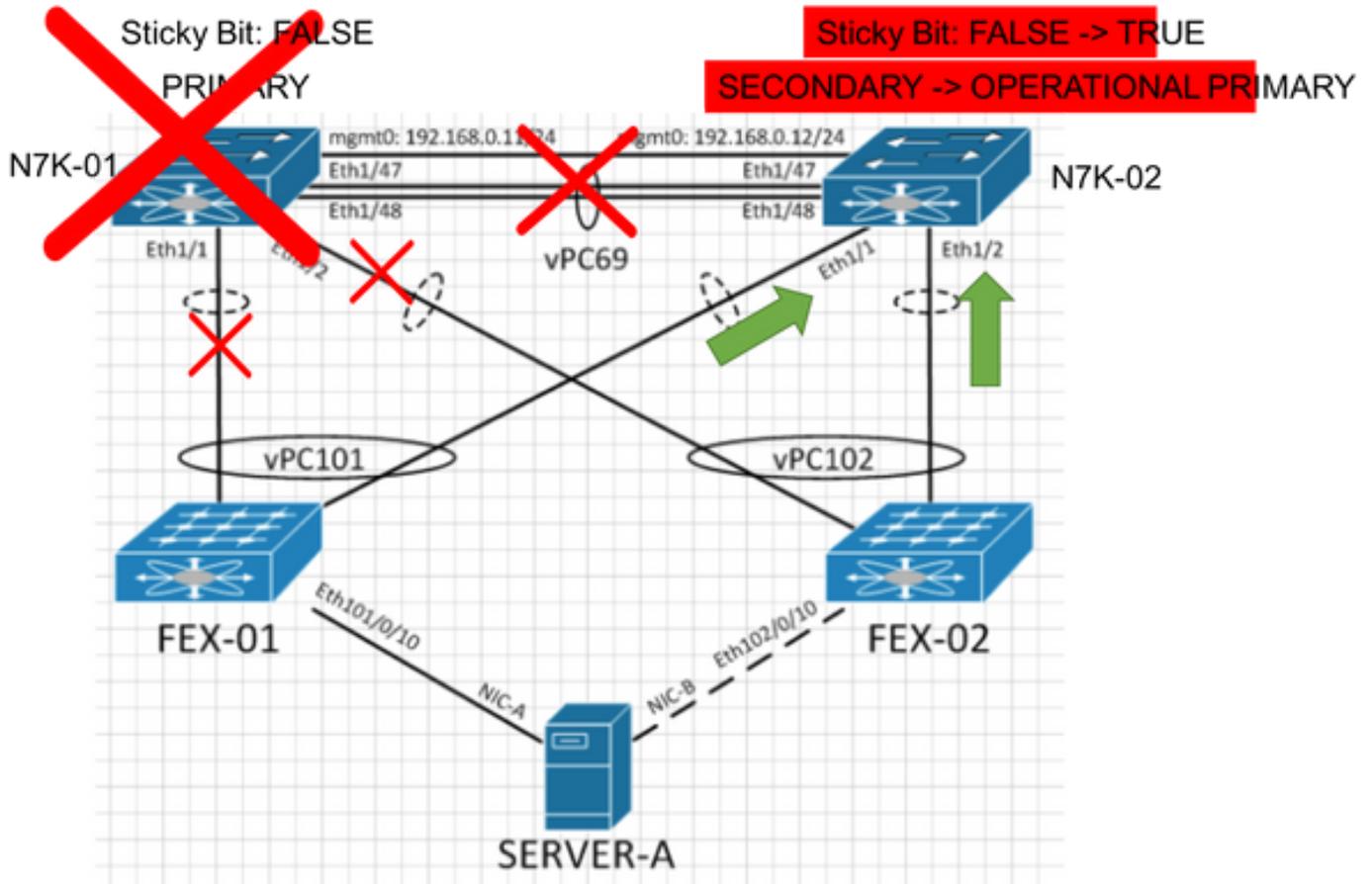


圖3

如圖所示，當斷電後Nexus-01重新上線時，Nexus-02保留運行主要角色，無論其角色優先順序如何（因為它具有TRUE粘滯位），並且Nexus-01在上線時充當輔助角色。只有Nexus-01會開始VPC初始化過程，而Nexus-02仍保持主用狀態，並照常轉發流量。因此，未發現網路中斷。

Nexus-01上的vPC初始化過程有兩個計時器，它現在是vPC運行的輔助裝置：

- 延遲還原SVI（預設情況下為10秒）
- 延遲恢復（預設情況下為30秒）

因此，在將Nexus-01重新引入網路作為vPC輔助裝置後，Nexus-01的恢復時間預計為40秒。但是，由於Nexus-02扮演主要角色，因此所有流量現在均透過上述的Nexus-01，因此不會出現網路中斷。

