了解 vPC 选择过程

目录

[简介](#)

[先决条件](#)

- [要求](#)
- [使用的组件](#)

[虚拟 PortChannel 技术](#)

- [vPC 角色](#)
- [vPC 角色优先级](#)
- [无中断 vPC 角色更改](#)
- [vPC 对等链路关闭时的 vPC 系统行为](#)
- [vPC 主粘滞位](#)
- [vPC 延迟恢复](#)
- [vPC 延迟恢复接口 VLAN](#)
- [使用 4000 SVI 的扩展 SVI 设置时的 vPC 延迟恢复](#)

[vPC 选择过程](#)

[vPC 恢复场景](#)

[与错误设置粘滞位相关的网络中断示例](#)

简介

本文档介绍 Nexus 系列交换机上的虚拟端口通道 (vPC) 角色选择过程。

先决条件

要求

Cisco 建议您了解以下主题：

- Nexus 系列交换机上的 vPC
- 生成树协议 (STP)

使用的组件

本文档中的信息基于 Nexus 9000 系列交换机平台。

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

虚拟 PortChannel 技术

虚拟端口通道 (vPC) 可以使物理上连接到两台不同思科交换机的链路显示为连接到第三台设备的单一端口通道。第三台设备可以是交换机、服务器或支持 IEEE 802.3ad 端口通道的任何其他网络设备。vPC 还支持创建跨两台交换机的第 2 层端口通道。目前，vPC 已在 Cisco Nexus 9000、7000、5000 和 3000 系列平台（带或不带 Cisco Nexus 2000 系列交换矩阵扩展器）上实施。

注意：Cisco NX-OS软件vPC和Cisco Catalyst虚拟交换系统(VSS)具有类似技术。对于 Cisco EtherChannel 技术，术语“多机箱 EtherChannel (MCEC)”可互换使用，指代任一技术。

---

vPC 角色

尽管两台vPC交换机显示为通向下游设备的单个交换机，但其中两台vPC交换机具有明确的vPC角色：vPC 主设备和 vPC 辅助设备。

vPC 角色是非预占式角色，这意味着设备可以配置为 vPC 主设备，但作为 vPC 辅助对等设备运行。在以下场景中，可能会发生这种情况：

1. 当原先的主设备发生故障时，辅助 vPC 设备将成为新的主设备。
2. 当系统恢复运行时，之前的主设备现在成为辅助设备，反之亦然。

vPC 角色定义了两台 vPC 对等设备中的哪台设备处理网桥协议数据单元 (BPDU) 并响应地址解析协议 (ARP) 请求。vPC 角色还定义了 vPC 主设备和 vPC 辅助设备为响应 vPC 对等链路关闭情况而采取的一系列操作。

vPC 角色优先级

您还可以在vPC域模式命令中使用角色优先级来影响vPC选举过程。值范围为1至65636，默认值为32667。值越低，意味着此交换机越有可能成为主vPC。

更改 vPC 对等设备的优先级时，可能会导致网络中的接口启动和关闭。如果要再次配置角色优先级，以使一台 vPC 设备成为主设备，请在主 vPC 设备上配置具有较低优先级值的角色优先级，并在辅助 vPC 设备上配置具有较高优先级值的角色优先级。然后，关闭两台设备上的 vPC 对等链路并输入 shutdown 命令，最后重新启用两台设备上的端口通道并输入 no shutdown 命令。

无中断 vPC 角色更改

vPC 无中断角色更改功能提供了可在 vPC 对等体之间切换 vPC 角色而不会影响流量的框架。根据 vPC 域下设备的角色优先级值完成 vPC 角色交换。执行 vpc role preempt 命令时，选择角色优先级较低的 vPC 对等设备作为主 vPC 设备。

