对具有重复NAT终端的SD-WAN集线器隧道进行故障排除

目录

简介

先决条件

- 要求
- 使用的组件

背景信息

- 控制平面NAT学习
- 数据平面NAT学习和VCONTROL
- 端口保留、端口重叠和端口随机化

拓扑示例

冲突发生的原因

- 不同的vBonds可以观察到相同的公共端点
- 同一个通告到达一个集线器
- 重复的分支机构地址发生私有 — 公共冲突
- NAT终端重新分配

症状和影响

诊断问题

- 步骤1.记录软件和分支传输状态
- 步骤2.在OMP中查找重复的终端
- 步骤3.比较逻辑TLOC与中心会话所有权
- 步骤4.检验NAT网关实际执行的操作
- 步骤5.检查Wireshark中的VCONTROL
- 步骤6.确认受影响路径并保留证据

解决方法和限制

- 使用具有唯一分配的稳定独立于终端的映射
- 禁用不同辐射点之间的端口重叠
- 对小型受控组使用端口偏移
- 禁用源端口保留或启用随机分配
- 单独的公有地址或保留端口块
- 保留备用传输并协调恢复

相关软件缺陷

附录：解码IPv4 VCONTROL报尾

相关信息

简介

本文档介绍当分支共享一个公共网络地址转换(NAT)端点时如何识别和缓解中心隧道故障。

先决条件

要求

Cisco 建议您了解以下主题：

- Cisco Catalyst软件定义的广域网(SD-WAN)架构和控制连接
- 重叠管理协议(OMP)、传输定位器(TLOC)、双向转发检测(BFD)和IP安全(IPsec)
- 多协议标签交换(MPLS)传输网络
- NAT映射和过滤行为
- 使用Wireshark进行基本数据包分析

使用的组件

本文档中的信息基于以下软件和硬件版本：

- 思科IOS® XE Catalyst SD-WAN路由器
- Cisco Catalyst SD-WAN控制器(vSmart)
- Cisco Catalyst SD-WAN验证器(vBond)

中心会话所有权故障在Cisco IOS XE 17.18.2上重现，成对密钥可运行。命令可用性和输出字段因平台和版本而异。

示例使用私有地址或RFC 5737保留的文档地址，以及通用设备名称。CLI摘录将缩短并使用规范化的标识符。接口名称和某些输出字段因平台和版本而异。

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。

⚠ 警告：配置更改和连接重置可能会中断流量。使用批准的维护计划。初始诊断命令是只读的。

背景信息

从正确识别对称NAT的场景开始。分支A注册到vBond 1，分支B注册到vBond 2。如RFC 7857所允

许，当这些连接的目的地不同时，NAT网关可以为它们分配相同的转换源IP地址和端口。因此，即使两个分支都报告NAT类型A，它们都可以为其公共互联网TLOC通告相同的公共终端。

示例拓扑使用公共Internet颜色。同一场景可能涉及其他公共颜色，例如商业互联网。将诊断过滤器应用到实际受影响的传输颜色；字面公共internet颜色不是必需条件。

注意：Cisco Catalyst SD-WAN不支持来自不同站点的TLOC，这些站点向同一个接收边缘通告相同的公共IP地址和UDP端口。在此冲突条件下，接收方的行为是不可预测的。即使正确识别了对称NAT，也可能发生这种情况。仅对NAT进行正确分类并不能保证通告的终端是唯一的。

控制平面NAT学习

在vBond注册过程中，边缘会发送其专用传输地址和端口。vBond将观察该连接上的源地址和端口，并将观察的终端返回到边缘。边缘可以通过OMP将此信息通告为TLOC的公有地址和端口属性，包括具有公共颜色（例如公共互联网或商业互联网）的TLOC。

该观察描述了与该vBond的连接。使用与目的地相关的NAT，到集线器的映射可以不同。路由器会比较NAT观察结果，以便对映射行为进行分类。在show sdwan control local-properties中，A表示依赖于地址和端口的映射，E表示独立于终端的映射。图中所示的分类描述了路由器的决策；nat转换表或数据包捕获会验证NAT网关的实际行为。

例如，运行show sdwan control local-properties并找到传输的NAT TYPE列。此缩短的输出保留了NAT类型图例和选定的接口列，并具有规范化的地址：

<#root>

Spoke-A# show sdwan control local-properties

```
...
 NAT TYPE: E -- indicates End-point independent mapping
           A -- indicates Address-port dependent mapping
           N -- indicates Not learned
           Note: Requires minimum two vbonds to learn the NAT type

                          PUBLIC         PUBLIC PRIVATE         PRIVATE                          NAT
INTERFACE                 IPv4           PORT   IPv4            PORT    COLOR          STATE     TYPE
GigabitEthernet2          192.168.9.11    12346  192.168.9.11    12346   mpls           up        E
GigabitEthernet3          203.0.113.10    23012  192.168.11.11   12346   biz-internet   up        A
```

商务Internet传输报告A并通告203.0.113.10:23012作为其公共终端。MPLS传输报告E，具有相同的私有和公共终端。从NAT TYPE中读取分类；仅地址或端口相等并不能识别NAT类型。N表示路由器尚未获知分类。

数据平面NAT学习和VCONTROL

集线器必须先找到对应对等体和隧道的入站IPsec安全关联(SA)，然后才能处理加密的BFD数据包。SA标识加密密钥和验证和解密数据包所需的其他状态。接收器使用终端信息和封装安全负载(ESP)安全参数索引(SPI)来定位该状态。SPI本身不是全局唯一的分支标识符。