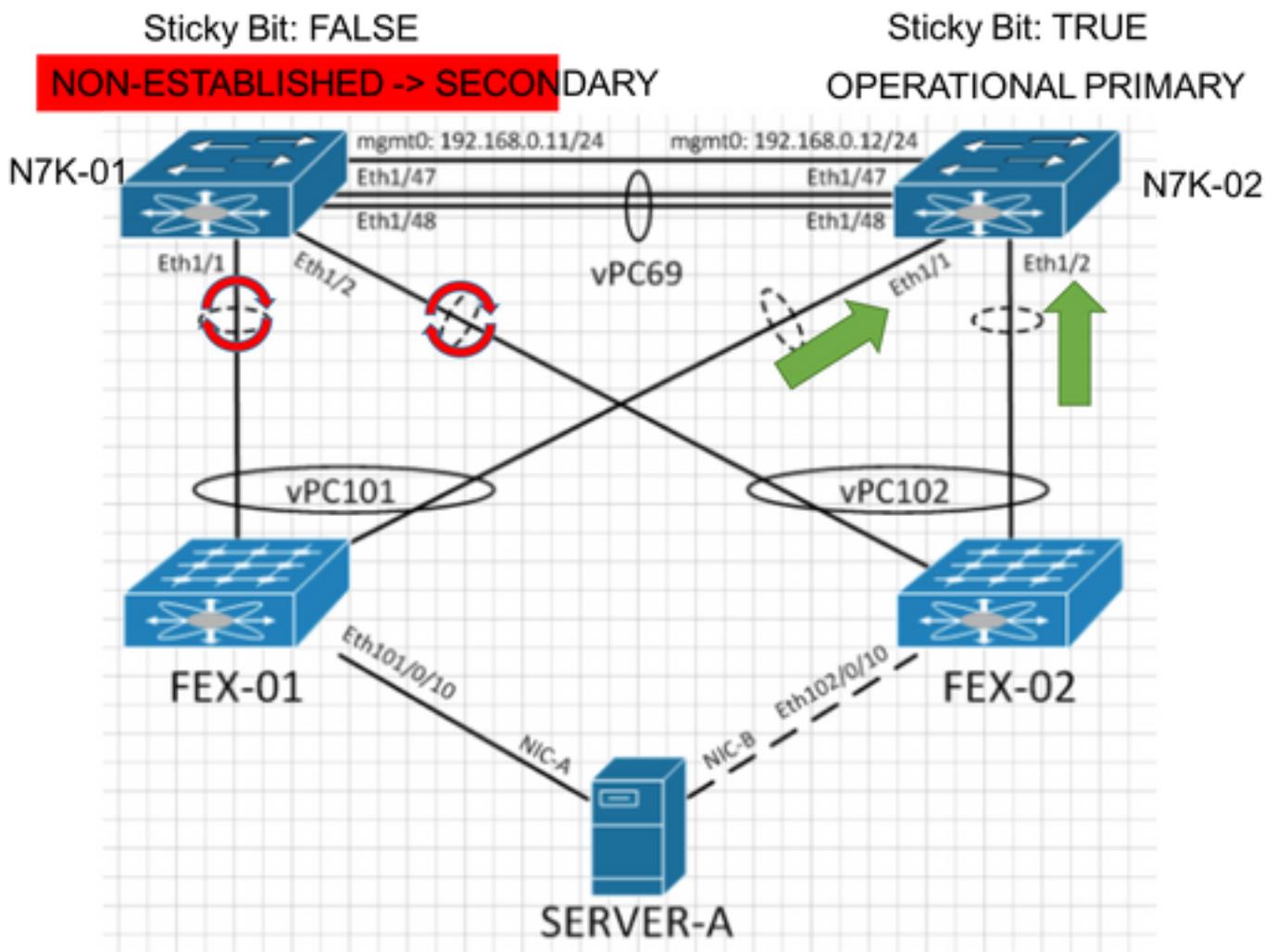


圖4

與錯誤設定粘滯位相關的網路中斷示例

當隔離交換機(Nexus-02)引入回VPC域時，由於粘滯位設定不正確，導致網路中斷

但是，如果在兩個Nexus交換機上未正確設定粘滯位，則隔離的交換機重新引入VPC域後可能會出現網路中斷。在將隔離交換機引入回VPC域之前，其粘滯位必須設定為FALSE。(有關更換N7K機箱的過程，請參閱[Nexus 7000機箱更換過程](#)。)