有关详细信息，请参阅无中断 vPC 角色更改的使用案例场景。

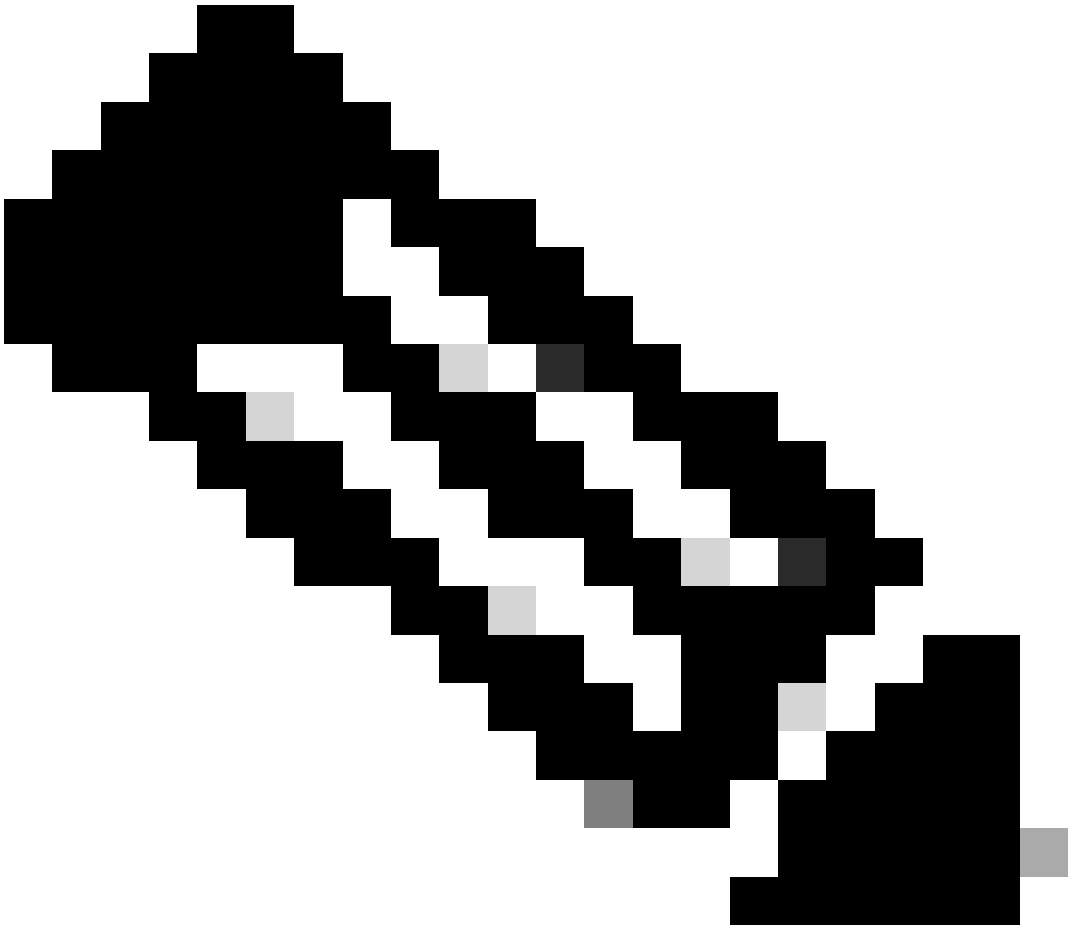
vPC 对等链路关闭时的 vPC 系统行为

当 vPC 对等链路出现故障而 vPC 对等保持连接链路仍处于开启状态时，vPC 辅助对等设备将执行以下操作：

1. 挂起其 vPC 成员端口.

2. 关闭与 vPC VLAN 关联的 SVI。

这种来自 vPC 的保护行为会将所有南北向的流量重定向到 vPC 主设备。



> 注意：当vPC对等链路断开时，两台vPC对等设备无法再相互同步，因此所设计的保护机制导致其中一个对等设备（发生时，即辅助对等设备）与数据路径隔离。

---

vPC 主粘滞位

vPC主粘滞位是一种编程保护机制，用于在主交换机意外重新加载时避免不必要的角色更改（可能导致网络中断）。vPC主粘滞位允许处于活动状态的交换机在死机交换机重新变为活动状态或隔离交换机重新集成到VPC域中时保持主交换机角色。

切换 vPC 主粘滞位：

1.在以下情况下，vPC主粘滞位值设置为TRUE:

- 当前 vPC 主设备重新启动，且支持 vPC 的交换机将其角色从“vPC 辅助设备”更改为“vPC 主

运行设备”。如果角色从 vPC 辅助运行设备更改为 vPC 主设备，则不会设置粘滞位。
- 当重新加载恢复计时器（默认为 240 秒）到期时，支持 vPC 的交换机将其角色从“未建立”更改为“vPC 主设备”。

2.在以下情况下，vPC主粘滞位值设置为FALSE:

- 支持 vPC 的交换机重新启动（粘滞位默认设置为 FALSE）。
- vPC 角色优先级被更改或重新输入。

vPC 主粘滞位在 vPC Manager 软件组件结构下报告，可以使用此 NX-OS exec 模式命令进行检查。

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

vPC 延迟恢复

vPC 对等设备重新加载并恢复后，路由协议需要时间重新收敛。在上行链路第 3 层连接重新建立之前，恢复中的 vPC 分支会将从接入层路由到汇聚层/核心层的流量转发到黑洞。

vPC延迟恢复功能延迟恢复的vPC对等设备上的vPC分支启动。vPC延迟恢复允许第3层路由协议在允许vPC支路上的任何流量之前进行收敛。这可实现更平稳的恢复，并在恢复阶段实现零丢包（流量仍会在活动的 vPC 对等设备上转移）。 此功能默认启用，vPC 恢复的默认计时器为 30 秒。可根据特定的第 3 层收敛基线在 1 秒至 3600 秒的范围内调整该计时器。

vPC 延迟恢复接口 VLAN

要延迟恢复的vPC对等设备上的VLAN接口启动，请使用delay restore命令的interfaces-vlan选项。默认情况下，此功能已启用，vPC恢复默认计时器为10秒。

