Solucionar problemas de túneis de hub SD-WAN com pontos finais NAT duplicados

Contents

- Introdução
- Pré-requisitos
  - Requisitos
  - Componentes Utilizados
- Informações de Apoio
  - Aprendizado de NAT de plano de controle
  - NAT Learning e VCONTROL de plano de dados
  - Preservação de portas, sobreposição de portas e aleatorização de portas
- Topologia de exemplo
- Por que a colisão acontece
  - vBonds diferentes podem observar o mesmo endpoint público
  - Os Mesmos Anúncios Atingem Um Hub
  - Colisões privadas como públicas com endereços de filial repetidos
  - Reatribuição de ponto final de NAT
- Sintomas e impacto
- Faça o diagnóstico do problema
  - Etapa 1. Registrar o Software e o Estado de Transporte Spoke
  - Etapa 2. Localizar endpoints duplicados no OMP
  - Etapa 3. Comparar TLOCs Lógicos com Propriedade de Sessão de Hub
  - Etapa 4. Verificar o que o Gateway NAT realmente faz
  - Etapa 5. Inspecionar o VCONTROL no Wireshark
  - Etapa 6. Confirmar o caminho afetado e preservar evidências
- Soluções alternativas e limitações
  - Use mapeamentos estáveis independentes de endpoint com alocações exclusivas
  - Desative a sobreposição de portas em spokes diferentes
  - Use o deslocamento de porta para um grupo pequeno e controlado
  - Desabilitar Preservação de Porta de Origem ou Habilitar Alocação Aleatória
  - Endereços Públicos Separados ou Blocos de Portas de Reserva
  - Manter um transporte alternativo e coordenar a recuperação
- Defeitos de software relacionados
- Anexo: Decodificar o Trailer IPv4 VCONTROL
- Informações Relacionadas

Introdução

Este documento descreve como identificar e mitigar falhas de túnel de hub quando os spokes compartilham um ponto final de Network Address Translation (NAT) público.

Pré-requisitos

Requisitos

A Cisco recomenda que você tenha conhecimento destes tópicos:

- Arquitetura de rede de longa distância (SD-WAN) definida pelo software Cisco Catalyst e conexões de controle
- Protocolo de Gerenciamento de Sobreposição (OMP - Overlay Management Protocol), Localizadores de Transporte (TLOCs - Transport Locators), Detecção de Encaminhamento Bidirecional (BFD - Bidirectional Forwarding Detection) e Segurança IP (IPsec)
- Redes de transporte Multiprotocol Label Switching (MPLS)
- Mapeamento de NAT e comportamento de filtro
- Análise básica de pacotes com o Wireshark

Componentes Utilizados

As informações neste documento são baseadas nestas versões de software e hardware:

- Roteadores Cisco IOS® XE Catalyst SD-WAN
- Controlador Cisco Catalyst SD-WAN (vSmart)
- Validador Cisco Catalyst SD-WAN (vBond)

A falha de propriedade da sessão de hub foi reproduzida no Cisco IOS XE 17.18.2 com chaves em pares operacionais. Os campos de disponibilidade e saída do comando variam de acordo com a plataforma e a versão.

Os exemplos usam endereços privados ou endereços de documentação reservados pelo RFC 5737, com nomes de dispositivos genéricos. Trechos do CLI são encurtados e usam identificadores normalizados. Os nomes das interfaces e alguns campos de saída variam por plataforma e versão.

As informações neste documento foram criadas a partir de dispositivos em um ambiente de laboratório específico. Todos os dispositivos utilizados neste documento foram iniciados com uma configuração (padrão) inicial. Se a rede estiver ativa, certifique-se de que você entenda o impacto potencial de qualquer comando.

⚠ Caution: Alterações de configuração e redefinições de conexão podem interromper o tráfego. Usar um plano de manutenção aprovado. Os comandos de diagnóstico iniciais são somente leitura.

Informações de Apoio

Comece com um cenário em que o NAT simétrico seja identificado corretamente. Spoke A registra com vBond 1 e Spoke B registra com vBond 2. O gateway NAT pode atribuir o mesmo endereço IP de origem traduzido e a mesma porta a essas conexões quando seus destinos diferem, como o RFC 7857 permite. Ambos os spokes podem, consequentemente, anunciar o mesmo endpoint público para seus TLOCs de internet pública, mesmo quando ambos relatam NAT tipo A.

A topologia de exemplo usa a cor da Internet pública. O mesmo cenário pode envolver outras cores públicas, como biz-internet. Aplicar os filtros de diagnóstico à cor de transporte afetada real; a cor literal da internet pública não é uma condição necessária.

✎ Observação: o Cisco Catalyst SD-WAN não suporta TLOCs de sites diferentes que anunciam o mesmo endereço IP público e a mesma porta UDP em direção a uma borda de recebimento comum. O comportamento do receptor nessa condição de colisão é imprevisível. Isso pode ocorrer mesmo quando o NAT simétrico é identificado corretamente. A classificação correta do NAT, por si só, não garante que os pontos de extremidade anunciados sejam exclusivos.

Aprendizado de NAT de plano de controle

Um Edge envia seu endereço de transporte privado e porta durante o registro com vBond. O vBond observa o endereço de origem e a porta nessa conexão e retorna o ponto final observado para o Edge. O Edge pode anunciar essas informações através do OMP como o endereço público e os atributos de porta de um TLOC, incluindo um TLOC com uma cor pública, como internet pública ou internet de negócios.

A observação descreve a conexão com aquele vBond. Com o NAT dependente de destino, o mapeamento em direção a um hub pode ser diferente. O roteador compara as observações de NAT para classificar o comportamento de mapeamento. Em show sdwan control local-properties, A significa mapeamento dependente de endereço e porta, e E significa mapeamento independente de ponto final. A classificação exibida descreve a decisão do roteador; uma tabela de conversão NAT ou captura de pacotes verifica o comportamento real do gateway NAT.

