Risoluzione dei problemi dei tunnel hub SD-WAN con endpoint NAT duplicati

Sommario

- Introduzione
- Prerequisiti
  - Requisiti
  - Componenti usati
- Premesse
  - Control-Plane NAT Learning
  - Apprendimento NAT Data-Plane e VCONTROL
  - Conservazione delle porte, sovrapposizione delle porte e casualizzazione delle porte
- Topologia di esempio
- Perché avviene la collisione
  - Diversi vBonds possono osservare lo stesso endpoint pubblico
  - Le Stesse Pubblicità Raggiungono Un Unico Hub
  - Collisioni tra privati e pubblici con indirizzi di branch ripetuti
  - Riassegnazione endpoint NAT
- Sintomi e impatto
- Diagnosticare il problema
  - Passaggio 1. Registrare lo stato del trasporto del software e dello spoke
  - Passaggio 2. Trovare endpoint duplicati in OMP
  - Passaggio 3. Confrontare i TLOC logici con la proprietà della sessione hub
  - Passaggio 4. Verificare le funzionalità effettive del gateway NAT
  - Passaggio 5. Ispezionare VCONTROL in Wireshark
  - Passaggio 6. Confermare il percorso interessato e conservare le prove
- Soluzioni e limitazioni
  - Utilizzo di mapping stabili indipendenti dagli endpoint con allocazioni univoche
  - Disabilita Sovrapposizione Porte Tra Spoke Diversi
  - Usa offset porta per un gruppo controllato di piccole dimensioni
  - Disabilita conservazione porta di origine o abilita allocazione casuale
  - Indirizzi pubblici o blocchi di porte di riserva separati
  - Mantieni trasporto alternativo e coordina recupero
- Problemi software correlati
- Appendice: Decodificare la sequenza finale del controllo IPv4
- Informazioni correlate

Introduzione

In questo documento viene descritto come identificare e ridurre gli errori del tunnel di hub quando gli spoke condividono un endpoint NAT (Network Address Translation) pubblico.

Prerequisiti

Requisiti

Cisco raccomanda la conoscenza dei seguenti argomenti:

- Architettura SD-WAN (Wide Area Network) definita dal software Cisco Catalyst e connessioni di controllo
- Overlay Management Protocol (OMP), Transport Locator (TLOC), Bidirectional Forwarding Detection (BFD) e IP Security (IPsec)
- Reti di trasporto Multiprotocol Label Switching (MPLS)
- Mapping NAT e comportamento filtro
- Analisi di base dei pacchetti con Wireshark

Componenti usati

Le informazioni fornite in questo documento si basano sulle seguenti versioni software e hardware:

- Router Cisco IOS® XE Catalyst SD-WAN
- Controller Cisco Catalyst SD-WAN (vSmart)
- Cisco Catalyst SD-WAN Validator (vBond)

L'errore di proprietà della sessione dell'hub è stato riprodotto su Cisco IOS XE 17.18.2 con chiavi pairwise operative. I campi relativi alla disponibilità dei comandi e all'output variano a seconda della piattaforma e della versione.

Negli esempi vengono utilizzati indirizzi privati o di documentazione riservati dalla RFC 5737 con nomi di dispositivo generici. Gli estratti della CLI sono abbreviati e usano identificatori normalizzati. I nomi delle interfacce e alcuni campi di output variano a seconda della piattaforma e della versione.

Le informazioni discusse in questo documento fanno riferimento a dispositivi usati in uno specifico ambiente di emulazione. Su tutti i dispositivi menzionati nel documento la configurazione è stata ripristinata ai valori predefiniti. Se la rete è operativa, valutare attentamente eventuali conseguenze derivanti dall'uso dei comandi.

⚠ Attenzione: Le modifiche alla configurazione e le reimpostazioni della connessione possono interrompere il traffico. Utilizzare un piano di manutenzione approvato. I comandi diagnostici iniziali sono di sola lettura.

Premesse

Iniziare con uno scenario in cui NAT simmetrico è identificato correttamente. Il spoke A si registra con vBond 1, mentre il spoke B si registra con vBond 2. Il gateway NAT può assegnare lo stesso indirizzo IP di origine tradotto e la stessa porta a queste connessioni quando le loro destinazioni differiscono, come permesso dalla RFC 7857. Entrambi gli spoke possono pertanto pubblicizzare lo stesso endpoint pubblico per i propri TLOC Internet pubblici, anche quando entrambi segnalano NAT di tipo A.

La topologia di esempio utilizza il colore Internet pubblico. Lo stesso scenario può coinvolgere altri colori pubblici, come il biz-internet. Applicare i filtri diagnostici al colore di trasporto effettivo; il valore letterale public-internet color non è una condizione obbligatoria.

✎ Nota: Cisco Catalyst SD-WAN non supporta i TLOC di siti diversi che pubblicizzano lo stesso indirizzo IP pubblico e la stessa porta UDP verso un perimetro di ricezione comune. Il comportamento del ricevitore in questa condizione di collisione è imprevedibile. Questo problema può verificarsi anche quando il protocollo NAT simmetrico è identificato correttamente. La corretta classificazione NAT da sola non garantisce l'univocità degli endpoint annunciati.

Control-Plane NAT Learning

Durante la registrazione con vBond, An Edge invia il proprio indirizzo di trasporto privato e la porta. vBond osserva l'indirizzo di origine e la porta su tale connessione e restituisce l'endpoint osservato al perimetro. The Edge può pubblicizzare queste informazioni tramite OMP come l'indirizzo pubblico e gli attributi della porta di un TLOC, incluso un TLOC con un colore pubblico come internet pubblico o internet biz.

L'osservazione descrive la connessione a tale vBond. Con un NAT dipendente dalla destinazione, la mappatura verso un hub può essere diversa. Il router confronta le osservazioni NAT per classificare il comportamento della mappatura. In show sdwan control local-properties, A indica il mapping dipendente dall'indirizzo e dalla porta e E indica il mapping indipendente dall'endpoint. La classificazione visualizzata descrive la decisione del router; una tabella di conversione NAT o l'acquisizione di pacchetti verifica il comportamento effettivo del gateway NAT.