使用 4000 SVI 的扩展 SVI 设置时的 vPC 延迟恢复

为了配置 pacer 以批量启动 VLAN 或网桥域接口（一次启动一批共 200 个 SVI），引入了新命令 delay restore interface-VLAN batch<1-4094>。可将 vPC 延迟恢复计时器命令 delay restore<Timeout value> 的值配置为大于所有已配置的批处理计时器的和。如此一来，仅当所有 SVI 都已完全启动后，vPC 分支才会启动，因此可以避免出现任何流量黑洞。

示例：

4000个Vlan，200个批处理，15秒延迟

delay restore > (4000/200)x 15

vPC 选择过程

在 vPC 系统中，根据这些参数和顺序，一台 vPC 对等设备定义为 vPC 主设备，一台定义为 vPC 辅助设备

1. vPC 主粘滞位设置为 0 或 1。
2. 用户定义的 vPC 角色优先级（Cisco NX-OS 软件按照最小数值选择主设备）。
3. 系统 MAC 地址值（Cisco NX-OS 软件按照最低 MAC 地址选择主设备）。

此流程图（图 1）总结了两台 vPC 对等设备在 vPC 主交换机选择过程中经历的步骤。

1. 在 vPC 主设备选择过程中，两台设备之间的第一个检查参数是 vPC 主粘滞位。如果 vPC 对等设备在比较中获胜，它将成为 vPC 主设备，而不考虑两台对等设备已配置的 vPC 角色优先级值或系统 MAC 地址。
2. 如果两台 vPC 对等交换机的粘滞位值相同，则选择过程将继续执行下一步，比较用户定义的 vPC 角色优先级。
3. 如果两个 vPC 角色配置为相同的值，则选择过程将继续比较系统 MAC 地址。