如圖5所示，Nexus-01配置的VPC角色優先順序高於Nexus-02，而Nexus-02的粘滯位設定為TRUE。Nexus-01的鏈路E1/1和E1/2處於轉發狀態，而E1/1和E1/2處於關閉狀態。

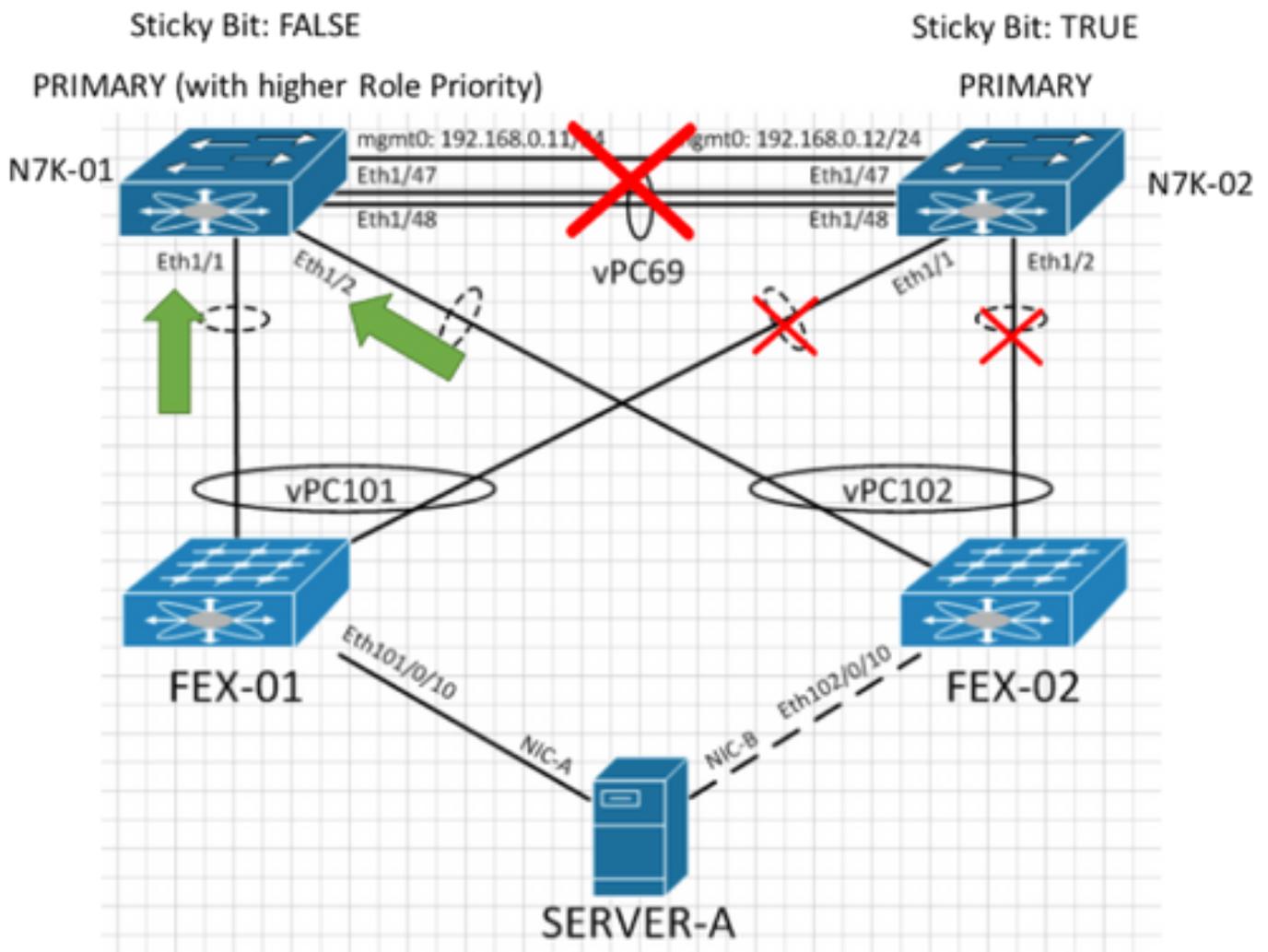


圖5

當PKA和對等鏈路恢復後，Nexus-02將扮演主角色，而不管其角色優先順序如何（因為它具有TRUE粘滯位），並強制Nexus-01成為輔助NEXUS，並且VPC初始化過程在Nexus-01上開始。因此，Nexus-01的鏈路E1/1和E1/2被VPC掛起，並在中繼恢復計時器（預設情況下為40秒）過期後進入線上狀態。在這種情況下，在PKA和對等鏈路恢復後會出現40秒的網路中斷，如圖6所示。