Ad esempio, eseguire show sdwan control local-properties e individuare la colonna NAT TYPE per

il trasporto. Questo output abbreviato conserva la legenda di tipo NAT e le colonne di interfaccia scelte, con indirizzi normalizzati:

<#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
```

Il trasporto biz-internet segnala A e pubblicizza 203.0.113.10:23012 come endpoint pubblico. Il trasporto MPLS segnala E e ha endpoint pubblici e privati uguali. leggere la classificazione da NAT TYPE; l'uguaglianza di indirizzi o porte da sola non identifica il tipo NAT. N indica che il router non ha appreso una classificazione.

Apprendimento NAT Data-Plane e VCONTROL

Prima di poter elaborare un pacchetto BFD crittografato, un hub deve trovare l'associazione di sicurezza IPsec (SA) in entrata per il peer e il tunnel corrispondenti. L'associazione di sicurezza identifica la chiave crittografica e l'altro stato necessario per autenticare e decrittografare il pacchetto. Per individuare lo stato, il destinatario utilizza le informazioni sull'endpoint insieme all'indice dei parametri di sicurezza (SPI) ESP (Encapsulating Security Payload). L'indice SPI da solo non è un identificatore spoke univoco globale.

Un spoke dietro un NAT simmetrico in genere avvia il traffico verso un hub raggiungibile. In questo modo viene creato un mapping per la destinazione specifica. L'indirizzo di origine o la porta visualizzata dall'hub può essere diversa dall'endpoint pubblico appreso in precedenza tramite vBond e pubblicizzato in OMP. Di conseguenza, le intestazioni esterne ricevute non sempre corrispondono all'endpoint nello stato del tunnel installato. L'hub deve essere in grado di associare il nuovo endpoint in transito al TLOC remoto noto e al relativo stato IPSec prima di poter decrittografare i dati BFD.

Per il formato IPv4 BFD-over-IPsec discusso qui, SD-WAN inserisce le informazioni dell'endpoint in una sequenza terminale VCONTROL visibile alla fine del payload UDP, all'esterno del payload BFD criptato. L'hub è in grado di leggere queste informazioni prima della decrittografia. VCONTROL trasporta le informazioni sull'endpoint, non le chiavi di crittografia, e non stabilisce

l'identità peer da sola.

Ad esempio, durante la creazione iniziale del tunnel, prima del completamento dell'apprendimento del piano dati:

```
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
```

Il ricevitore tenta per primo la normale ricerca di un endpoint/SPI. Se la ricerca non trova un'associazione di protezione, l'endpoint di origine di VCONTROL fornisce un fallback. Nell'esempio, l'hub può utilizzare il router 203.0.113.10:12346 trasportato, insieme al contesto di destinazione locale e all'interfaccia SPI, per individuare l'associazione di protezione in entrata associata al TLOC remoto noto, anche se il pacchetto arriva da 203.0.113.10:45678.

L'associazione di protezione scelta fornisce il contesto della chiave di ricezione. L'indice SPI sceglie la generazione di chiavi applicabile all'interno di tale contesto. L'hub esegue quindi l'autenticazione e la decrittografia IPsec. solo un pacchetto che supera i controlli richiesti è disponibile per l'elaborazione BFD. Una sequenza finale leggibile o una ricerca riuscita non consentono di eseguire correttamente l'autenticazione, la decrittografia o la ricezione BFD.

La differenza tra le porte di origine osservate e quelle trasportate può quindi essere normale con il protocollo NAT simmetrico. Affinché la relazione peer funzioni, l'hub deve associare il pacchetto allo stato corretto del tunnel e della chiave e conoscere l'endpoint di ritorno effettivo. In questo esempio, l'apprendimento riuscito consente all'hub di inviare il traffico di ritorno a 203.0.113.10:45678. Si tratta di un'operazione distinta dall'apprendimento del control plane tramite vBond.

VCONTROL non impone al ricevitore di ignorare un endpoint/SPI corrispondente esistente. Se lo stato dell'endpoint duplicato associa un pacchetto al peer errato o alla generazione della chiave, la sequenza terminale non garantisce il ripristino. Fornisce un suggerimento di ricerca; non è possibile distinguere due peer che trasportano lo stesso endpoint senza lo stato di crittografia e il tunnel circostante corretto.

VCONTROL è presente anche sui pacchetti BFD applicabili quando è segnalato il tipo NAT E. La sua presenza non è un flag symmetric-NAT e non è inclusa in ogni pacchetto di dati IPsec ordinario. Per il formato BFD IPv4 descritto qui, le informazioni public-TLOC vengono omesse tra due colori privati.

L'endpoint specifico del peer appreso tramite il piano dati non significa che il spoke deve sostituire il proprio annuncio OMP con un nuovo endpoint per ogni destinazione. Mapping appresi diversi

sono una parte prevista dell'attraversamento NAT simmetrico.

Quando un tunnel è attivo, un mittente che ha appreso la propria conversione dell'endpoint per quel tunnel può trasportare quell'endpoint come informazioni di origine. Questo è l'endpoint pubblico per tunnel del mittente, non l'endpoint remoto o un nuovo annuncio OMP globale. Confrontare sempre il pacchetto, l'annuncio OMP e lo stato sessione dallo stesso periodo di osservazione.

Conservazione delle porte, sovrapposizione delle porte e casualizzazione delle porte

Si tratta di comportamenti diversi:

Comportamento	Significato
Conservazione della porta di origine	Il gateway NAT tenta di mantenere la porta di origine UDP originale dello spoke, ad esempio, 12346, come porta di origine tradotta.
Sovrapposizione delle porte tra le destinazioni	Il gateway NAT riutilizza un IP tradotto: per endpoint interni diversi che comunicano con destinazioni esterne diverse.
Assegnazione casuale delle porte	Il gateway NAT sceglie le porte tradotte utilizzando un metodo di allocazione casuale.
Mapping indipendente dagli endpoint	Un endpoint interno conserva lo stesso endpoint tradotto tra le destinazioni esterne mentre il mapping esiste.

La RFC 7857, sezione 3, consente la sovrapposizione delle porte esterne per endpoint interni diversi quando le destinazioni sono diverse. Il NAT distingue i flussi usando le loro cinque tuple. Ciò non consente due flussi UDP simultanei e indistinguibili verso la stessa destinazione. L'allocazione casuale e la conservazione sono scelte di allocazione separate; nessuna delle due stabilisce l'indipendenza dell'endpoint.

Topologia di esempio

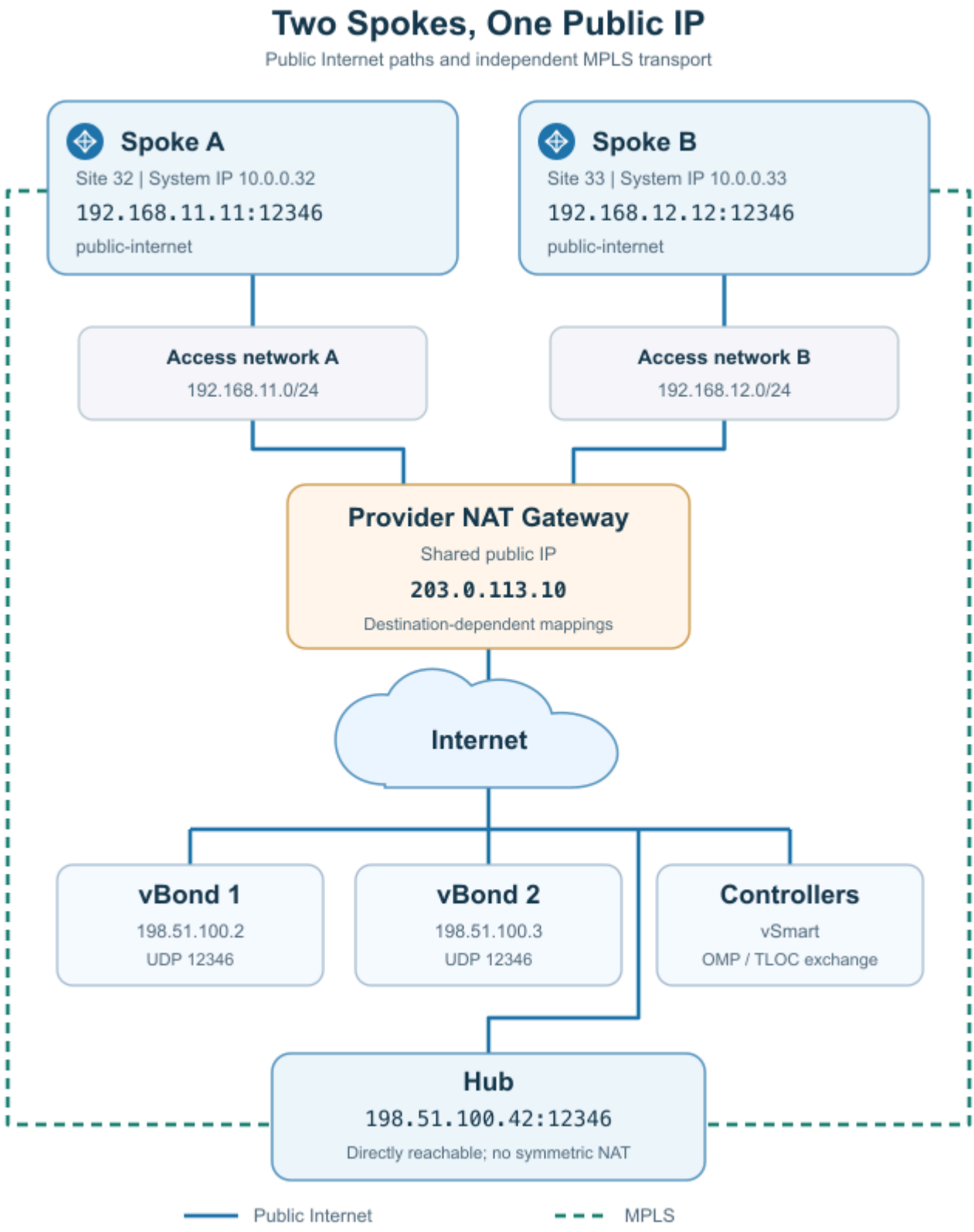

Figura 1. Gateway NAT condiviso, reti di accesso spoke separate e trasporto MPLS indipendente.

Il raggio A utilizza 192.168.11.11:12346, mentre il raggio B utilizza 192.168.12.12:12346. Entrambi condividono l'indirizzo pubblico 203.0.113.10 tramite il gateway NAT. L'hub ha un indirizzo raggiungibile direttamente e non si trova dietro un NAT simmetrico. I due raggi possono pubblicizzare la stessa porta IP:port pubblica anche se i loro endpoint privati sono diversi.

Perché avviene la collisione

Diversi vBonds possono osservare lo stesso endpoint pubblico

Prendiamo in considerazione i seguenti mapping UDP simultanei:

Peer	Origine privata	Origine tradotta	Destinazione
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

Le risposte vengono restituite correttamente perché NAT controlla anche l'endpoint remoto. Una risposta da vBond 1 appartiene al mapping per A; una risposta da vBond 2 appartiene al mapping per B.

Entrambi i router possono inoltre contattare l'altro vBond e rilevare correttamente il mapping dipendente dalla destinazione. Ad esempio, mantenete invariata la tupla di origine privata per ogni raggio e tenete presente questo ordine di osservazione:

Peer	Prima osservazione accettata su vBond	Confronto con altri vBond	Risultato
A	vBond 1 restituisce 203.0.113.10:12346	vBond 2 restituisce 203.0.113.10:40001	NAT A tupla pubblica mantenuta 203.0.113.10:12346
B	vBond 2 restituisce 203.0.113.10:12346	vBond 1 restituisce 203.0.113.10:4002	NAT A tupla pubblica mantenuta 203.0.113.10:12346

Le osservazioni di vBond sono diverse per ciascun router, quindi il rilevamento di NAT A è corretto. Le prime osservazioni mantenute corrispondono tuttavia tra i due router. La Classificazione A non assegna una porta pubblica unica tra tutti i raggi e visitare entrambi i vBonds non garantisce tale unicità.

Ogni spoke riceve l'indirizzo pubblico e il porto osservati da vBond nella risposta di registrazione e utilizza l'osservazione mantenuta per popolare le sue proprietà di trasporto locale. Lo spoke include l'endpoint pubblico, insieme al relativo endpoint privato, negli attributi TLOC che pubblicizza al controller SD-WAN tramite OMP. Il controllore distribuisce quindi il TLOC ad altri spigoli in base alla politica di controllo applicabile; vBond stesso non genera la pubblicità OMP.

Se entrambi i raggi mantengono lo stesso endpoint pubblico osservato, OMP può pertanto contenere:

```
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
```

Questi TLOC logici rimangono diversi perché i loro IP di sistema sono diversi.

Le Stesse Pubblicità Raggiungono Un Unico Hub

Il gateway NAT deve distinguere entrambi i spoke quando contattano lo stesso hub. Una delle possibili allocazioni è:

Peer	Origine convertita verso hub	Destinazione
A	203.0.113.10:12346	198.51.100.42:12346
B	203.0.113.10:45678	198.51.100.42:12346

L'hub ha due peer logici che pubblicizzano lo stesso endpoint di destinazione, mentre i pacchetti possono arrivare da endpoint effettivi diversi. In normali operazioni NAT simmetriche, l'apprendimento del piano dati consente di stabilire le mappature di ritorno.

La limitazione della proprietà della sessione hub si verifica quando diversi peer condividono lo stato della sessione scelto da questa chiave di trasporto:

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

Se entrambi gli annunci scelgono la stessa chiave, un aggiornamento da un peer successivo può sostituire le informazioni sulla proprietà nello stato condiviso. Un successivo ritiro, eliminazione o ricreazione della sessione per un peer può disturbare lo stato IPsec/BFD per l'altro peer. L'apprendimento del piano dati non garantisce il ripristino quando tale stato è già associato al peer errato o alla generazione del flusso.

Il trailer visibile non è un identificatore peer indipendente: entrambi i raggi possono avere lo stesso endpoint di origine. Se l'hub sceglie uno stato IPSec esistente che appartiene al peer errato o alla generazione di chiavi, l'autenticazione e l'apprendimento possono avere esito negativo anche se i

pacchetti raggiungono l'hub. Per stabilire la sequenza per un pacchetto specifico, il Cisco Technical Assistance Center (TAC) deve correlare l'associazione di protezione scelta con il peer e la generazione. Il trailer del pacchetto da solo non può dimostrarlo.

La stessa proprietà intellettuale pubblica da sola non determina una collisione. Porte remote diverse producono in genere chiavi di trasporto diverse. Verificare inoltre l'indirizzo locale dell'hub, la porta locale e il protocollo scelti. Se un terzo peer annuncia lo stesso endpoint, può anche competere per tale chiave; il modello di errore dipende dall'ordine degli eventi.

Collisioni tra privati e pubblici con indirizzi di branch ripetuti

L'ID bug Cisco CSCwx12986 può causare la pubblicità da parte di un edge di una tupla RFC 1918 privata negli attributi TLOC pubblici dopo un hop della porta di origine automatico durante il confronto NAT. Se diramazioni diverse riutilizzano lo stesso indirizzo privato e la stessa porta, gli attributi pubblici non corretti possono corrispondere tra peer ed esporre il problema di proprietà della sessione dell'hub.

Ad esempio, entrambi i raggi possono usare 192.168.11.11 dietro firewall di diramazione separati. Ogni firewall mantiene separata la propria rete di filiali e converte l'indirizzo di origine. Se un gateway NAT di un provider upstream lo converte di nuovo, il percorso ha due stadi NAT, o un doppio NAT. I contesti VRF (Virtual Routing and Forwarding) separati rappresentano un altro modo per mantenere isolati gli indirizzi di diramazione sovrapposti.

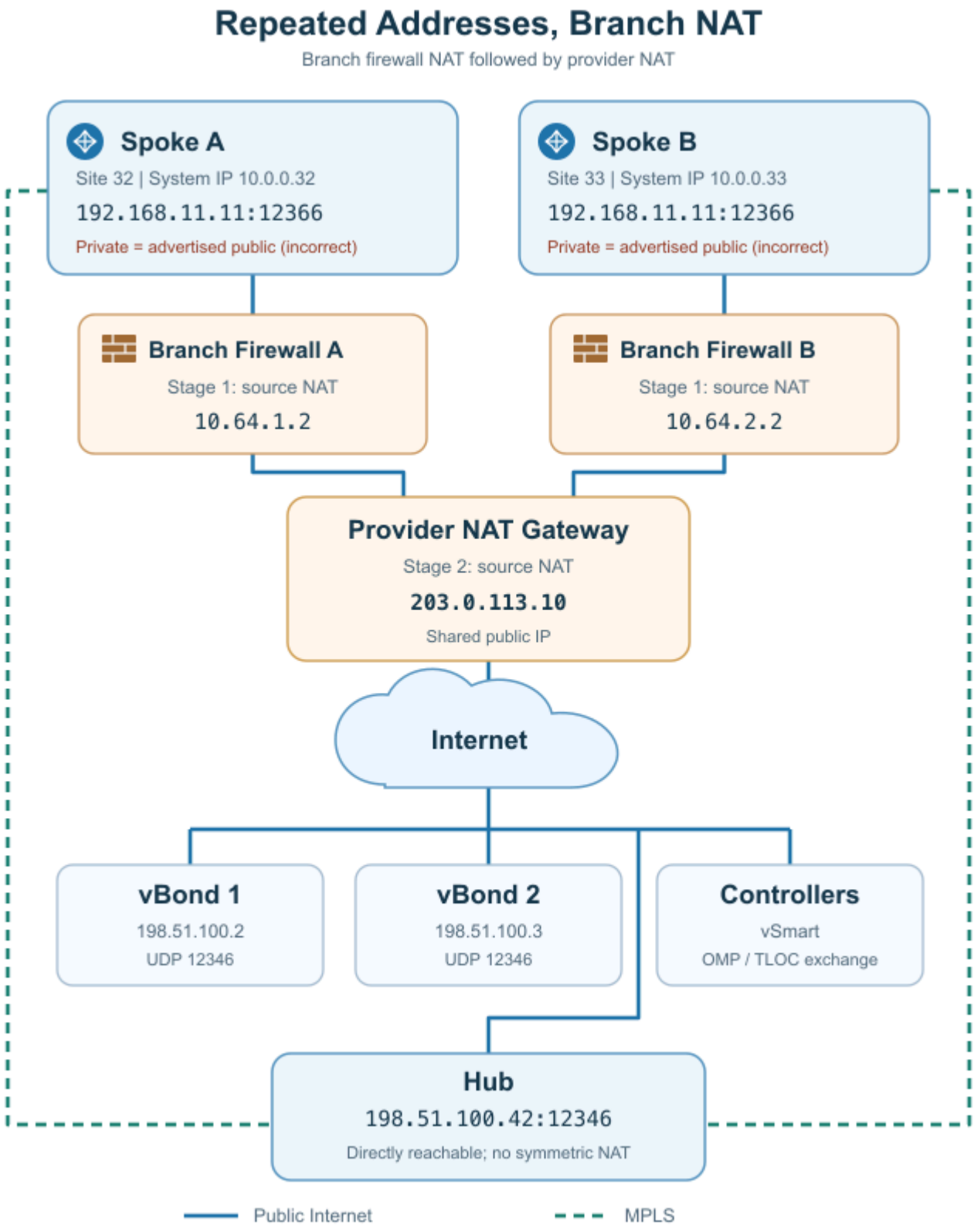

Figura 2. Creazione di diramazioni del firewall e traduzioni NAT del provider. Gli attributi TLOC pubblici non corretti mantengono l'endpoint privato originale, mentre i pacchetti utilizzano gli indirizzi di origine tradotti.