对称NAT后的分支通常向可达中心发起流量。这将为该特定目标创建一个映射。集线器看到的源地址或端口可能与之前通过vBond获知并在OMP中通告的公共终端不同。因此，收到的外部报头并不总是与已安装隧道状态下的终端匹配。集线器需要一种方法将此新的在线终端与已知的远程TLOC及其IPsec状态相关联，然后才能解密BFD。

对于此处讨论的IPv4 BFD-over-IPsec格式，SD-WAN将终端信息放在UDP负载末尾的可见VCONTROL报尾中，不在加密BFD负载之内。集线器可以在解密前读取此信息。VCONTROL传输端点信息而不是加密密钥，并且自身不建立对等体身份。

例如，在初始隧道建立期间，数据平面学习完成之前：

```
Spoke B advertised through OMP:       203.0.113.10:12346
Spoke B observed at the hub:          203.0.113.10:45678
VCONTROL carried source endpoint:    203.0.113.10:12346
```

接收方首先尝试其正常终端/SPI查找。如果查找未找到SA，则VCONTROL源终端提供回退。在本示例中，集线器可以使用承载的203.0.113.10:12346及其本地目标环境和SPI，以便找到与已知远程TLOC关联的入站SA，即使数据包从203.0.113.10:45678到达。

所选SA提供接收密钥上下文。SPI会选择在该情景中适用的密钥生成。然后，集线器执行IPsec身份验证和解密；只有通过所需检查的数据包才可用于BFD处理。仅可读的尾部或成功的查找并不能证明身份验证、解密或BFD接收是成功的。

因此，观察的源端口和携带的源端口之间的差异可以使用对称NAT正常。要使对等关系正常工作，集线器必须将数据包与正确的隧道和密钥状态相关联，并了解实际的返回端点。在本例中，成功获知允许集线器将返回流量发送到203.0.113.10:45678。这不同于通过vBond进行的控制平面学习。

VCONTROL不会强制接收方覆盖现有终端/SPI匹配。如果重复的终端状态将数据包与错误的对等体或密钥生成相关联，则帧尾不保证恢复。它提供查找提示；它无法区分没有正确周围隧道和加密状态而承载同一端点的两个对等体。

当报告NAT类型E时，适用的BFD数据包上也存在VCONTROL。它的存在不是对称NAT标志，并且不包括在每个普通IPsec数据包中。对于此处介绍的IPv4 BFD格式，在两个专用颜色之间省略公共

TLOC信息。

通过数据平面了解的对等特定终端并不意味着分支必须为每个目标使用新终端替换其OMP通告。不同获知的映射是对称NAT遍历的预期部分。

隧道启动后，获知该隧道的自身转换端点的发送方可以将该端点作为源信息传送。这是发送方的每隧道公共终端，而不是远程终端或新的全局OMP通告。始终比较来自同一观察期的数据包、OMP通告和会话状态。

端口保留、端口重叠和端口随机化

以下是不同的行为：

行为	含义
源端口保留	NAT网关尝试保留分支的原始UDP源端口（例如12346）作为转换后的源端口。
目标之间的端口重叠	NAT网关会重复使用一个转换后的IP:用于与不同外部目标通信的不同内部端点的端口。
端口随机化	NAT网关使用随机分配方法选择转换端口。
独立于终端的映射	当映射存在时，一个内部端点在外部目标之间保留相同的转换后端点。

RFC 7857第3部分允许在目标不同时重叠不同内部端点的外部端口。NAT使用流的五元组来区分流。这不允许两个无法区分的同步UDP流到达同一个目的地。随机分配和保留是独立的分配选择；二者均不建立终端独立性。

