Dépannage des tunnels de concentrateur SD-WAN avec des terminaux NAT dupliqués

Table des matières

- Introduction
- Conditions préalables
  - Exigences
  - Composants utilisés
- Informations générales
  - Apprentissage NAT sur le plan de contrôle
  - Apprentissage NAT du plan de données et VCONTROL
  - Conservation des ports, chevauchement des ports et randomisation des ports
- Exemple de topologie
- Pourquoi la collision se produit
  - Différents vBonds peuvent observer le même terminal public
  - Les Mêmes Annonces Atteignent Un Concentrateur
  - Collisions « privé-public » avec adresses de filiale répétées
  - Réaffectation des terminaux NAT
- Symptômes et impact
- Diagnostiquer le problème
  - Étape 1 : enregistrement de l'état de transport du logiciel et du satellite
  - Étape 2. Recherche des terminaux en double dans OMP
  - Étape 3. Comparaison des TLOC logiques avec la propriété de session Hub
  - Étape 4. Vérification de la fonction réelle de la passerelle NAT
  - Étape 5. Inspection de VCONTROL dans Wireshark
  - Étape 6. Confirmation du chemin affecté et conservation des preuves
- Solutions de contournement et limitations
  - Utiliser des mappages stables indépendants des terminaux avec des allocations uniques
  - Désactiver le chevauchement de ports sur différents rayons
  - Utiliser le décalage de port pour un petit groupe contrôlé
  - Désactiver la conservation du port source ou activer l'allocation aléatoire
  - Adresses publiques distinctes ou blocs de ports de réserve
  - Conserver un autre transport et coordonner la récupération
- Défauts logiciels associés
- Annexe : Décodage de la queue de bande IPv4 VCONTROL
- Informations connexes

Introduction

Ce document décrit comment identifier et atténuer les défaillances du tunnel de concentrateur lorsque les rayons partagent un point d'extrémité NAT (Network Address Translation) public.

Conditions préalables

Exigences

Cisco vous recommande de prendre connaissance des rubriques suivantes :

- Architecture de réseau étendu (SD-WAN) définie par logiciel Cisco Catalyst et connexions de contrôle
- Protocole de gestion de superposition (OMP), localisateurs de transport (TLOC), détection de transfert bidirectionnel (BFD) et sécurité IP (IPsec)
- Réseaux de transport MPLS (Multiprotocol Label Switching)
- Mappage NAT et comportement du filtre
- Analyse de base des paquets avec Wireshark

Composants utilisés

Les informations contenues dans ce document sont basées sur les versions de matériel et de logiciel suivantes :

- Routeurs SD-WAN Cisco IOS® XE Catalyst
- Contrôleur SD-WAN Cisco Catalyst (vSmart)
- Cisco Catalyst SD-WAN Validator (vBond)

La défaillance de propriété de session de concentrateur a été reproduite sur Cisco IOS XE 17.18.2 avec des clés par paires opérationnelles. Les champs de disponibilité et de sortie des commandes varient selon la plate-forme et la version.

Les exemples utilisent des adresses privées ou des adresses de documentation réservées par la RFC 5737, avec des noms de périphériques génériques. Les extraits CLI sont raccourcis et utilisent des identificateurs normalisés. Les noms des interfaces et certains champs de sortie varient selon la plate-forme et la version.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Si votre réseau est en ligne, assurez-vous de bien comprendre l’incidence possible des commandes.

⚠ Mise en garde : Les modifications de configuration et les réinitialisations de connexion peuvent interrompre le trafic. Utiliser un plan de maintenance approuvé. Les commandes de diagnostic initiales sont en lecture seule.

Informations générales

Commencez par un scénario dans lequel la NAT symétrique est correctement identifiée. Spoke A s'enregistre avec vBond 1 et Spoke B s'enregistre avec vBond 2. La passerelle NAT peut attribuer la même adresse IP source traduite et le même port à ces connexions lorsque leurs destinations diffèrent, comme le permet la RFC 7857. Les deux rayons peuvent par conséquent annoncer le même point de terminaison public pour leurs TLOC Internet public, même lorsque les deux rapportent le type NAT A.