Por exemplo, execute show sdwan control local-properties e localize a coluna NAT TYPE para o

transporte. Essa saída reduzida retém a legenda do tipo NAT e as colunas de interface escolhidas, com endereços normalizados:

<#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
```

O transporte biz-internet informa A e anuncia 203.0.113.10:23012 como seu endpoint público. O transporte MPLS relata E e tem endpoints públicos e privados iguais. Leia a classificação do TIPO NAT; a igualdade de endereço ou porta sozinha não identifica o tipo de NAT. N significa que o roteador não aprendeu uma classificação.

NAT Learning e VCONTROL de plano de dados

Antes que um hub possa processar um pacote BFD criptografado, ele deve encontrar a associação de segurança (SA) IPsec de entrada para o par e o túnel correspondentes. A SA identifica a chave criptográfica e outro estado necessário para autenticar e descriptografar o pacote. O receptor usa informações de endpoint junto com o Índice de Parâmetros de Segurança (SPI - Security Payload) ESP para localizar esse estado. O SPI sozinho não é um identificador de spoke globalmente exclusivo.

Um spoke atrás de NAT simétrico normalmente inicia o tráfego em direção a um hub alcançável. Isso cria um mapeamento para esse destino específico. O endereço de origem ou a porta vista pelo hub pode diferir do endpoint público previamente aprendido por vBond e anunciado no OMP. Como resultado, os cabeçalhos externos recebidos nem sempre correspondem ao ponto final no estado de túnel instalado. O hub precisa de uma maneira de associar esse novo endpoint on-wire com o TLOC remoto conhecido e seu estado IPsec antes de poder descriptografar o BFD.

Para o formato IPv4 BFD-over-IPsec discutido aqui, a SD-WAN coloca as informações de endpoint em um trailer VCONTROL visível no final do payload de UDP, fora do payload de BFD criptografado. O hub pode ler essas informações antes da descriptografia. O VCONTROL transporta informações de endpoint, não chaves de criptografia, e não estabelece a identidade do peer sozinho.

Por exemplo, durante o estabelecimento inicial do túnel, antes que o aprendizado do plano de dados seja concluído:

```
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
```

O receptor primeiro tenta sua consulta normal de ponto final/SPI. Se essa pesquisa não encontrar uma SA, o ponto final de origem VCONTROL fornecerá um fallback. Neste exemplo, o hub pode usar o 203.0.113.10:12346 transportado, juntamente com seu contexto de destino local e o SPI, para localizar a SA de entrada associada ao TLOC remoto conhecido, mesmo que o pacote chegue de 203.0.113.10:45678.

A SA escolhida fornece o contexto da chave de recebimento. O SPI escolhe a geração de chave aplicável nesse contexto. Em seguida, o hub executa a autenticação e a descriptografia do IPsec; somente um pacote que passa pelas verificações necessárias está disponível para processamento BFD. Um trailer legível ou apenas uma pesquisa bem-sucedida não comprovam autenticação, descriptografia ou recepção BFD bem-sucedida.

A diferença entre as portas de origem observadas e transportadas pode, portanto, ser normal com o NAT simétrico. Para que a relação de peer funcione, o hub deve associar o pacote ao túnel e ao estado de chave corretos e aprender o ponto final de retorno real. Neste exemplo, o aprendizado bem-sucedido permite que o hub envie o tráfego de retorno para 203.0.113.10:45678. Isso é separado do aprendizado do plano de controle através do vBond.

O VCONTROL não força o receptor a substituir uma correspondência de ponto final/SPI existente. Se o estado de ponto final duplicado associar um pacote com a geração incorreta de peer ou chave, o trailer não garantirá a recuperação. Ele fornece uma dica de consulta; não é possível distinguir dois peers que transportam o mesmo ponto final sem o túnel circundante correto e o estado criptográfico.

O VCONTROL também está presente nos pacotes BFD aplicáveis quando o NAT tipo E é relatado. Sua presença não é um flag NAT simétrico e não está incluído em todos os pacotes de dados IPsec comuns. Para o formato BFD IPv4 descrito aqui, as informações de TLOC público são omitidas entre duas cores privadas.

O ponto final específico do peer aprendido através do plano de dados não significa que o spoke deve substituir seu anúncio OMP por um novo ponto final para cada destino. Diferentes mapeamentos aprendidos são uma parte esperada do percurso simétrico-NAT.

Uma vez que um túnel está Ativo, um remetente que aprendeu seu próprio ponto final traduzido

para esse túnel pode transportar esse ponto final como sua informação de origem. Este é o ponto final público por túnel do remetente, não o ponto final remoto ou um novo anúncio global OMP. Sempre compare o pacote, o anúncio OMP e o estado da sessão do mesmo período de observação.

Preservação de portas, sobreposição de portas e aleatorização de portas

Estes são comportamentos diferentes:

Comportamento	Significado
Preservação da porta de origem	O gateway NAT tenta reter a porta de origem UDP original do spoke, por exemplo, 12346, como a porta de origem convertida.
Sobreposição de portas entre destinos	O gateway NAT reutiliza um IP convertido: para diferentes endpoints internos que se comunicam com diferentes destinos externos.
Aleatoriedade de portas	O gateway NAT escolhe as portas convertidas usando um método de alocação aleatória.
Mapeamento independente de endpoint	Um ponto final interno retém o mesmo ponto final convertido em destinos externos enquanto esse mapeamento existir.

O RFC 7857, Seção 3 permite a sobreposição de portas externas para diferentes endpoints internos quando os destinos são diferentes. O NAT distingue os fluxos usando suas cinco tuplas. Isso não permite dois fluxos UDP simultâneos indistinguíveis para o mesmo destino. A atribuição e a preservação aleatórias são opções de atribuição separadas; nem estabelece a independência do endpoint.

Topologia de exemplo

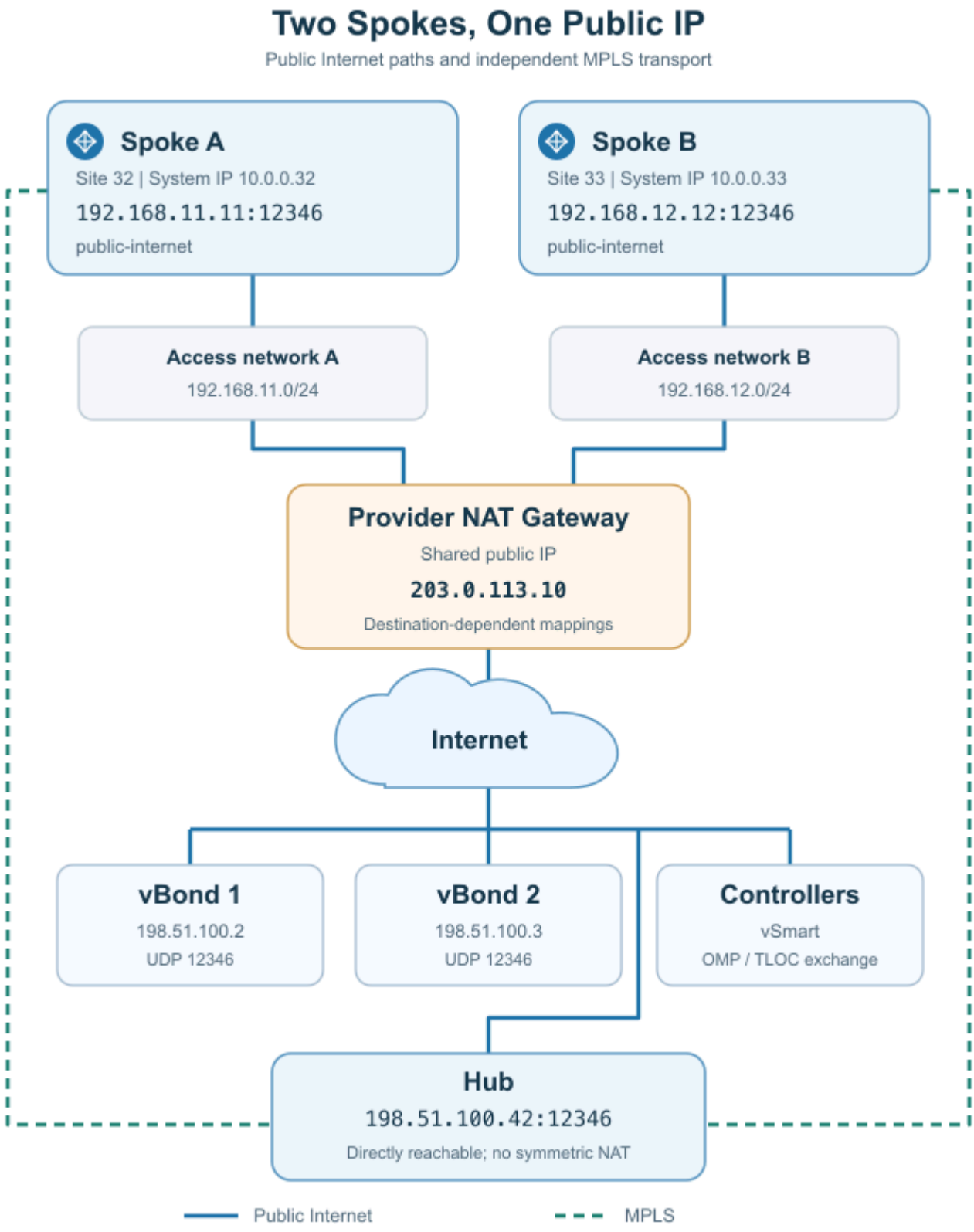

Figura 1. Gateway NAT compartilhado, redes de acesso spoke separadas e transporte MPLS independente.

O spoke A usa 192.168.11.11:12346, e o spoke B usa 192.168.12.12:12346. Ambos compartilham o endereço público 203.0.113.10 através do gateway NAT. O hub tem um endereço diretamente acessível e não está atrás de NAT simétrico. Os dois spokes podem anunciar o

mesmo IP:porta público, mesmo que seus endpoints privados sejam diferentes.

Por que a colisão acontece

vBonds diferentes podem observar o mesmo endpoint público

Considere estes mapeamentos UDP simultâneos:

Correspondente	Fonte Privada	Fonte traduzida	Destino
R	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

As respostas retornam corretamente porque o NAT também verifica o ponto final remoto. Uma resposta do vBond 1 pertence ao mapeamento de A; uma resposta de vBond 2 pertence ao mapeamento para B.

Ambos os roteadores também podem entrar em contato com o outro vBond e detectar corretamente o mapeamento dependente de destino. Por exemplo, mantenha inalterada a tupla de fonte privada para cada spoke e considere esta ordem de observação:

Correspondente	Primeira observação de vBond aceita	Comparação com o outro vBond	Resultado
R	vBond 1 retorna 203.0.113.10:12346	vBond 2 retorna 203.0.113.10:40001	NAT A; tupla pública retida 203.0.113.10:12346
B	vBond 2 retorna 203.0.113.10:12346	vBond 1 retorna 203.0.113.10:40002	NAT A; tupla pública retida 203.0.113.10:12346

As observações de vBond diferem para cada roteador, portanto, a detecção do NAT A está correta. As primeiras observações retidas, no entanto, coincidem nos dois roteadores. A classificação A não aloca uma porta pública exclusiva em todos os spokes e visitar ambos os vBonds não garante essa exclusividade.

Cada raio recebe o endereço público e o porto observado por vBond na resposta de registro e usa a observação retida para preencher suas propriedades de transporte local. O spoke inclui esse endpoint público, junto com seu endpoint privado, nos atributos TLOC que ele anuncia ao controlador SD-WAN por meio do OMP. Em seguida, o controlador distribui o TLOC a outras

arestas de acordo com a política de controle aplicável; O próprio vBond não gera o anúncio OMP.

Se ambos os raios mantiverem o mesmo ponto final público observado, o OMP pode, portanto, conter:

```
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
```

Esses TLOCs lógicos permanecem diferentes porque seus IPs de sistema são diferentes.

Os Mesmos Anúncios Atingem Um Hub

O gateway NAT deve distinguir os dois spokes quando eles entram em contato com o mesmo hub. Uma alocação possível é:

Correspondente	Origem Traduzida em Direção ao Hub	Destino
R	203.0.113.10:12346	198.51.100.42:12346
B	203.0.113.10:45678	198.51.100.42:12346

O hub tem dois pares lógicos anunciando o mesmo ponto final de destino, enquanto os pacotes podem chegar de diferentes pontos finais reais. Sob operação NAT simétrica normal, a aprendizagem do plano de dados ajuda a estabelecer os mapeamentos de retorno.

A limitação de propriedade de sessão de hub surge quando diferentes pares compartilham o estado da sessão escolhido por esta chave de transporte:

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

Se ambos os anúncios escolherem a mesma chave, uma atualização de um peer posterior pode substituir as informações de propriedade no estado compartilhado. Uma retirada, exclusão ou recriação subsequente da sessão para um peer pode perturbar o estado de IPsec/BFD para o outro peer. A aprendizagem do plano de dados não garante a recuperação quando esse estado já está associado ao peer ou à geração de fluxo errado.

O reboque visível não é um identificador de pares independente: ambos os spokes podem

transportar o mesmo ponto final de origem. Se o hub escolher um estado IPsec existente que pertença à geração incorreta de peer ou chave, a autenticação e a aprendizagem poderão falhar mesmo que os pacotes cheguem ao hub. Para estabelecer essa sequência para um pacote específico, o Cisco Technical Assistance Center (TAC) deve correlacionar o SA escolhido com seu par e geração. O trailer do pacote sozinho não pode comprová-lo.

O mesmo IP público sozinho não estabelece uma colisão. Diferentes portas remotas normalmente produzem diferentes chaves de transporte. Verifique também o endereço hub-local, a porta local e o protocolo escolhidos. Se um terceiro peer anunciar o mesmo endpoint, ele também poderá competir por essa chave; o padrão de falha depende da ordem do evento.

Colisões privadas como públicas com endereços de filial repetidos

O bug da Cisco ID CSCwx12986 pode fazer com que um Edge anuncie uma tupla privada RFC 1918 em seus atributos TLOC públicos após um salto de porta de origem automática durante a comparação de NAT. Se filiais diferentes reutilizarem o mesmo endereço privado e porta, esses atributos públicos incorretos poderão corresponder entre pares e expor o problema de propriedade da sessão de hub.

Por exemplo, ambos os spokes podem usar 192.168.11.11 atrás de firewalls de filial separados. Cada firewall mantém sua rede de filial separada e converte o endereço de origem. Se um gateway NAT do provedor upstream o converter novamente, o caminho terá dois estágios NAT, ou NAT duplo. Contextos de roteamento e encaminhamento virtual (VRF) separados são outra maneira de manter isolados os endereços de filial sobrepostos.

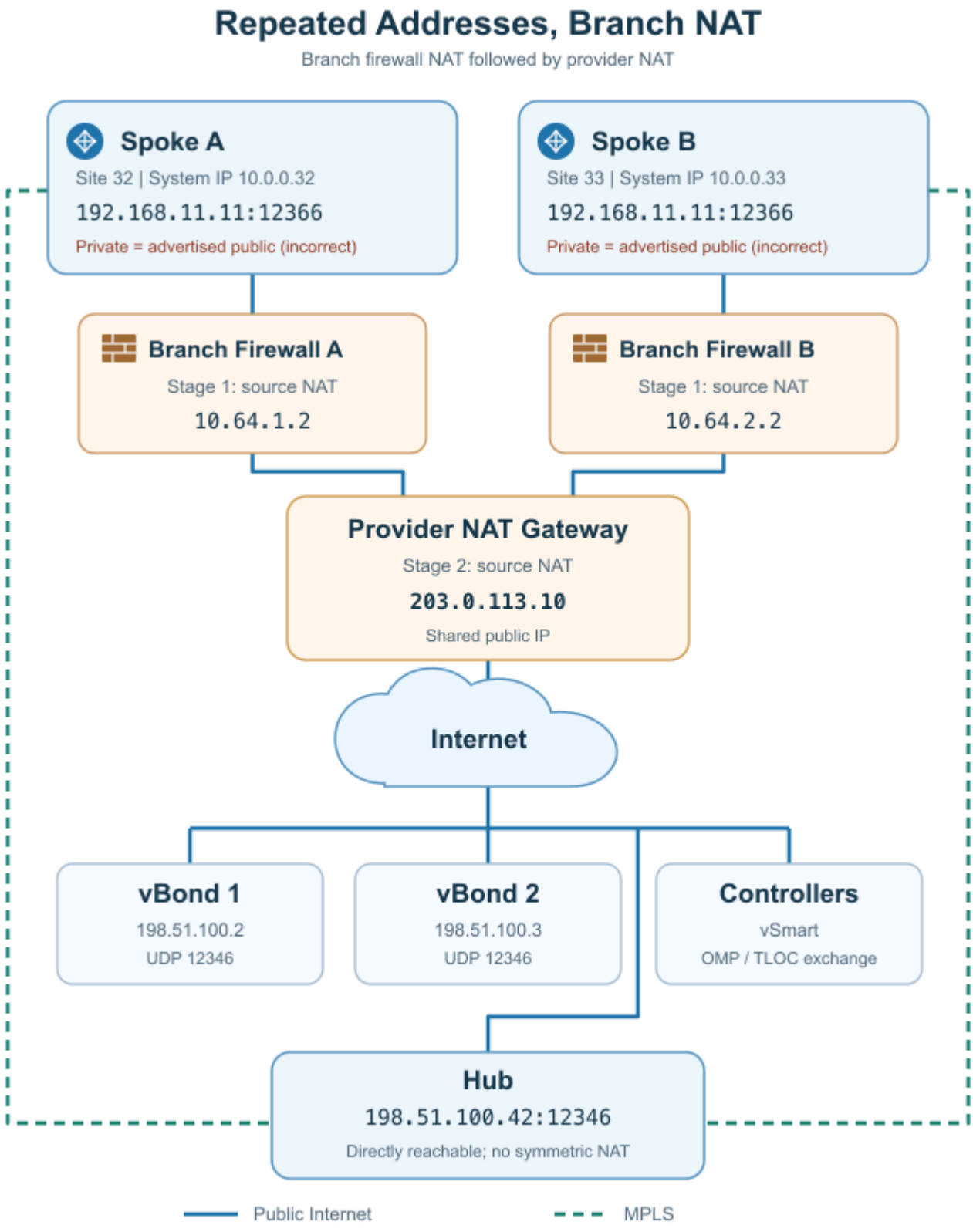

Figura 2. Firewall da filial e conversões NAT do provedor. Os atributos públicos TLOC incorretos retêm o ponto final privado original, enquanto os pacotes usam os endereços origem convertidos.

Se ambos os spokes encontrarem o defeito em seus TLOCs de internet pública e usarem a mesma porta de origem de pós-salto, o hub poderá receber estes anúncios:

Correspondente	Ponto de Extremidade Privado Após Salto de Porta	Ponto de Extremidade TLOC Público Incorreto
R	192.168.11.11:12366	192.168.11.11:12366
B	192.168.11.11:12366	192.168.11.11:12366

Note: Endereços privados repetidos e NAT duplo não são pré-requisitos para a ID de bug Cisco CSCwx12986 em si. A colisão adicional entre pares exige a correspondência dos valores anunciados de IP:porta e o mesmo endereço hub-local, porta e protocolo. O mesmo IP privado com portas anunciadas diferentes não estabelece essa colisão.

A igualdade de endereço público e privado é normal em um transporte sem NAT. Um endereço RFC 1918 em um campo público é suspeito apenas quando contradiz o caminho pretendido e a tradução observada. Os endereços privados exclusivos da filial evitam a sobreposição de tupla privada repetida neste exemplo; eles não impedem o cenário de reutilização de porta público traduzido separado descrito anteriormente.

Reatribuição de ponto final de NAT

Um ponto final também pode ser reatribuído a outro spoke depois que um mapeamento NAT expira, enquanto o hub ainda retém o estado antigo do peer. Compare o histórico de mapeamento com o ciclo de vida da sessão do hub para identificar a sobreposição entre os proprietários antigos e novos.

Sintomas e impacto

Os indicadores típicos incluem:

- As conexões de controle permanecem Ativas enquanto os caminhos BFD afetados estão Desativados ou não conseguem se recuperar.
- Dois IPs de sistema diferentes anunciam o mesmo IP:porta público.
- O hub retém os TLOCs lógicos, mas um ponto de vista filtrado da sessão da plataforma está faltando ou mostra informações inconsistentes de IP do sistema, cor remota, discriminador, túnel ou SA.
- Uma retirada de transporte ou recriação para um peer perturba outro peer com o mesmo endpoint.
- As quedas de entrada de IPsec podem aumentar no caminho afetado.

A falha pode tornar o caminho de transporte afetado inutilizável. Se esse for o único caminho de

transporte utilizável para um spoke, o spoke será isolado da sobreposição. Em um hub grande, muitos spokes que compartilham um endereço NAT aumentam a população que requer correlação, mas o número de caminhos afetados deve ser medido; não pode ser deduzido a partir do número de locais por trás desse endereço.

Faça o diagnóstico do problema

Etapa 1. Registrar o Software e o Estado de Transporte Spoke

Em cada Borda afetada do Cisco IOS XE, substitua **<affected-color>** pela cor de transporte real, por exemplo biz-internet ou public-internet:

<#root>