Se entrambi gli spoke rilevano il difetto sui propri TLOC Internet pubblici e utilizzano la stessa porta di origine post-hop, l'hub può ricevere questi annunci:

Peer	Endpoint privato dopo l'hop della porta	Endpoint TLOC pubblico non corretto
A	192.168.11.11:12366	192.168.11.11:12366
B	192.168.11.11:12366	192.168.11.11:12366

Nota: Gli indirizzi privati ripetuti e il doppio NAT non sono prerequisiti per l'ID bug Cisco CSCwx12986 stesso. Per l'ulteriore collisione tra peer è necessario che i valori IP:port annunciati corrispondano e che l'indirizzo locale dell'hub, la porta e il protocollo siano uguali. La collisione non viene stabilita dallo stesso IP privato con porte annunciate diverse.

L'uguaglianza degli indirizzi pubblici e privati è normale su un trasporto senza NAT. Un indirizzo RFC 1918 in un campo pubblico è sospetto solo quando contraddice il percorso previsto e la traduzione osservata. Gli indirizzi privati di diramazione univoci evitano la sovrapposizione di tuple private ripetute in questo esempio; non impediscono lo scenario separato di riutilizzo della porta pubblico convertito descritto in precedenza.

Riassegnazione endpoint NAT

Un endpoint può inoltre essere riassegnato a un altro spoke dopo la scadenza di un mapping NAT mentre l'hub mantiene ancora lo stato peer precedente. Confrontare la cronologia del mapping con il ciclo di vita della sessione hub per identificare la sovrapposizione tra i proprietari vecchi e nuovi.

Sintomi e impatto

Gli indicatori tipici includono:

- Le connessioni di controllo rimangono attive mentre i percorsi BFD interessati sono inattivi o non ripristinati.
- Due IP di sistema diversi pubblicizzano la stessa porta IP:port pubblica.
- L'hub conserva entrambi i TLOC logici, ma in una vista di sessione di piattaforma filtrata manca un peer o vengono visualizzate informazioni incoerenti relative all'IP del sistema, al colore remoto, al discriminatore, al tunnel o all'associazione di protezione.
- Un ritiro o uno svago del trasporto per un peer disturba un altro peer con lo stesso endpoint.
- Le perdite di input IPSec possono aumentare sul percorso interessato.

L'errore può rendere inutilizzabile il percorso di trasporto interessato. Se questo è l'unico percorso di trasporto utilizzabile per un raggio, il raggio viene isolato dalla sovrapposizione. In un hub di grandi dimensioni, molti spoke che condividono un indirizzo NAT aumentano la popolazione che

richiede correlazione, ma il numero di percorsi interessati deve essere misurato; non può essere dedotto dal numero di siti che si trovano dietro tale indirizzo.

Diagnosticare il problema

Passaggio 1. Registrare lo stato del trasporto del software e dello spoke

Su ciascun perimetro Cisco IOS XE interessato, sostituire **<**AFFECTED-color**>** con il colore di trasporto effettivo, ad esempio biz-internet o 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

Registrare l'indirizzo pubblico/privato e la porta, l'interfaccia, il colore, il tipo NAT, l'ID del sito e lo stato del controllo. Dopo il primo esempio completo di proprietà locali, utilizzare questo filtro compatto per conservare le intestazioni di interfaccia e la riga di trasporto scelta. Sostituire TenGigabit Ethernet0/1/0.10 con l'interfaccia interessata:

<#root>

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

Per lo scenario principale, i campi rilevanti possono essere letti su entrambi i raggi. Questo è un riepilogo di campi, non il layout di colonna 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
```