拓扑示例

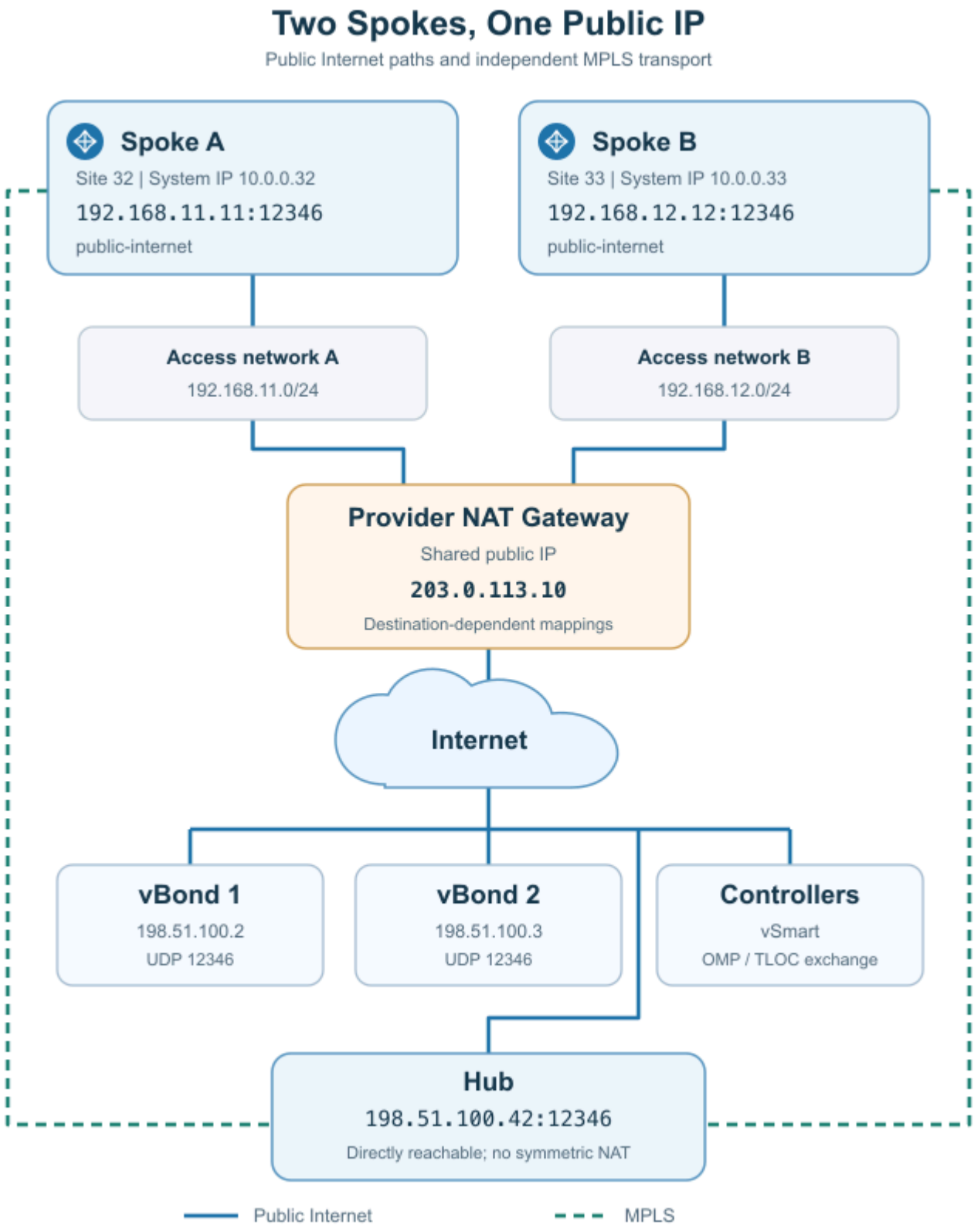

图1.共享NAT网关、独立的分支接入网络和独立的MPLS传输。

分支A使用192.168.11.11:12346，分支B使用192.168.12.12:12346。两者通过NAT网关共享公有地址203.0.113.10。集线器有一个可直接到达的地址，不在对称NAT的后面。两个分支可以通告相同的公共IP：端口，即使它们的专用终端不同。

冲突发生的原因

不同的vBonds可以观察到相同的公共端点

请考虑以下同步UDP映射：

对等体	私有源	转换后的源	目的地
A	192.168.11.11:12346	203.0.113.10:12346	vBond 1, 198.51.100.2:12346
B	192.168.12.12:12346	203.0.113.10:12346	vBond 2, 198.51.100.3:12346

由于NAT也会检查远程终端，因此回复会正确返回。vBond 1的回复属于A的映射；vBond 2的回复属于B的映射。

两个路由器也可以联系其他vBond并正确检测与目的地相关的映射。例如，请保持每个辐条的专用源元组不变，并考虑以下观察顺序：

对等体	首次接受的虚拟绑定观察	与其他vBond的比较	结果
A	vBond 1返回 203.0.113.10:12346	vBond 2返回 203.0.113.10:40001	NAT A;保留公共元组 203.0.113.10:12346
B	vBond 2返回 203.0.113.10:12346	vBond 1返回 203.0.113.10:40002	NAT A;保留公共元组 203.0.113.10:12346

每个路由器的vBond观察结果不同，因此检测NAT A是正确的。尽管如此，所保留的第一个观察结果在两台路由器上仍然相符。分类A不会在所有分支之间分配唯一公共端口，访问两个vBonds也不能保证唯一性。

每个分支都接收注册应答中由vBond观察到的公有地址和端口，并使用保留的观察填充其本地传输属性。分支在其通过OMP向SD-WAN控制器通告的TLOC属性中包括公共终端及其专用终端。然后，控制器根据适用的控制策略将TLOC分配到其他Edge;vBond本身并不发起OMP通告。

如果两个分支保留相同的观察公共端点，则OMP可以包含：

```
TLOC system IP   Color            Public endpoint        Private endpoint
10.0.0.32        public-internet   203.0.113.10:12346     192.168.11.11:12346
10.0.0.33        public-internet   203.0.113.10:12346     192.168.12.12:12346
```

由于它们的系统IP不同，因此这些保持不同的逻辑TLOC。

同一个通告到达一个集线器

NAT网关必须区分两个分支，当它们联系同一个集线器时。一个可能的分配是：

对等体	转换为集线器的源	目的地
A	203.0.113.10:12346	198.51.100.42:12346
B	203.0.113.10:45678	198.51.100.42:12346

集线器有两个通告相同目标端点的逻辑对等体，而数据包可以来自不同的实际端点。在正常对称NAT操作下，数据平面学习有助于建立返回映射。

当不同的对等体共享此传输密钥选择的会话状态时，会出现集线器会话所有权限制：

```
Hub-local IP + hub-local port + remote IP + remote port + protocol
```

如果两个通告都选择相同的密钥，则来自后续对等体的更新可以替换共享状态的所有权信息。一个对等体的会话随后退出、删除或重新创建可能会干扰另一个对等体的IPsec/BFD状态。当该状态已经与错误的对等体或流生成相关联时，数据平面学习无法保证恢复。

可见尾部不是独立的对等标识符：两个分支可以承载相同的源端点。如果集线器选择属于错误对等体或密钥生成的现有IPsec状态，则即使数据包到达集线器，身份验证和学习也可能失败。为了建立特定数据包的顺序，思科技术支持中心(TAC)必须将所选SA与其对等设备和生成相关联。仅数据包报尾无法证明这一点。

仅相同的公有IP不会造成冲突。不同的远程端口通常产生不同的传输密钥。还要检查所选的集线器本地地址、本地端口和协议。如果第三个对等体通告相同的端点，它还可以竞争该密钥；故障模式取决于事件顺序。