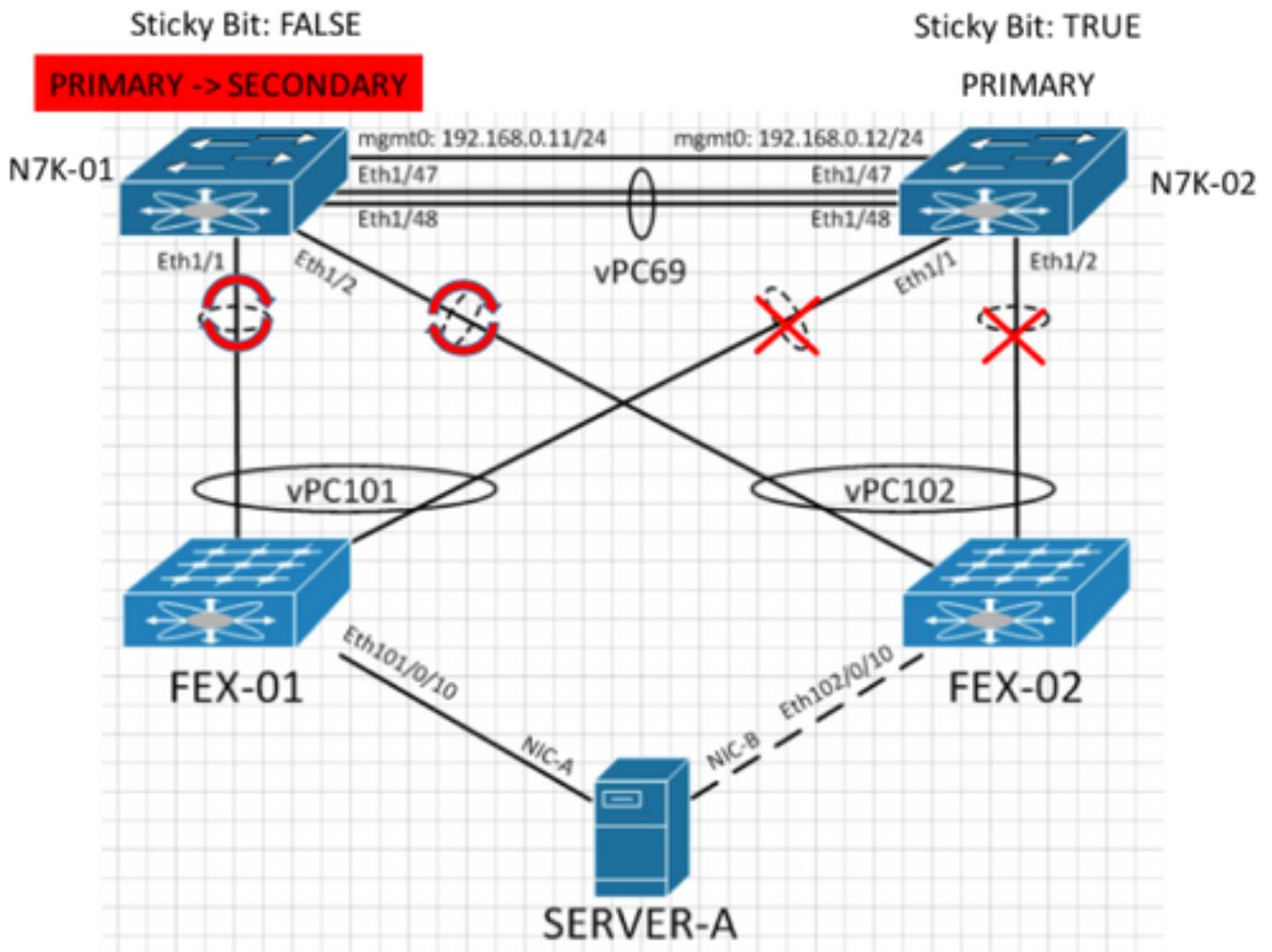


圖6

註：將Nexus重新引入vPC域時，必須確保活動vPC裝置中的vPC角色沒有更改。為避免將兩台交換機的粘滯位設定為相同值時更改vPC角色，活動vPC裝置必須具有更高的角色優先順序，才能保持其PRIMARY角色。有關VPC角色選舉過程的詳細資訊，請參閱本文檔中的圖1。

關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。