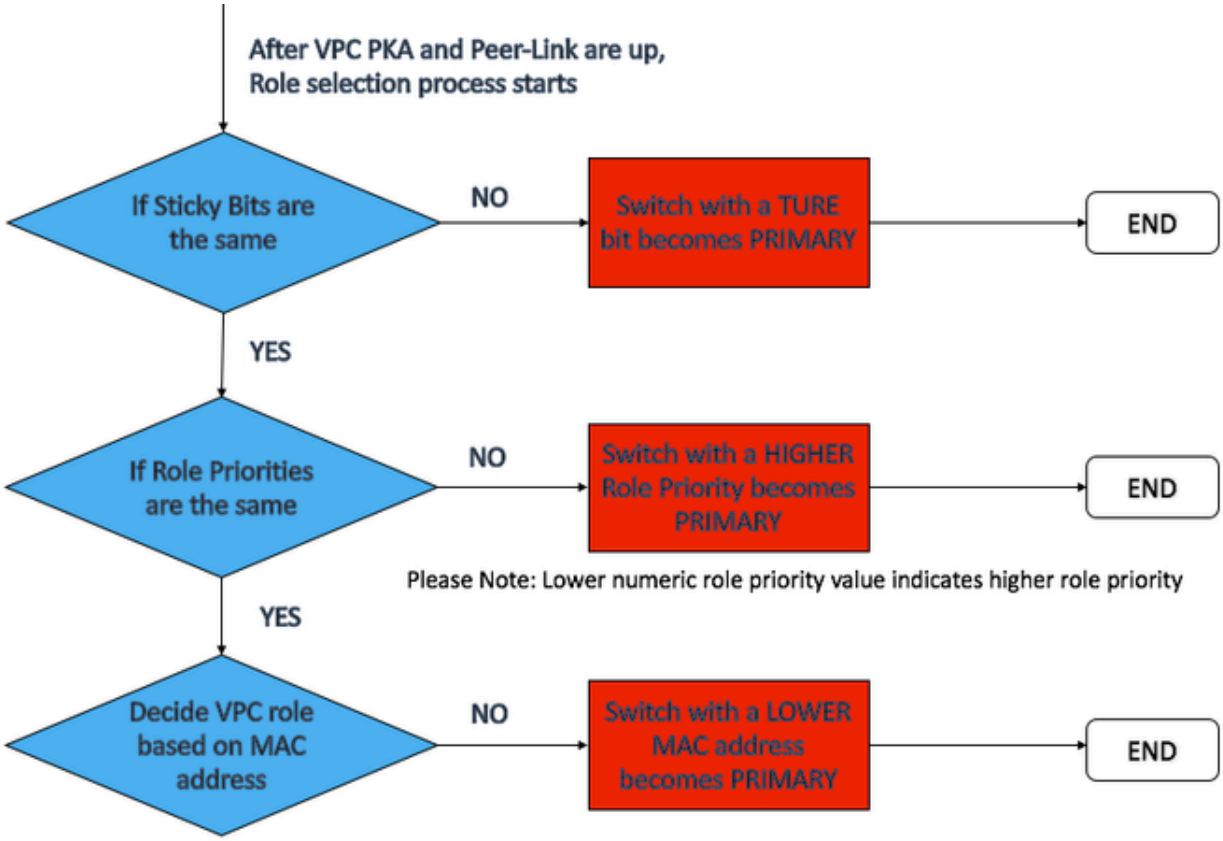


图 1

如图所示，当 vPC 交换机将 vPC 主粘滞位设置为 1（TRUE 条件）而其对等体的粘滞位设置为 0（FALSE 条件）时，TRUE 端将在选择中获胜，并承担 vPC 主设备角色。

vPC 对等设备 1 粘滞位设置为 1	vPC 对等设备 2 粘滞位设置为 1	vPC 主设备
False (0)	False (0)	Tie
True (1)	False (0)	vPC 对等设备 1
False (0)	True (1)	vPC 对等设备 2

True (1)	True (1)	Tie

vPC 恢复场景

了解 vPC 选择过程十分重要，不可轻忽，尤其是在 vPC 恢复场景中。

图 2 显示了一个典型的 vPC 设置，其中 Nexus-01 是 vPC 主设备，Nexus-02 是 vPC 辅助设备。两台设备都默认将其粘滞位重置为 FALSE。