重复的分支机构地址发生私有 — 公共冲突

Cisco bug ID CSCwx12986可使边缘在NAT比较期间自动源端口跳之后在其公共TLOC属性中通告专用RFC 1918元组。如果不同的分支重复使用相同的私有地址和端口，这些不正确的公共属性可能会在对等体之间匹配并暴露集线器会话所有权问题。

例如，两个分支都可以在单独的分支机构防火墙后面使用192.168.11.11。每个防火墙都将其分支机构网络分隔开，并转换源地址。如果上游提供程序NAT网关再次转换它，则路径具有两个NAT阶段或两个NAT。独立的虚拟路由和转发(VRF)环境是隔离重叠分支地址的另一种方式。

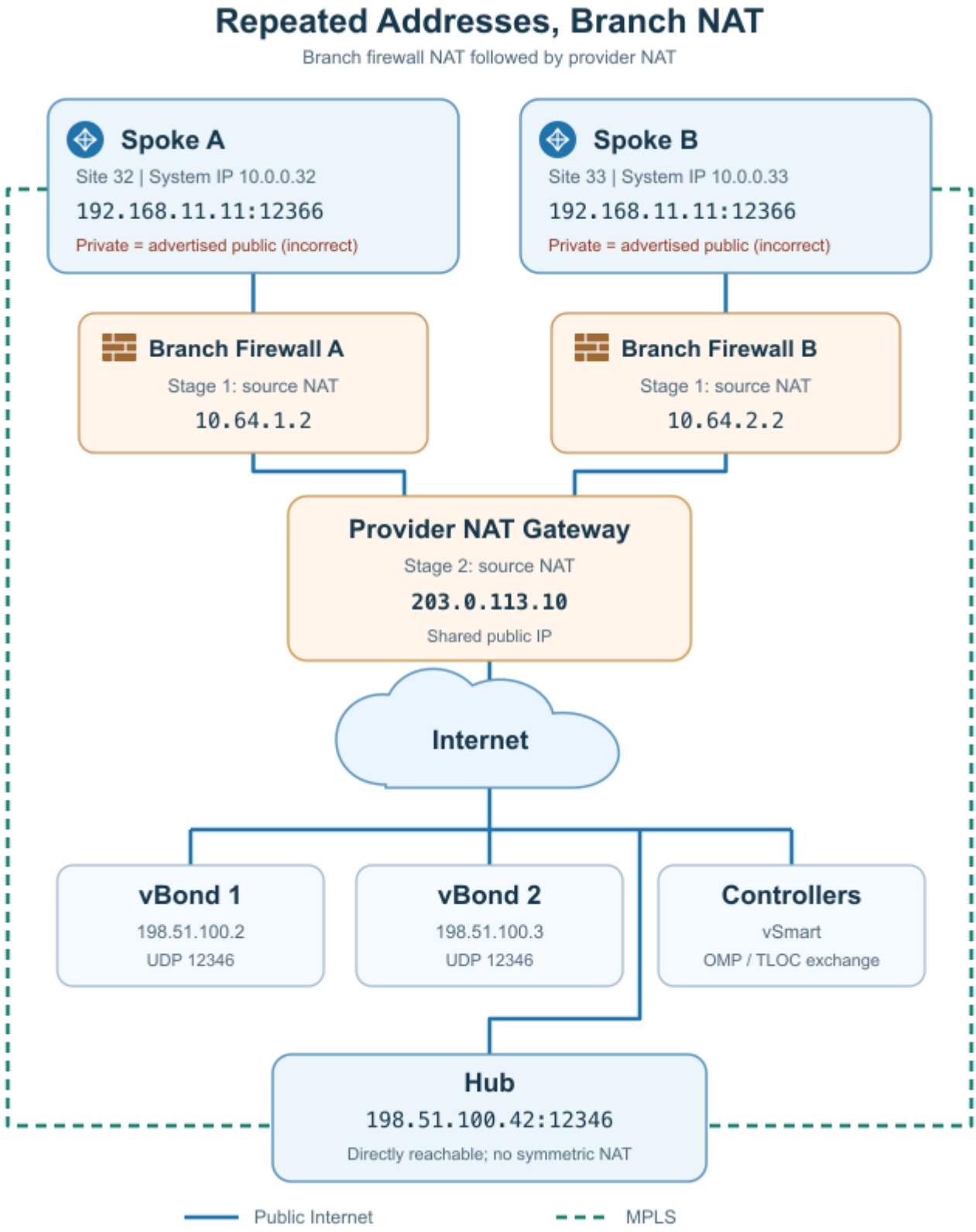

图2.分支机构防火墙和提供商NAT转换。不正确的公共TLOC属性保留原始专用端点，而数据包使用转换后的源地址。

如果两个分支在其公共互联网TLOC上遇到缺陷并使用相同的后跳源端口，集线器可以接收以下通告：

对等体	端口跳后的专用终端	公共TLOC终结点不正确
A	192.168.11.11:12366	192.168.11.11:12366
B	192.168.11.11:12366	192.168.11.11:12366

注意：重复私有地址和双NAT不是思科漏洞ID CSCwx12986本身的必备条件。额外的跨对等冲突需要匹配通告的IP：端口值和相同的集线器本地地址、端口和协议。具有不同通告端口的相同私有IP不会建立此冲突。

在没有NAT的传输中，公有和私有地址相等是正常现象。公共字段中的RFC 1918地址只有在与预期路径和观察到的转换相抵触时才可疑。唯一的分支专用地址可避免本示例中的重复私有元组重叠；它们不会阻止前面介绍的单独转换公共端口重用场景。