```
show clock detail
```

```
show version | include Cisco IOS XE Software
```

```
show sdwan control local-properties
```

```
show sdwan control connections | include <affected-color>
```

```
show sdwan omp peers
```

Registre o endereço privado/público e a porta, a interface, a cor, o tipo de NAT, a ID do local e o estado de controle. Após a primeira amostra completa das propriedades locais, use este filtro compacto para reter os cabeçalhos da interface e a linha de transporte escolhida. Substitua TenGigabitEthernet0/1/0.10 pela interface afetada:

<#root>

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

Para o cenário principal, os campos relevantes podem ser lidos em ambos os spokes. Este é um resumo de campo, não o layout de coluna nativo:

```
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
```

Uma conexão de controle Up verifica a relação do controlador. Ele não verifica um caminho BFD spoke-to-hub nem a exclusividade do endpoint público.

Etapa 2. Localizar endpoints duplicados no OMP

No hub Cisco IOS XE:

<#root>

show clock detail

show sdwan omp tlocs table | include 203.0.113.10

No controlador SD-WAN (vSmart):

<#root>

show clock

show omp tlocs | tab | include 203.0.113.10

Esta saída de Controlador usa identificadores normalizados. IPv6 e colunas de região são omitidos:

<#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
```

Compare o IP do sistema TLOC, a cor, o encapsulamento, o IP público e a porta pública. Vários caminhos para o mesmo TLOC lógico de diferentes controladores são normais. O candidato à colisão são duas identidades de peer diferentes com o mesmo endpoint efetivo.