Una connessione di controllo Su verifica la relazione tra controller. Non verifica un percorso BFD spoke-to-hub o l'univocità dell'endpoint pubblico.

Passaggio 2. Trovare endpoint duplicati in OMP

Sull'hub Cisco IOS XE:

<#root>

show clock detail

show sdwan omp tlocs table | include 203.0.113.10

Sul controller SD-WAN (vSmart):

<#root>

show clock

show omp tlocs | tab | include 203.0.113.10

Questo output del controller utilizza identificatori normalizzati. Le colonne IPv6 e region sono omesse:

<#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
```

Confrontare l'IP del sistema TLOC, il colore, l'incapsulamento, l'IP pubblico e la porta pubblica. Più percorsi allo stesso TLOC logico da controller diversi sono normali. Il candidato per la collisione è costituito da due identità peer diverse con lo stesso endpoint effettivo.

Ripetere il comando filtrato dopo 30-60 secondi e conservare entrambi i timestamp. Una sovrapposizione persistente è un'utile prova; non sono richiesti comandi perfettamente simultanei. Per un breve periodo transitorio, raccogliere gli eventi OMP e i record del ciclo di vita delle sessioni corrispondenti con TAC.

Per due peer sospetti, mantenere le singole voci della tabella TLOC:

<#root>

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

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

Per una visualizzazione compatta dei dettagli TLOC IPv4 su Cisco IOS XE, usare questo filtro. Sostituire l'indirizzo IP del sistema con il peer sotto indagine e mantenere le virgolette attorno al selettore 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|::
```

Confrontare gli attributi di porta/indirizzo pubblico e privato per entrambi i peer. Il filtro omette da questa visualizzazione le righe non impostate, con valore zero, predefinite e IPv6 corrispondenti.

Passaggio 3. Confrontare i TLOC logici con la proprietà della sessione hub

Ispezionare lo stato TLOC consegnato a FTM. Confrontare gli IP di sistema, i colori, gli indirizzi IP pubblici e le porte UDP pubbliche di entrambi i peer.

<#root>

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

Eseguire questi comandi per confrontare i record di Forwarding Table Manager (FTM), BFD e sessione di piattaforma:

<#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

Nell'esempio viene mostrato il modello con identificatori normalizzati. Per maggiore chiarezza, le etichette di colonna vengono aggiunte alle righe FTM filtrate; non fanno parte dell'output del comando filtrato:

<#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 e FTM mantengono entrambe le identità peer logiche, mentre questa ricerca di sessione mostra solo il raggio A per la destinazione condivisa. Anche lo stato FTM di Spoke A è inattivo; la presenza nella tabella della sessione della piattaforma non indica un percorso integro. L'indice FTM TLOC è distinto dal discriminatore locale BFD (BFD-LD) nella tabella delle sessioni. Questo pattern è un motivo per analizzare la proprietà tra i livelli. Una riga CLI mancante da sola non determina una perdita di pacchetto.

In questa tabella, Color rappresenta il trasporto locale; utilizzare Remote color nella vista dettagliata per identificare il TLOC remoto. Confrontare anche DstIp e DPubIp: una destinazione effettiva appresa può essere diversa dall'endpoint pubblico associato alla sessione. Utilizzare la visualizzazione specifica del peer se un filtro di sola destinazione non corrisponde più all'indirizzo appreso.

Un hop successivo non valido, l'identificatore SA 0 o In HW: FALSE fornisce ulteriori prove anziché una firma univoca. Questi campi possono inoltre verificarsi durante le normali operazioni di configurazione o in caso di errori di sovrapposizione.

Passaggio 4. Verificare le funzionalità effettive del gateway NAT

Chiedere all'operatore NAT o al provider di servizi i record di conversione UDP simultanei con questi campi:

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

Questa è una tabella di traduzione illustrativa. Utilizzare il comando di diagnostica equivalente per la piattaforma NAT. Su un gateway Cisco IOS XE NAT, iniziare con:

<#root>

show ip nat translations | include 203.0.113.10

Conservare sia le colonne interne di origine che di destinazione. Un elenco contenente solo l'origine tradotta non può mostrare perché il riutilizzo è valido per vBonds diversi o quale spoke possiede il mapping hub.

Acquisire i flussi UDP interessati prima e dopo NAT per vBonds e l'hub. Mantenere intere le lunghezze dei pacchetti in modo che la sequenza terminale di VCONTROL non venga troncata. Utilizzare una raccolta SPAN (Switched Port Analyzer) o EPC (Embedded Packet Capture) concordata con TAC. confrontare lo stesso flusso tra i punti di acquisizione; non dedurre la latenza dagli orologi di acquisizione non sincronizzati.

Passaggio 5. Ispezionare VCONTROL in Wireshark

Usare il decodificatore Lua descritto nell'appendice per il formato IPv4 BFD/VCONTROL qui descritto. Salvare il codice come sdwan_vctrl.lua nella stessa cartella dell'acquisizione. Questo decodificatore diagnostico visualizza i campi visibili; non decrittografa ESP né verifica l'autenticazione. Il suo ambito è costituito da pacchetti IPv4/UDP completi e non nidificati con il flag BFD/VCONTROL previsto e la struttura finale.

Eseguire il comando Wireshark dalla cartella contenente il decoder e la cattura:

<#root>

```
wireshark -n \

  -X lua_script:sdwan_vctrl.lua \

  -d udp.port==12346,sdwan_vctrl \

  -d udp.port==45678,sdwan_vctrl \

  -r hub.pcap
```

Sostituire le porte di esempio con quelle nell'acquisizione. In Windows o macOS, utilizzare il percorso completo dell'eseguibile Wireshark se non è in PATH.

In alternativa, caricare lo script Lua all'avvio di Wireshark, scegliere un pacchetto UDP rilevante, quindi selezionare Analyze > Decode As. Assegnare le voci della porta UDP di origine e di destinazione applicabili a SDWAN_VCTRL / SD-WAN VCTRL (IPv4). Una voce UDPENCAP esplicita sull'altra porta può avere la precedenza su questo decoder. Applicate la modifica a questa sessione di analisi.