NAT终端重新分配

在NAT映射到期后，当中心点仍保留旧的对等体状态时，也可以将终端重新分配给另一个分支。将映射历史记录与中心会话生命周期进行比较，以确定新旧所有者之间的重叠。

症状和影响

典型指标包括：

- 当受影响的BFD路径关闭或无法恢复时，控制连接保持运行。
- 两个不同的系统IP通告相同的公共IP：端口。
- 集线器保留两个逻辑TLOC，但过滤的平台会话视图缺少一个对等体，或显示不一致的系统IP、远程颜色、鉴别器、隧道或SA信息。
- 一个对等体的传输撤销或重新创建会干扰具有相同端点的另一个对等体。
- 受影响路径上的IPsec输入丢弃可能会增加。

此故障可能导致受影响的传输路径不可用。如果这是分支的唯一可用传输路径，则该分支与重叠隔离。在大型网络枢纽中，共享一个NAT地址的多个分支会增加需要关联的数量，但必须测量受影响路径的数量；无法从该地址后面的站点数量推断出该地址。

诊断问题

步骤1.记录软件和分支传输状态

在每个受影响的Cisco IOS XE Edge上，将<affected-color>替换为实际传输颜色，例如biz-internet或public-internet:

<#root>

show clock detail

show version | include Cisco IOS XE Software

show sdwan control local-properties

show sdwan control connections | include

show sdwan omp peers

记录私有/公有地址以及端口、接口、颜色、NAT类型、站点ID和控制状态。在第一个完整的本地属性示例之后，使用此压缩过滤器来保留接口标题和所选的传输行。将TenGigabitEthernet0/1/0.10替换为受影响的接口：

<#root>

show sdwan control local-properties "wan-interface-list" | include INTERFACE|PUBLIC|TenGigabitEthernet0/

对于主场景，两个辐射点上的相关字段均可读取。这是字段摘要，而不是本地列布局：

```
Field                        Spoke A                 Spoke B
System IP                    10.0.0.32               10.0.0.33
Color                        public-internet         public-internet
Private endpoint             192.168.11.11:12346      192.168.12.12:12346
Public endpoint              203.0.113.10:12346       203.0.113.10:12346
NAT type                     A                       A
vSmart controls on transport Up                      Up
```

Up控制连接验证该控制器关系。它不会验证分支到集线器BFD路径或公共端点的唯一性。

步骤2.在OMP中查找重复的终端

在Cisco IOS XE集线器上：

<#root>

show clock detail

show sdwan omp tlocs table | include 203.0.113.10

在SD-WAN控制器(vSmart)上：

<#root>

show clock

show omp tlocs | tab | include 203.0.113.10

此控制器输出使用规范化的标识符。省略IPv6和区域列：

<#root>

vSmart# show omp tlocs | tab | include 203.0.113.10

```
TLOC IP     COLOR            ENCAP FROM PEER  STATUS KEY PUBLIC IP     PORT  PRIVATE IP     PORT
10.0.0.32   public-internet ipsec 10.0.0.5   C,R    1   203.0.113.10 12346 192.168.11.11 12346
                               10.0.0.32  C,I,R  1   203.0.113.10 12346 192.168.11.11 12346
10.0.0.33   public-internet ipsec 10.0.0.5   C,R    1   203.0.113.10 12346 192.168.12.12 12346
                               10.0.0.33  C,I,R  1   203.0.113.10 12346 192.168.12.12 12346
```

比较TLOC系统IP、颜色、封装、公共IP和公共端口。从不同控制器到同一逻辑TLOC的多条路径是正常的。冲突候选是具有相同有效端点的两个不同的对等身份。

在30-60秒后重复过滤的命令，并保留两个时间戳。持续的重叠是有用的证据；完全同步的命令不是必需的。对于短时间瞬变，请通过TAC收集OMP事件和相应的会话生命周期记录。

对于两个可疑对等体，保留其各自的TLOC表条目：

<#root>

```
show sdwan omp tlocs table | include 10.0.0.32
```

```
show sdwan omp tlocs table | include 10.0.0.33
```

要获得有关Cisco IOS XE的IPv4 TLOC详细信息的精简视图，请使用此过滤器。用所调查的对等设备替换系统IP，并保留ip <system-IP>选择器周围的引号：

<#root>

```
show sdwan omp tlocs "ip 10.0.0.32" | exclude not set| 0|None|efault|::
```

```
show sdwan omp tlocs "ip 10.0.0.33" | exclude not set| 0|None|efault|::
```

比较两个对等体的公有和私有地址/端口属性。过滤器会忽略此视图中的匹配未设置、零值、默认和IPv6行。

步骤3.比较逻辑TLOC与中心会话所有权

检查传送到FTM的TLOC状态。比较两个对等体的系统IP、颜色、公共IP地址和公共UDP端口。

<#root>

```
show platform software sdwan ttmd tlocs
```

运行以下命令以比较转发表管理器(FTM)、BFD和平台会话记录：

<#root>

```
show sdwan ftm bfd tloc | include 10.0.0.32
```

```
show sdwan ftm bfd tloc | include 10.0.0.33
```

```
show sdwan bfd sessions | include 10.0.0.32

show sdwan bfd sessions | include 10.0.0.33

show platform software sdwan session dstip 203.0.113.10

show platform software sdwan session remote-sysip 10.0.0.32 detail

show platform software sdwan session remote-sysip 10.0.0.33 detail
```