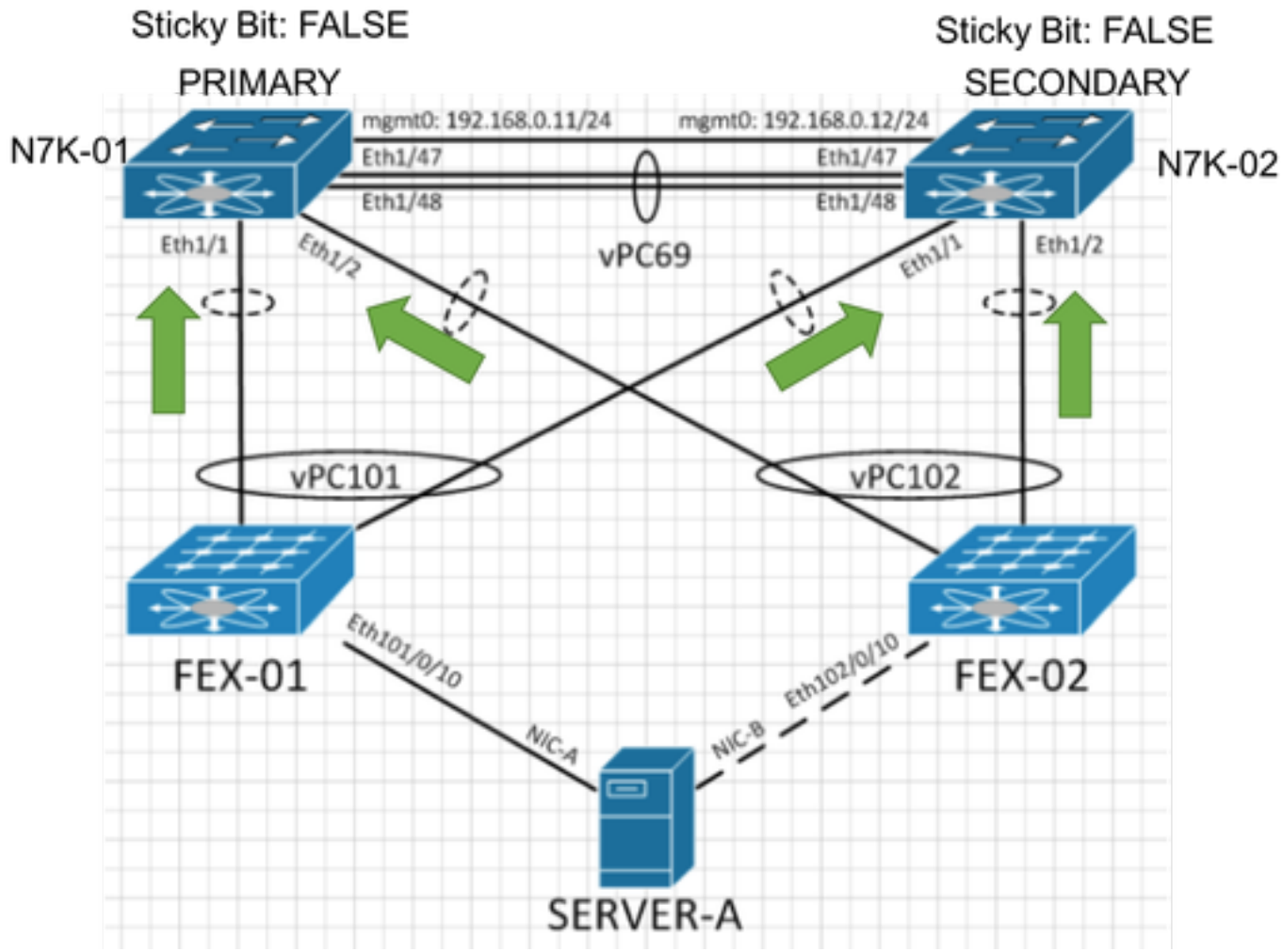


图 2

如上图所示，Nexus-01 现在电源中断，并已与网络隔离。Nexus-02将其自身提升为vPC主设备，并将vPC粘滞位设置为TRUE。Nexus-02现在变为主设备运行，粘滞位现在设置为TRUE。

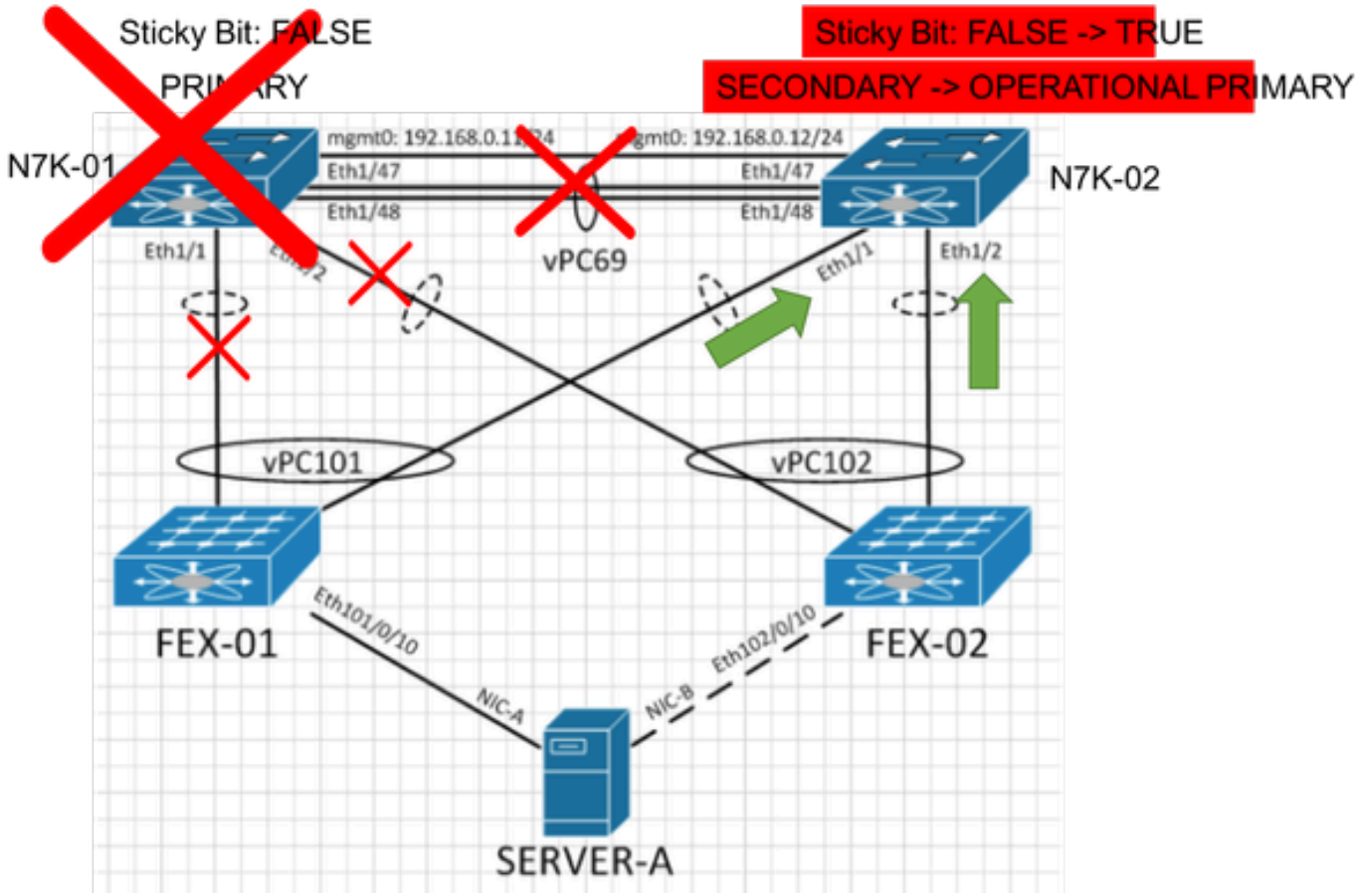


图 3

如上图所示，当 Nexus-01 在断电恢复后恢复在线状态时，Nexus-02 无论角色优先级如何都将保留“主运行设备”角色（因其粘滞位为 TRUE），而 Nexus-01 在线时承担“辅助设备”角色。只有Nexus-01才开始VPC初始化过程，而Nexus-02仍保持主用状态，并照常转发流量。因此，不会出现网络中断。