Applica questo filtro di visualizzazione:

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

Espandere SD-WAN VCTRL > Trailer di Visible VCONTROL e confrontare:

```
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
```

Questi valori illustrativi descrivono la creazione iniziale del tunnel prima del completamento dell'apprendimento del piano dati. Mostrano l'effettiva mappatura hub per B lungo il relativo endpoint trasportato. NAT modifica le intestazioni esterne; non riscrive le informazioni sull'endpoint all'interno della sequenza finale. Questa mancata corrispondenza da sola non dimostra un errore.

Eseguire il comando TShark per un'esportazione compatta:

<#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
```

Senza un decodificatore personalizzato, il formato IPv4 qui descritto ha questa sequenza terminale TLV (Type-Length-Value) di 17 byte alla fine del payload UDP. La spaziatura interna a livello di collegamento è esterna al 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
```

Non applicare questo layout fisso a IPv6, altri formati di controllo, pacchetti troncati o dati crittografati arbitrari. Un filtro decodificatore vuoto può indicare che il pacchetto non è nel formato supportato. Anche un endpoint di origine leggibile non è un'etichetta IP di sistema autenticata.

Nota: Gli avvisi di sequenza ESP generici in Wireshark non sono diagnostica anti-replay del router. Correlare il pacchetto con i contatori e i registri del router prima di trarre una conclusione sull'autenticazione o sulla proprietà.

Passaggio 6. Confermare il percorso interessato e conservare le prove

Sull'hub, prelevare due campioni di contatore del processore QuantumFlow (QFP) attorno al guasto osservato:

<#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

Utilizzare le differenze dei contatori anziché i totali assoluti. Le interruzioni IPsec aggregate possono avere altre cause. Confermare l'arrivo del pacchetto sul trasporto interessato e correlarlo al peer, all'endpoint, al discriminatore BFD, al tunnel, all'associazione di sicurezza e all'ora dell'evento.

Fornire a TAC le due istantanee OMP, le proprietà e i controlli locali spoke, i dettagli della

sessione hub per entrambi i peer, lo stato BFD, i record di traduzione, le acquisizioni limitate e le versioni software. Se già disponibili, includere gli admin-tech rilevanti e la cronologia delle connessioni/revoche. Evitare la raccolta ripetuta di intere tabelle in un hub di grandi dimensioni.

Verificare la raggiungibilità del servizio separatamente e registrare il trasporto effettivamente scelto. Un'indicazione di aumento BFD di 0 secondi non indica un ripristino sostenuto.

Soluzioni e limitazioni

Valuta queste opzioni di mitigazione con Cisco TAC e l'operatore NAT. La limitazione riguarda i duplicati degli endpoint pubblicizzati che raggiungono un hub comune; ciò non significa che NAT simmetrico, indirizzi pubblici condivisi o indirizzi privati ripetuti non siano generalmente supportati. Le procedure di ripristino e l'efficacia della riduzione dei rischi dipendono dall'implementazione.

Utilizzo di mapping stabili indipendenti dagli endpoint con allocazioni univoche

Un'opzione di prevenzione lato NAT per la sequenza illustrata è quella di fornire a ciascun endpoint di trasporto spoke un endpoint UDP stabile e tradotto in modo univoco tra destinazioni vBond, controller e hub. Richiedi indipendenza endpoint e univocità tra i raggi. Verificare il risultato nei record di traduzione e nelle acquisizioni, anche dopo il failover del gateway NAT o la ricreazione dello stato.

Ad esempio, A può utilizzare coerentemente 203.0.113.10:40001, mentre B utilizza coerentemente 203.0.113.10:40002. Traduzioni statiche o criteri di allocazione appropriati possono fornire questo comportamento, in base alla piattaforma NAT e alla progettazione degli accessi.

Il mapping indipendente dall'endpoint non richiede l'accesso in ingresso senza restrizioni. Mantenere i controlli firewall richiesti. Una correzione della policy NAT non rimuove automaticamente lo stato di sessione annunciato o programmato in precedenza; coordinare un nuovo ciclo di scoperta e recupero con TAC.

Disabilita Sovrapposizione Porte Tra Spoke Diversi

Se supportato, riservare gli endpoint UDP tradotti in modo univoco nei relativi spoke SD-WAN. Non riutilizzarle solo perché le loro destinazioni sono diverse. In questo modo viene rimosso direttamente il mapping vBond duplicato mostrato in precedenza, anche se il gateway NAT mantiene alcuni comportamenti dipendenti dalla destinazione.

L'impostazione può utilizzare più porte o richiedere indirizzi pubblici aggiuntivi. Non rende il mapping per ogni endpoint spoke indipendente da se stesso e non risolve lo stato obsoleto del ricevitore dopo una riassegnazione successiva dell'endpoint.

Con più stadi NAT o gateway NAT ridondanti, verificare l'allocazione visibile al limite Internet finale. L'univocità all'interno di un'istanza NAT non determina l'univocità tra altre istanze che condividono lo stesso indirizzo pubblico.