此示例显示具有规范化标识符的模式。为清楚起见，列标签会添加到已过滤的FTM行；它们不是已过滤命令输出的一部分：

<#root>

```
Hub# show sdwan ftm bfd tloc | include 10.0.0.32

System IP  Remote color     Encap  TLOC index  State
10.0.0.32  public-internet  IPSEC  32827       DOWN

Hub# show sdwan ftm bfd tloc | include 10.0.0.33

System IP  Remote color     Encap  TLOC index  State
10.0.0.33  public-internet  IPSEC  32828       DOWN

Hub# show platform software sdwan session dstip 203.0.113.10

RemoteSysIP Color        Proto SrcIp         SPort DstIp        DPort DPubIp       PPort BFD-LD TUN-ID
10.0.0.32   mpls         IPSEC 192.168.9.42  12346 203.0.113.10 12346 203.0.113.10 12346 20184  798
10.0.0.32   biz-internet IPSEC 192.168.10.42 12346 203.0.113.10 12346 203.0.113.10 12346 20183  799
```

OMP和FTM保留两个逻辑对等身份，而此会话查找仅显示共享目标的分支A。分支A也具有下行FTM状态；平台会话表中的状态不表示正常路径。FTM TLOC索引与会话表中的BFD本地标识符(BFD-LD)不同。此模式是调查各层所有权的原因之一。仅缺少CLI行无法建立数据包丢失。

在此表中，Color表示本地传输；在详细视图中使用Remote color以标识远程TLOC。请比较DstIp和DPubIp:获知的实际目标可能与与会话关联的公共终端不同。如果仅目标过滤器不再匹配已获取的地址，请使用对等体特定视图。

Invalid next hop、SA标识符为0或In HW:FALSE提供其他证据，而不是唯一签名。这些字段也可能

会在普通设置或衬底故障期间出现。

步骤4.检验NAT网关实际执行的操作

请向NAT或服务提供商操作员询问带有以下字段的并发UDP转换记录：

Spoke	Inside source	Translated source	External destination
A	192.168.11.11:12346	203.0.113.10:12346	198.51.100.2:12346
B	192.168.12.12:12346	203.0.113.10:12346	198.51.100.3:12346
A	192.168.11.11:12346	203.0.113.10:12346	198.51.100.42:12346
B	192.168.12.12:12346	203.0.113.10:45678	198.51.100.42:12346

这是一个说明性转换表。对NAT平台使用等效的诊断命令。在Cisco IOS XE NAT网关上，从以下内容开始：

<#root>

show ip nat translations | include 203.0.113.10

保留内部源列和目标列。仅包含已转换源的列表无法显示为什么重用对不同的vBonds有效，或哪个分支拥有中心映射。

在vBonds和集线器的NAT之前和之后捕获受影响的UDP流量。保留完整的数据包长度，以便VCONTROL尾部不会被截断。使用与TAC商定的限定交换端口分析器(SPAN)或嵌入式数据包捕获(EPC)集合。比较捕获点之间的相同流量；不要从非同步捕获时钟推断延迟。

步骤5.检查Wireshark中的VCONTROL

对于此处介绍的IPv4 BFD/VCONTROL格式，请使用附录中的Lua解码器。将代码另存为sdwan_vctrl.lua，与捕获保存在同一文件夹中。此诊断解码器显示可见字段；它不会解密ESP或验证身份验证。其作用域是完整的、具有预期的BFD/VCONTROL标志和尾部结构的非嵌套IPv4/UDP数据包。

从包含解码器和捕获的文件夹运行Wireshark命令：

<#root>

**wireshark -n **

```
  -X lua_script:sdwan_vctrl.lua \

  -d udp.port==12346,sdwan_vctrl \

  -d udp.port==45678,sdwan_vctrl \

  -r hub.pcap
```

将示例端口替换为捕获中的端口。在Windows或macOS上，如果Wireshark可执行文件不在PATH中，请使用它的完整路径。

或者，在Wireshark启动时加载Lua脚本，选择相关的UDP数据包，然后导航到分析>解码为。将适用的源和目标UDP端口条目分配到SDWAN_VCTRL / SD-WAN VCTRL(IPv4)。 另一个端口上的显式UDPENCAP条目可以优先于此解码器。将更改应用于此分析会话。

应用此显示筛选器：

```
sdwan_vctrl && ip.addr == 203.0.113.10
```

展开SD-WAN VCTRL > Visible VCONTROL trailer并比较：

```
Outer IPv4 source:       203.0.113.10
Outer UDP source port:   45678
Carried source IPv4:     203.0.113.10
Carried source UDP port: 12346
Carried destination:     198.51.100.42:12346
```

这些示例值描述数据平面学习完成之前初始隧道建立。它们显示B的实际中心映射及其承载端点。NAT修改外部报头；它不会重写此帧尾中的终端信息。仅此不匹配并不表明存在故障。

运行TShark命令进行紧凑导出：

<#root>

```
tshark -n -X lua_script:sdwan_vctrl.lua \

  -d udp.port==12346,sdwan_vctrl \
```

```
  -d udp.port==45678,sdwan_vctrl \

  -r hub.pcap -Y sdwan_vctrl -T fields \

  -e frame.number -e ip.src -e udp.srcport -e ip.dst -e udp.dstport \

  -e sdwan_vctrl.spi \

  -e sdwan_vctrl.src_ip -e sdwan_vctrl.src_port \

  -e sdwan_vctrl.dst_ip -e sdwan_vctrl.dst_port
```

如果没有自定义解码器，此处描述的IPv4格式在UDP负载的末尾具有此17字节类型长度值(TLV)报尾。链路层填充在UDP负载之外：

```
81 00 0f cb 00 71 0a 30 3a c6 33 64 2a 30 3a 00 11
|  |     |           |     |           |     |
|  |     |           |     |           |     +-- Complete trailer length: 17
|  |     |           |     |           +-------- Destination port: 12346
|  |     |           |     +-------------------- Destination: 198.51.100.42
|  |     |           +-------------------------- Source port: 12346
|  |     +-------------------------------------- Source: 203.0.113.10
|  +-------------------------------------------- TLV length: 15
+----------------------------------------------- Type: 0x81
```