Nexus-01 上有两个计时器与 vPC 初始化过程相关联，Nexus-01 现在是 vPC 辅助运行设备：

- delay restore SVI（默认为 10 秒）
- delay restore（默认为 30 秒）

因此，在 Nexus-01 作为 vPC 辅助设备被重新引入网络后，您可以在 Nexus-01 上获得 40 秒的恢复时间。但是，由于 Nexus-02 承担“主设备”角色，现在所有流量都会流经 Nexus-01（如上所述），因此不会出现网络中断。

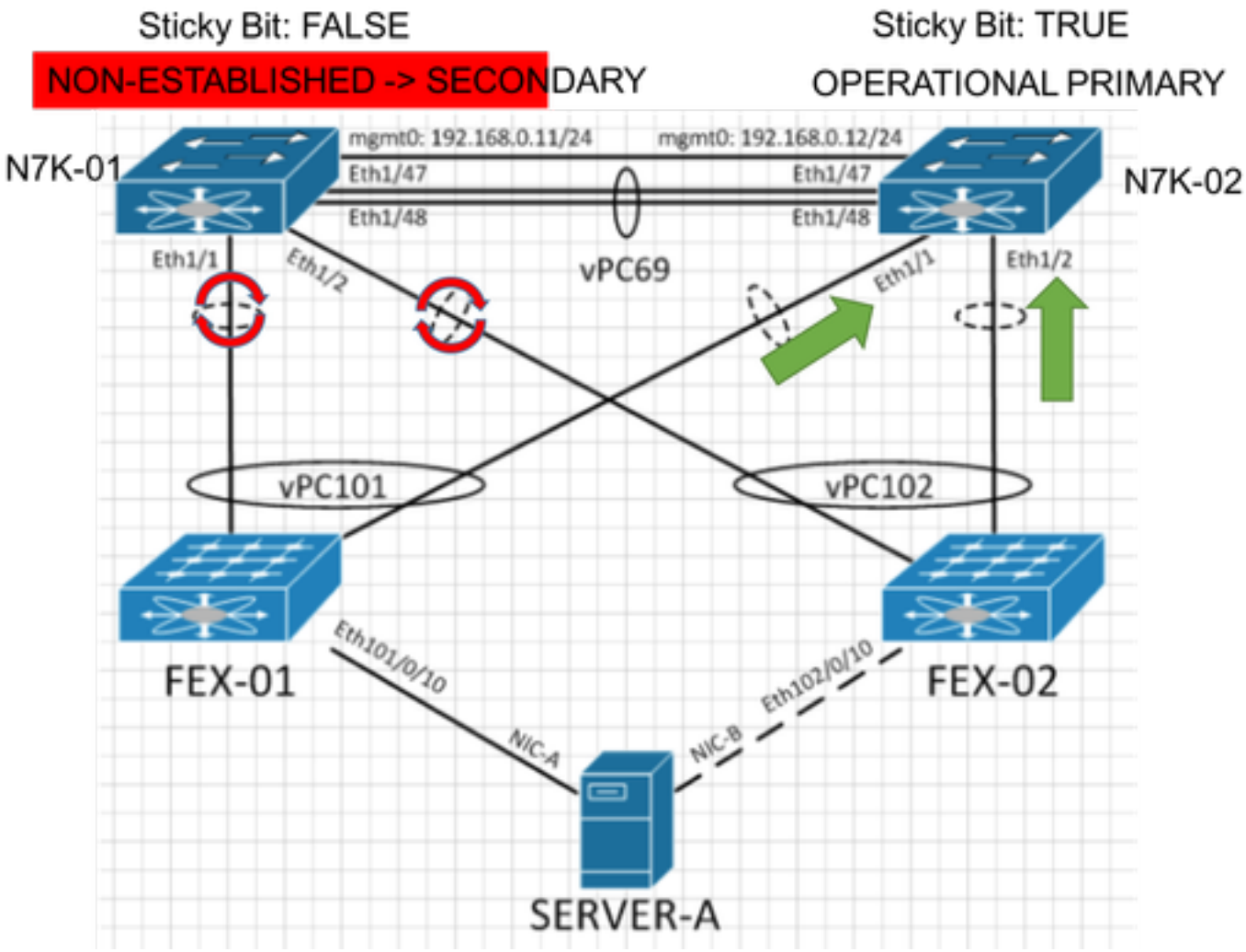


图 4

与错误设置粘滞位相关的网络中断示例

当隔离交换机 (Nexus-02) 重新被引入 vPC 域时，粘滞位设置错误会导致网络中断

但是，如果未在两台 Nexus 交换机上正确设置粘滞位，则在将隔离交换机重新引入 vPC 域后，可能会发生网络中断。在将隔离交换机重新引入 vPC 域之前，必须将其粘滞位设置为 FALSE。（有关更换 N7K 机箱的程序，请参阅 Nexus 7000 机箱更换程序。）