Usa offset porta per un gruppo controllato di piccole dimensioni

Il comando port-offset fornisce i valori da 0 a 19. Ad esempio, il raggio A può mantenere l'offset 0 e il raggio B può utilizzare l'offset 1. Eseguire questi comandi di configurazione sul raggio 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
```

Verificare le porte di origine attive e convertite dopo la modifica. Una modifica a questa impostazione di sistema può ricreare le connessioni dati e di controllo tra i trasporti. Non si tratta di un passaggio diagnostico senza interruzioni.

I 20 offset forniscono 20 gruppi di porte di base distinti, con offset 0 come scelta unica. Non sono un limite generale di 20 router SD-WAN dietro un IP pubblico. Questa particolare soluzione fornisce solo 20 scelte di offset univoche e richiede il gateway NAT per conservarle e distinguerle. Quando gli offset vengono riutilizzati, lo schema fallisce come schema di univocità. L'hop della porta non aggiunge altre identità di offset.

Ad esempio, 124 router non possono ricevere tutti offset univoci dietro un indirizzo condiviso. Uno

schema di sola offset con un massimo di 20 router per gruppo di indirizzi pubblici richiede almeno sette gruppi. Questa aritmetica non è una raccomandazione sulla capacità né un sostituto per la verifica del comportamento NAT. Un gateway NAT che riscrive o si sovrappone alle porte tradotte può ignorare lo schema anche con un numero inferiore di router.

Disabilita conservazione porta di origine o abilita allocazione casuale

Una politica senza una preferenza di conservazione può ridurre la tendenza per molti raggi a ricevere :12346 verso diversi vBonds. L'allocazione casuale può distribuire questi mapping su un pool di porte di dimensioni maggiori.

Tuttavia, né il cambiamento da solo garantisce la prevenzione delle collisioni. Un allocatore casuale può ancora riutilizzare una porta tra le destinazioni, modificare un mapping verso l'hub o riassegnare un endpoint scaduto mentre rimane lo stato precedente dell'hub. Verificare se l'allocazione è univoca tra sottoscrittori e destinazioni. Le sole porte casuali non stabiliscono l'univocità.

Verificare le impostazioni di conservazione e allocazione della porta di origine sul gateway NAT con il relativo operatore. La configurazione applicabile dipende dalla piattaforma NAT; verificare i mapping risultanti verso vBonds e l'hub con i record di conversione e le acquisizioni di pacchetti.

Le porte convertite possono non rientrare nell'intervallo di porte base SD-WAN originale. Esaminare le regole del firewall in modo che non rifiutino il traffico tradotto legittimo solo perché la porta di origine è stata modificata.

Indirizzi pubblici o blocchi di porte di riserva separati

Indirizzi tradotti dedicati o blocchi di porte non sovrapposti riservati per spoke possono fornire endpoint annunciati distinti per spoke attivi diversi. La policy deve essere applicata in modo coerente sui percorsi di discovery e hub e sui nodi NAT ridondanti. È possibile richiedere ulteriore spazio di indirizzi e supporto del provider. Il riutilizzo dinamico dopo il ritiro del mapping richiede ancora una corretta gestione del ciclo di vita della sessione hub.

Mantieni trasporto alternativo e coordina recupero

Un MPLS verificato o un altro percorso idoneo può mantenere il servizio durante l'analisi del percorso Internet interessato. Un elenco di controllo di accesso (ACL) sottostante che blocca l'indirizzo NAT condiviso fornisce solo il contenimento: può rimuovere la ridondanza di Internet per ogni spoke dietro l'indirizzo e non ritira automaticamente la loro pubblicità OMP.

Lo stato può essere temporaneamente ricreato tramite cancellazione del controllo, reimpostazione dell'interfaccia, cancellazione dello stato NAT o ricaricamento dell'hub. Nessuno garantisce la prevenzione del ripetersi. Non eseguire cancellazioni ampie come primo passo diagnostico. Preservare le prove, correggere le condizioni di assegnazione ove possibile e concordare l'azione di recupero più piccola con il TAC. Convalidare il traffico BFD attivo e bidirezionale sostenuto sul percorso ripristinato.

Problemi software correlati

ID bug Cisco	Ambito di applicazione del presente articolo
CSCww45867	Documentazione della limitazione della proprietà della sessione hub quando peer diversi pubblicizzano lo stesso endpoint TLOC pubblico. Questo record tiene traccia dell'aggiornamento della documentazione, anziché di una correzione software.
CSCwx12986	Una sequenza di apprendimento automatica port-hop/NAT può lasciare una tupla privata nei campi TLOC pubblici. Se i branch riutilizzano la stessa porta IP:port privata, gli endpoint pubblici errati possono corrispondere tra peer. La correzione di questo difetto del produttore non impedisce il riutilizzo valido della porta pubblica tradotta.
CSCwx48929	La classificazione NAT può diventare erroneamente E quando il rilevamento torna allo stesso vBond. La corretta classificazione non impedisce lo scenario di riutilizzo della porta differente-vBond valido.

Le correzioni per l'ID bug Cisco CSCwx12986 e l'ID bug Cisco CSCwx48929 non garantiscono endpoint pubblici annunciati univoci e non eliminano da sole questo scenario di collisione. La chiusura del record della documentazione relativa all'ID bug Cisco CSCww45867 non determina una modifica nel comportamento del software. Confermare le limitazioni e le mitigazioni applicabili con Cisco TAC.

Appendice: Decodificare la sequenza finale del controllo IPv4

Salvare il codice come sdwan_vctrl.lua. Il decodificatore diagnostico non autentica i pacchetti né decrittografa i relativi payload. Utilizza il filtro di visualizzazione sdwan_vctrl e registra un'opzione Decodifica come esplicita. L'acquisizione del pacchetto rimane invariata.

```
-- 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)
```

Informazioni correlate

- RFC 7857, sezione 3: Comportamento di sovrapposizione delle porte
- RFC 4787: Requisiti comportamentali NAT per UDP unicast
- RFC 5737: Blocchi di indirizzi IPv4 riservati alla documentazione
- Guida alla progettazione di Cisco Catalyst SD-WAN
- Comandi di sistema Cisco IOS XE Catalyst SD-WAN: port-offset
- Cisco SD-WAN Port Offset Configuration
- Informazioni di riferimento per la riga di comando di Wireshark
- Informazioni di riferimento per la riga di comando di TShark
- Documentazione e supporto tecnico – Cisco Systems

Informazioni su questa traduzione

Cisco ha tradotto questo documento utilizzando una combinazione di tecnologie automatiche e umane per offrire ai nostri utenti in tutto il mondo contenuti di supporto nella propria lingua. Si noti che anche la migliore traduzione automatica non sarà mai accurata come quella fornita da un traduttore professionista. Cisco Systems, Inc. non si assume alcuna responsabilità per l’accuratezza di queste traduzioni e consiglia di consultare sempre il documento originale in inglese (disponibile al link fornito).