请勿将此固定布局应用于IPv6、其他控制格式、截断的数据包或任意加密的数据。空的解码器过滤器可能意味着数据包超出其支持的格式。可读的源终端也不是经过身份验证的系统IP标签。

注意：Wireshark中的通用ESP序列警告不是路由器反重播诊断。在得出身份验证或所有权结论之前，请将数据包与路由器计数器和日志关联起来。

步骤6.确认受影响路径并保留证据

在集线器上，围绕观察到的故障取两个QuantumFlow处理器(QFP)计数器样本：

<#root>

```
show clock detail
```

```
show platform hardware qfp active statistics drop | include Ipsec

show sdwan bfd sessions | include 10.0.0.32

show sdwan bfd sessions | include 10.0.0.33
```

使用计数器差异而不是绝对总数。聚合IPsec丢弃可能有其他原因。确认数据包到达受影响的传输，并将其与对等设备、终端、BFD鉴别器、隧道、SA和事件时间关联。

为TAC提供两个OMP快照、分支本地属性和控件、两个对等体的中心会话详细信息、BFD状态、转换记录、限定捕获和软件版本。如果可用，请包含相关管理技术和连接/撤销历史记录。避免在大型集线器上重复收集整个表。

单独测试服务可达性，并记录实际选择的传输。零秒BFD启动指示不是持续恢复。

解决方法和限制

通过Cisco TAC和NAT操作员评估这些缓解选项。限制涉及到达公共集线器的重复通告端点；这并不意味着通常不支持对称NAT、共享公有地址或重复私有编址。恢复程序和降低效率取决于部署。

使用具有唯一分配的稳定独立于终端的映射

图示序列的一个NAT端防御选项是为每个不同的分支传输端点提供跨vBond、控制器和集线器目标的稳定、唯一的转换UDP端点。请求辐射点之间的终端独立性和唯一性。在NAT网关故障切换或状态重新创建之后，验证转换记录和捕获的结果。

例如，A可以始终使用203.0.113.10:4001，而B始终使用203.0.113.10:4002。静态转换或合适的分配策略可以提供此行为，具体取决于NAT平台和访问设计。

独立于终端的映射不需要不受限制的入站访问。保留所需的防火墙控制。NAT策略更正不会自动删除以前通告或编程的会话状态；与TAC协调新的发现周期和恢复。

禁用不同辐射点之间的端口重叠

在受支持的情况下，在相关SD-WAN分支中唯一保留转换后的UDP终端。不要仅仅因为目的地不同而重复使用它们。这将直接删除前面显示的同时重复vBond映射，即使NAT网关保留一些与目标相关的行为。

该设置可能会占用更多端口或需要额外的公有地址。它本身不会使每个分支的映射独立于端点，并且它不会在后续端点重新分配后解析过时的接收器状态。

对于多个NAT阶段或冗余NAT网关，检查最终互联网边界处可见的分配。一个NAT实例内的唯一性不会在共享同一公有地址的其他实例间建立唯一性。

对小型受控组使用端口偏移

port-offset命令提供值0到19。例如，分支A可以保留偏移0，分支B可以使用偏移1。在分支B上运行以下配置命令：

<#root>

```
Spoke-B# config-transaction

Spoke-B(config)# system

Spoke-B(config-system)# port-offset 1

Spoke-B(config-system)# commit

Spoke-B(config-system)# end

Spoke-B# show sdwan control local-properties
```

更改后检验活动源端口和转换端口。更改此系统设置可以重建跨传输的控制和数据连接。这不是一个无中断的诊断步骤。

这20个偏移量提供20个不同的基本端口组，偏移量0是唯一的选择。它们不限于一个公共IP后面的20个SD-WAN路由器。此特定解决方法仅提供20个唯一的偏移选项，需要NAT网关保留这些选项并区分它们。当偏移重复使用时，它作为唯一性方案失败。端口跳跃不会添加更多偏移标识。

例如，124台路由器不能都在一个共享地址后接收唯一偏移。每个公有地址组最多包含20台路由器的仅偏移方案至少需要七个组。此算法不是容量建议或替代NAT行为验证。重写或重叠转换端口的NAT网关即使使用更少的路由器，也可能使方案失败。

禁用源端口保留或启用随机分配

没有保留偏好的策略可以减少许多分支接受:12346的倾向于不同的vBonds。随机分配可以在更大的端口池上分布这些映射。

但是，单凭更改并不能保证避免冲突。随机分配器仍然可以跨目标重复使用端口，更改到集线器的映射，或在旧集线器状态保留期间重新分配过期的终端。验证在用户和目标之间分配是否唯一。仅随机端口不能建立唯一性。

检查NAT网关及其运营商的源端口保留和分配设置。适用的配置取决于NAT平台；使用转换记录和数据包捕获验证到vBonds和集线器的结果映射。

转换后的端口可能不在原始SD-WAN base-port范围之内。检查防火墙规则，使其不会仅仅因为源端口更改而拒绝合法转换的流量。

单独的公有地址或保留端口块

专用转换地址，或者每个分支保留的非重叠端口块，可以为不同的活动分支提供不同的通告终端。策略必须在发现和集线器路径以及冗余NAT节点间一致地应用。可能需要额外的地址空间和提供商支持。映射停用后的动态重用仍需要正确的中心会话生命周期管理。

保留备用传输并协调恢复

经过验证的MPLS或其他符合条件的路径可以保留服务，同时调查受影响的互联网路径。阻止共享NAT地址的基础访问控制列表(ACL)仅提供包含功能：它可为该地址后面的每个分支删除Internet冗余，并且不会自动撤消其OMP通告。