如图 5 所示，Nexus-01 配置的 vPC 角色优先级高于 Nexus-02，Nexus-02 的粘滞位设置为 TRUE。Nexus-01 的链路 E1/1和 E1/2 处于转发状态，而链路 E1/1 和 E1/2 处于关闭状态。

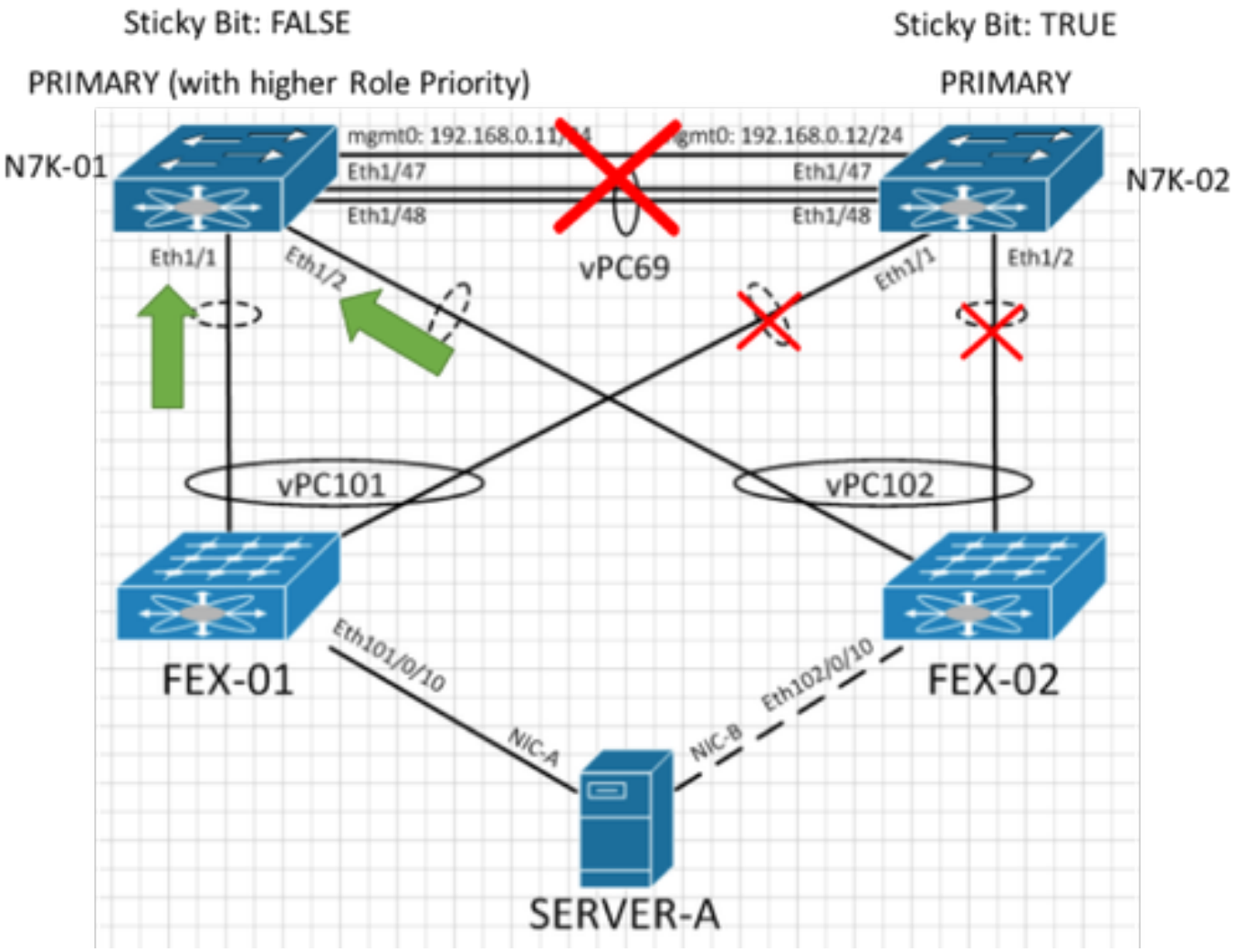


图 5

当PKA和对等链路恢复后，Nexus-02将承担主要角色，而不考虑其角色优先级(因为它具有TRUE sticky bit)，并强制Nexus-01成为SECONDARY，VPC初始化过程在Nexus-01上开始。因此，Nexus-01的链路E1/1和E1/2被VPC挂起，并在中继恢复计时器（默认情况下为40秒）到期后联机。 在这种情况下，当 PKA 和对等链路恢复后，会出现 40 秒的网络中断，如图 6 所示。