Repita o comando filtrado após 30-60 segundos e mantenha os dois timestamps. A sobreposição persistente é uma prova útil; comandos perfeitamente simultâneos não são necessários. Para um curto período transitório, colete eventos OMP e registros de ciclo de vida de sessão correspondentes com o TAC.

Para dois pares suspeitos, preserve suas entradas individuais da tabela TLOC:

<#root>

show sdwan omp tlocs table | include 10.0.0.32

show sdwan omp tlocs table | include 10.0.0.33

Para obter uma visualização compacta dos detalhes do TLOC IPv4 no Cisco IOS XE, use este filtro. Substitua o IP do sistema pelo peer em investigação e mantenha as aspas em torno do seletor ip <system-IP>:

<#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|::

Compare os atributos de endereço/porta público e privado para ambos os peers. O filtro omite as linhas unset, com valor zero, padrão e IPv6 correspondentes nesta exibição.

Etapa 3. Comparar TLOCs Lógicos com Propriedade de Sessão de Hub

Inspecione o estado TLOC entregue ao FTM. Compare os IPs do sistema, as cores, os endereços IP públicos e as portas UDP públicas de ambos os peers.

<#root>

show platform software sdwan ttmd tlocs

Execute estes comandos para comparar o FTM (Forwarding Table Manager), o BFD e os registros de sessão de plataforma:

<#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

Este exemplo mostra o padrão com identificadores normalizados. Os rótulos das colunas são acrescentados às linhas FTM filtradas para maior clareza; eles não fazem parte da saída do comando filtrado:

<#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
```

O OMP e o FTM mantêm as identidades lógicas dos pares, enquanto esta pesquisa de sessão mostra apenas o Spoke A para o destino compartilhado. O spoke A também tem um estado de FTM inativo; a presença na tabela de sessão de plataforma não indica um caminho íntegro. O índice FTM TLOC é diferente do discriminador local BFD (BFD-LD) na tabela de sessão. Esse padrão é um motivo para investigar a propriedade entre camadas. Uma linha CLI ausente sozinha não estabelece perda de pacotes.

Nesta tabela, Color representa o transporte local; use Remote color na visualização detalhada para identificar o TLOC remoto. Compare também DstIp e DPubIp: um destino real aprendido pode diferir do endpoint público associado à sessão. Use a visualização específica do peer se um filtro somente de destino não corresponder mais ao endereço aprendido.

Um próximo salto inválido, identificador SA de 0 ou Em HW: FALSO fornece evidência adicional em vez de uma assinatura única. Esses campos também podem ocorrer durante a configuração comum ou falha subjacente.

Etapa 4. Verificar o que o Gateway NAT realmente faz

Peça ao NAT ou ao operador do provedor de serviços os registros de conversão de UDP simultâneos com estes campos:

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

Esta é uma tabela de tradução ilustrativa. Use o comando de diagnóstico equivalente para sua plataforma NAT. Em um gateway NAT do Cisco IOS XE, comece com:

<#root>

```
show ip nat translations | include 203.0.113.10
```

Mantenha as colunas de origem e destino internas. Uma lista contendo apenas a fonte traduzida não pode mostrar por que a reutilização é válida para vBonds diferentes ou qual spoke possui o mapeamento de hub.

Capture os fluxos de UDP afetados antes e depois do NAT para vBonds e o hub. Preserve o comprimento total do pacote para que o trailer VCONTROL não fique truncado. Use uma coleção limitada Switched Port Analyzer (SPAN) ou Embedded Packet Capture (EPC) acordada com o TAC. Comparar o mesmo fluxo nos pontos de captura; não inferir a latência de relógios de captura não sincronizados.

Etapa 5. Inspecionar o VCONTROL no Wireshark

Use o decodificador Lua no apêndice para o formato IPv4 BFD/VCONTROL descrito aqui. Salve o código como sdwan_vctrl.lua na mesma pasta da sua captura. Este decodificador de diagnóstico exibe campos visíveis; ele não descriptografa o ESP ou verifica a autenticação. Seu escopo é composto por pacotes IPv4/UDP completos e não aninhados com o sinalizador BFD/VCONTROL e a estrutura de trailer esperados.

Execute o comando Wireshark da pasta que contém o decodificador e sua captura:

<#root>

```
wireshark -n \

  -X lua_script:sdwan_vctrl.lua \

  -d udp.port==12346,sdwan_vctrl \

  -d udp.port==45678,sdwan_vctrl \

  -r hub.pcap
```

Substitua as portas de exemplo por aquelas na captura. No Windows ou no macOS, use o caminho completo para o executável do Wireshark se ele não estiver em PATH.

Como alternativa, carregue o script Lua na inicialização do Wireshark, escolha um pacote UDP relevante e navegue para Analisar > Decodificar como. Atribua as entradas de porta UDP origem e destino aplicáveis a SDWAN_VCTRL / SD-WAN VCTRL (IPv4). Uma entrada UDPENCAP explícita na outra porta pode ter precedência sobre esse decodificador. Aplicar a alteração a esta

sessão de análise.

Aplicar este filtro de exibição:

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

Expanda SD-WAN VCTRL > trailer de VCONTROL visível e compare:

```
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
```

Esses valores ilustrativos descrevem o estabelecimento inicial do túnel antes da conclusão do aprendizado do plano de dados. Eles mostram o mapeamento de hub real para B ao lado de seu ponto final transportado. O NAT modifica os cabeçalhos externos; ele não regrava as informações de endpoint dentro desse trailer. Essa incompatibilidade por si só não demonstra uma falha.

Execute o comando TShark para uma exportação compacta:

<#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
```

Sem um decodificador personalizado, o formato IPv4 descrito aqui tem esse trailer TLV (type-length-value, valor de comprimento do tipo) de 17 bytes no final do payload UDP. O preenchimento da camada de enlace está fora do payload UDP:

```
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
```

Não aplique esse layout fixo ao IPv6, a outros formatos de controle, a pacotes truncados ou a dados criptografados arbitrários. Um filtro de decodificador vazio pode significar que o pacote está fora de seu formato suportado. Um ponto final de origem legível também não é um rótulo de IP de sistema autenticado.

Note: Avisos genéricos de sequência ESP no Wireshark não são diagnósticos de antirreprodução do roteador. Correlacione o pacote com os registros e contadores do roteador antes de tirar uma conclusão de autenticação ou propriedade.

Etapa 6. Confirmar o caminho afetado e preservar evidências

No hub, pegue duas amostras do contador do processador QuantumFlow (QFP) em torno da falha observada:

<#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

Use diferenças de contador em vez de totais absolutos. As quedas de IPsec agregadas podem ter outras causas. Confirme a chegada do pacote no transporte afetado e correlacione-o com o peer, o endpoint, o discriminador de BFD, o túnel, a AS e a hora do evento.

Forneça ao TAC os dois instantâneos OMP, propriedades e controles locais do spoke, detalhes da sessão do hub para os dois pontos, estado BFD, registros de conversão, capturas delimitadas e versões de software. Se já estiver disponível, inclua os técnicos de administração relevantes e o histórico de conexão/retirada. Evite a coleta repetida de tabelas inteiras em um hub grande.

Teste a acessibilidade do serviço separadamente e registre o transporte escolhido. Uma indicação de BFD Up de zero segundo não é recuperação sustentada.

Soluções alternativas e limitações

Avalie essas opções de mitigação com o Cisco TAC e o operador NAT. A limitação diz respeito a pontos terminais anunciados em duplicado que chegam a um hub comum; isso não significa que o NAT simétrico, os endereços públicos compartilhados ou o endereçamento privado repetido geralmente não sejam suportados. Os procedimentos de recuperação e a eficácia da mitigação dependem da implantação.

Use mapeamentos estáveis independentes de endpoint com alocações exclusivas

Uma opção de prevenção do lado NAT para a sequência ilustrada é dar a cada endpoint de transporte spoke distinto um endpoint UDP traduzido único e estável em vBond, controlador e destinos de hub. Solicite independência e exclusividade de endpoint nos spokes. Verifique o resultado em registros e capturas de conversão, também após failover de gateway NAT ou recriação de estado.

Por exemplo, A pode usar consistentemente 203.0.113.10:4001, enquanto B usa consistentemente 203.0.113.10:4002. Conversões estáticas ou uma política de alocação adequada podem fornecer esse comportamento, sujeito à plataforma NAT e ao design de acesso.

O mapeamento independente de ponto de extremidade não requer acesso de entrada irrestrito. Preserve os controles de firewall necessários. Uma correção de política de NAT não remove automaticamente o estado de sessão anunciado ou programado anteriormente; coordene um novo ciclo de descoberta e recuperação com o TAC.

Desative a sobreposição de portas em spokes diferentes

Onde houver suporte, reserve endpoints UDP convertidos exclusivamente nos spokes SD-WAN relevantes. Não os reutilize apenas porque seus destinos são diferentes. Isso remove diretamente o mapeamento de vBond duplicado simultâneo mostrado anteriormente, mesmo que o gateway NAT mantenha algum comportamento dependente do destino.

A configuração pode consumir mais portas ou exigir endereços públicos adicionais. Ele por si só não faz o mapeamento para cada ponto final de spoke independente e não resolve o estado do receptor obsoleto após a reatribuição posterior do ponto final.

Com vários estágios de NAT ou gateways de NAT redundantes, verifique a alocação visível no limite final da Internet. A exclusividade dentro de uma instância de NAT não estabelece exclusividade em outras instâncias que compartilham o mesmo endereço público.

Use o deslocamento de porta para um grupo pequeno e controlado

O comando port-offset fornece valores de 0 a 19. Por exemplo, Spoke A pode reter offset 0 e Spoke B pode usar offset 1. Execute estes comandos de configuração em Spoke B:

<#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
```

Verifique a origem ativa e as portas convertidas após a alteração. Uma alteração nessa configuração do sistema pode recriar o controle e as conexões de dados nos transportes. Não é uma etapa de diagnóstico sem interrupções.

Os 20 deslocamentos fornecem 20 grupos distintos de porta base com deslocamento 0 como uma escolha. Eles não são um limite geral de 20 roteadores SD-WAN atrás de um IP público.

Essa solução específica fornece apenas 20 opções de deslocamento exclusivas e exige que o gateway NAT as preserve e as diferencie. Ele falha como um esquema exclusivo quando as compensações são reutilizadas. O salto de porta não adiciona mais identidades de deslocamento.

Por exemplo, 124 roteadores não podem receber deslocamentos exclusivos atrás de um endereço compartilhado. Um esquema somente de compensação com no máximo 20 roteadores por grupo de endereço público requer pelo menos sete grupos. Essa aritmética não é uma recomendação de capacidade ou um substituto para a verificação do comportamento do NAT. Um gateway NAT que regrava ou sobrepõe portas convertidas pode anular o esquema mesmo com menos roteadores.

Desabilitar Preservação de Porta de Origem ou Habilitar Alocação Aleatória

Uma política sem uma preferência de preservação pode reduzir a tendência de muitos spokes receberem :12346 em direção a vBonds diferentes. A alocação aleatória pode distribuir esses mapeamentos em um pool de portas maior.

No entanto, nenhuma alteração isolada garante a prevenção de colisões. Um alocador aleatório ainda pode reutilizar uma porta entre destinos, alterar um mapeamento em direção ao hub ou reatribuir um ponto final expirado enquanto o estado antigo do hub permanece. Verifique se a alocação é exclusiva entre assinantes e destinos. Apenas portas aleatórias não estabelecem exclusividade.

Revise as configurações de preservação e alocação de porta de origem no gateway NAT com seu operador. A configuração aplicável depende da plataforma NAT; verifique os mapeamentos resultantes em direção a vBonds e ao hub com registros de conversão e capturas de pacotes.

As portas convertidas podem ficar fora do intervalo de portas base da SD-WAN original. Reveja as regras de firewall para que elas não rejeitem o tráfego traduzido legítimo apenas porque a porta de origem foi alterada.

Endereços Públicos Separados ou Blocos de Portas de Reserva

Endereços dedicados traduzidos, ou blocos de porta não sobrepostos reservados por spoke, podem fornecer endpoints anunciados distintos para diferentes spokes ativos. A política deve ser aplicada de forma consistente em todos os caminhos de descoberta e hub e em todos os nós NAT redundantes. Espaço de endereço adicional e suporte do provedor podem ser necessários. A reutilização dinâmica após a desativação do mapeamento ainda requer o gerenciamento correto do ciclo de vida da sessão do hub.

Manter um transporte alternativo e coordenar a recuperação

Um MPLS verificado ou outro caminho qualificado pode preservar o serviço enquanto o caminho da Internet afetado é investigado. Uma lista de controle de acesso (ACL) subjacente que bloqueia o endereço NAT compartilhado fornece somente contenção: ele pode remover a redundância da Internet para cada spoke atrás desse endereço e não retira seus anúncios OMP automaticamente.

Uma limpeza de controle, uma reinicialização de interface, uma limpeza de estado de NAT ou uma recarga de hub podem recriar temporariamente o estado. Nenhuma garante a prevenção de recorrência. Não execute limpezas amplas como a primeira etapa de diagnóstico. Preservar as provas, corrigir as condições de atribuição, sempre que possível, e chegar a acordo sobre a menor ação de recuperação com o TAC. Valide o BFD Up sustentado e o tráfego bidirecional no caminho recuperado.

Defeitos de software relacionados

ID de bug da Cisco	Âmbito de aplicação do presente artigo
CSCww45867	Documentação da limitação da propriedade da sessão do hub quando pares diferentes anunciam o mesmo ponto final TLOC público. Esse registro rastreia a atualização da documentação, em vez de uma correção de software.
CSCwx12986	Uma sequência automática de port-hop/NAT-learning pode deixar uma tupla privada nos campos TLOC públicos. Se as ramificações reutilizarem o mesmo IP:porta privado, os pontos finais públicos incorretos poderão corresponder entre pares. A correção desse defeito do produtor não impede a reutilização válida da porta pública traduzida.
CSCwx48929	A classificação NAT pode se tornar E incorretamente quando a descoberta retorna ao mesmo vBond. A classificação correta não impede o cenário válido de reutilização de porta vBond diferente.

As correções do bug da Cisco ID CSCwx12986 e do bug da Cisco ID CSCwx48929 não garantem endpoints públicos anunciados exclusivos e não eliminam, por si só, esse cenário de colisão. O fechamento do registro de documentação do ID de bug da Cisco CSCww45867 não estabelece uma alteração no comportamento do software. Confirme as limitações e mitigações aplicáveis com o Cisco TAC.

Anexo: Decodificar o Trailer IPv4 VCONTROL

Salve este código como sdwan_vctrl.lua. O decodificador de diagnóstico não autentica pacotes nem descriptografa suas cargas. Ele usa o filtro de exibição sdwan_vctrl e registra uma opção explícita Decodificar como. A captura de pacotes permanece inalterada.

```
-- 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)
```

Informações Relacionadas

- RFC 7857, Seção 3: Comportamento de sobreposição de portas
- RFC 4787: Requisitos comportamentais de NAT para UDP unicast
- RFC 5737: Blocos de endereços IPv4 reservados para documentação
- Guia de design do Cisco Catalyst SD-WAN
- Comandos do sistema Cisco IOS XE Catalyst SD-WAN: deslocamento de porta
- Configuração do deslocamento de porta do Cisco SD-WAN
- Referência de Linha de Comando do Wireshark
- Referência de linha de comando do TShark
- Suporte Técnico e Documentação - Cisco Systems

Sobre esta tradução
A Cisco traduziu este documento com a ajuda de tecnologias de tradução automática e humana para oferecer conteúdo de suporte aos seus usuários no seu próprio idioma, independentemente da localização.

Observe que mesmo a melhor tradução automática não será tão precisa quanto as realizadas por um tradutor profissional.

A Cisco Systems, Inc. não se responsabiliza pela precisão destas traduções e recomenda que o documento original em inglês (link fornecido) seja sempre consultado.