控制清除、接口重置、NAT状态清除或集线器重新加载可以临时重新创建状态。没有保证防止再次发生。不要执行宽清除作为诊断的第一步。保留证据，尽可能更正分配条件，并与TAC就最小恢复操作达成一致。验证恢复路径上的持续BFD Up和双向流量。

相关软件缺陷

Cisco Bug ID	与本文相关的范围
CSCww45867	不同对等体通告同一公共TLOC端点时的中心会话所有权限制的文档。此记录跟踪文档更新，而不是软件更正。
CSCwx12986	自动端口跳跃/NAT学习序列可以在公共TLOC字段中保留私有元组。如果分支机构重复使用相同的专用IP：端口，则不正确的公共终端可以在对等体之间匹配。更正此生成器缺陷不会阻止有效的转换公共端口重复使用。

Cisco Bug ID	与本文相关的范围
CSCwx48929	当发现返回到同一vBond时，NAT分类可能错误地变为E。正确的分类不会阻止有效的不同vBond端口重用场景。

对Cisco Bug ID CSCwx12986和Cisco Bug ID CSCwx48929的更正不能保证唯一通告的公共端点，也不能自行消除此冲突场景。关闭Cisco Bug ID CSCww45867文档记录不会改变软件行为。通过Cisco TAC确认适用的限制和缓解。

附录：解码IPv4 VCONTROL报尾

将此代码另存为sdwan_vctrl.lua。诊断解码器不对数据包进行身份验证或解密其负载。它使用sdwan_vctrl显示过滤器并注册一个显式Decode As选项。数据包捕获保持不变。

```
-- Cisco SD-WAN IPv4 VCONTROL trailer decoder.
-- Explicit UDP "Decode As" only; does not decrypt ESP.
-- Scope: complete ordinary IPv4/UDP packets, one IP/UDP layer.

local p = Proto("sdwan_vctrl", "SD-WAN VCTRL (IPv4)")

local f = {
    spi       = ProtoField.uint32("sdwan_vctrl.spi",
                    "Wire SPI including control flags", base.HEX),
    sequence  = ProtoField.uint32("sdwan_vctrl.sequence",
                    "ESP sequence number", base.DEC),
    type      = ProtoField.uint8("sdwan_vctrl.type",
                    "Trailer type", base.HEX),
    tlv_len   = ProtoField.uint16("sdwan_vctrl.tlv_len",
                    "TLV length", base.DEC),
    src_ip    = ProtoField.ipv4("sdwan_vctrl.src_ip",
                    "Carried source IPv4"),
    src_port  = ProtoField.uint16("sdwan_vctrl.src_port",
                    "Carried source UDP port", base.DEC),
    dst_ip    = ProtoField.ipv4("sdwan_vctrl.dst_ip",
                    "Carried destination IPv4"),
    dst_port  = ProtoField.uint16("sdwan_vctrl.dst_port",
                    "Carried destination UDP port", base.DEC),
    total_len = ProtoField.uint16("sdwan_vctrl.total_len",
                    "Complete trailer length", base.DEC)
}
p.fields = f

-- Field extractors must be created outside the dissector.
local get_udp_length = Field.new("udp.length")
local get_ip_version = Field.new("ip.version")
local get_ip_proto   = Field.new("ip.proto")

function p.dissector(tvb, pinfo, tree)
    local lengths = { get_udp_length() }
    local versions = { get_ip_version() }
    local protocols = { get_ip_proto() }

    -- Deliberately skip nested/tunnelled protocol stacks.
```

```
    if #lengths ~= 1 or #versions ~= 1 or #protocols ~= 1 then
        return 0
    end
    if tonumber(versions[1].value) ~= 4 or
       tonumber(protocols[1].value) ~= 17 then
        return 0
    end

    -- UDP length includes its 8-byte header.
    local n = tonumber(lengths[1].value) - 8

    -- Require the full declared payload, not a truncated fragment.
    if n < 25 or tvb:reported_len() ~= n or tvb:len() ~= n then
        return 0
    end

    -- Exact BFD + VCTRL control-bit combination.
    local spi = tvb(0, 4):uint()
    if math.floor(spi / 0x10000000) ~= 0xA then
        return 0
    end

    -- Trailer is at the end of UDP payload, not Ethernet frame.
    local o = n - 17

    if tvb(o, 1):uint() ~= 0x81 or
       tvb(o + 1, 2):uint() ~= 15 or
       tvb(o + 15, 2):uint() ~= 17 then
        return 0
    end

    local root = tree:add(p, tvb())
    root:add(f.spi,      tvb(0, 4))
    root:add(f.sequence, tvb(4, 4))

    local trailer = root:add(tvb(o, 17),
                             "Visible VCONTROL trailer")
    trailer:add(f.type,      tvb(o,      1))
    trailer:add(f.tlv_len,   tvb(o + 1,  2))
    trailer:add(f.src_ip,    tvb(o + 3,  4))
    trailer:add(f.src_port,  tvb(o + 7,  2))
    trailer:add(f.dst_ip,    tvb(o + 9,  4))
    trailer:add(f.dst_port,  tvb(o + 13, 2))
    trailer:add(f.total_len, tvb(o + 15, 2))

    pinfo.cols.protocol = "SDWAN-VCTRL"
    return n
end

DissectorTable.get("udp.port"):add_for_decode_as(p)
```

相关信息

- RFC 7857，第3部分：端口重叠行为
- RFC 4787：单播UDP的NAT行为要求
- RFC 5737：为文档保留的IPv4地址块

- Cisco Catalyst SD-WAN设计指南
- Cisco IOS XE Catalyst SD-WAN系统命令：port-offset
- Cisco SD-WAN端口偏移配置
- Wireshark命令行参考
- TShark命令行参考
- 技术支持和文档 - Cisco Systems

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）。