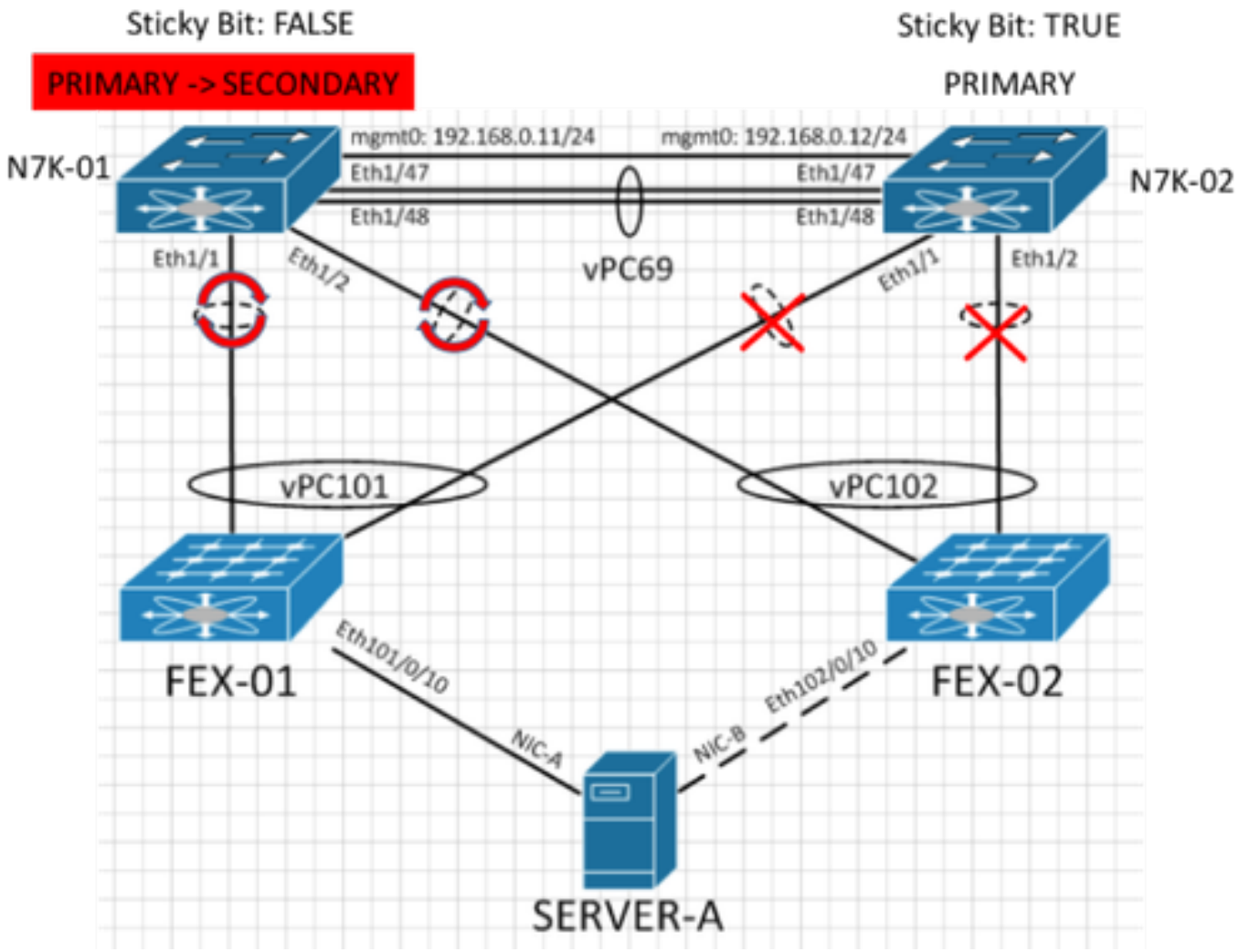


图 6

注意：当Nexus重新引入到vPC域时，必须确保活动vPC设备中的vPC角色没有更改。为了避免两台交换机的粘滞位设置为相同值时 vPC 角色更改，主 vPC 设备必须具有更高的角色优先级，才能保留其“主设备”角色。有关 vPC 角色选择过程的详细信息，请参阅本文档中的图 1。

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。