L'exemple de topologie utilise la couleur Internet public. Le même scénario peut impliquer d'autres couleurs publiques, telles que biz-internet. Appliquez les filtres de diagnostic à la couleur de transport réellement affectée ; la couleur internet public littérale n'est pas une condition requise.

✎ Remarque : Cisco Catalyst SD-WAN ne prend pas en charge les TLOC provenant de sites différents qui annoncent la même adresse IP publique et le même port UDP vers une périphérie réceptrice commune. Le comportement du récepteur dans cette condition de collision est imprévisible. Cela peut se produire même lorsque la NAT symétrique est correctement identifiée. Une classification NAT correcte seule ne garantit pas que les points de terminaison annoncés sont uniques.

Apprentissage NAT sur le plan de contrôle

Un périphérique envoie son adresse de transport et son port privés lors de l'enregistrement auprès de vBond. vBond observe l'adresse source et le port sur cette connexion et renvoie le point d'extrémité observé à l'extrémité. La périphérie peut annoncer ces informations via OMP en tant qu'adresse publique et attributs de port d'un TLOC, y compris un TLOC avec une couleur publique telle que public-internet ou biz-internet.

L'observation décrit le lien avec ce vBond. Avec la fonction NAT dépendante de la destination, le mappage vers un concentrateur peut être différent. Le routeur compare les observations NAT afin de classer le comportement de mappage. Dans show sdwan control local-properties, A signifie le mappage dépendant de l'adresse et du port, et E signifie le mappage indépendant du point d'extrémité. La classification affichée décrit la décision du routeur ; une table de traduction NAT ou une capture de paquets vérifie le comportement réel de la passerelle NAT.

Par exemple, exécutez show sdwan control local-properties et localisez la colonne NAT TYPE

pour le transport. Cette sortie raccourcie conserve la légende de type NAT et les colonnes d'interface choisies, avec des adresses normalisées :

<#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
```

Le transport biz-internet signale A et annonce 203.0.113.10:23012 comme son point d'extrémité public. Le transport MPLS signale E et possède des terminaux privés et publics égaux. Lisez la classification de NAT TYPE ; L'égalité d'adresse ou de port seule n'identifie pas le type NAT. N signifie que le routeur n’a pas appris de classification.

Apprentissage NAT du plan de données et VCONTROL

Avant qu'un concentrateur puisse traiter un paquet BFD chiffré, il doit trouver l'association de sécurité IPsec entrante pour l'homologue et le tunnel correspondants. L’association de sécurité identifie la clé de chiffrement et tout autre état requis pour authentifier et déchiffrer le paquet. Le récepteur utilise les informations de point d'extrémité avec l'index des paramètres de sécurité (SPI) ESP (Encapsulating Security Payload) afin de localiser cet état. Le SPI seul n'est pas un identificateur de rayon global unique.

Un rayon derrière la NAT symétrique initie normalement le trafic vers un concentrateur accessible. Cela crée un mappage pour cette destination particulière. L'adresse source ou le port vu par le concentrateur peut différer du point d'extrémité public précédemment appris via vBond et annoncé dans OMP. Par conséquent, les en-têtes externes reçus ne correspondent pas toujours au point d'extrémité dans l'état de tunnel installé. Le concentrateur a besoin d'un moyen d'associer ce nouveau point d'extrémité on-wire au TLOC distant connu et à son état IPsec avant de pouvoir décrypter BFD.

Pour le format IPv4 BFD-over-IPsec abordé ici, SD-WAN place les informations de point d'extrémité dans une queue de bande VCONTROL visible à la fin de la charge utile UDP, en dehors de la charge utile BFD chiffrée. Le concentrateur peut lire ces informations avant le déchiffrement. VCONTROL transporte les informations de point d'extrémité, pas les clés de

cryptage, et n'établit pas l'identité de l'homologue par lui-même.

Par exemple, lors de l'établissement initial du tunnel, avant la fin de l'apprentissage du plan de données :

```
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
```

Le récepteur tente d'abord sa recherche de point d'extrémité/SPI normale. Si cette recherche ne trouve pas d'association de sécurité, le point de terminaison source VCONTROL fournit un secours. Dans cet exemple, le concentrateur peut utiliser le 203.0.113.10:12346 transporté, avec son contexte de destination local et le SPI, afin de localiser la SA entrante associée au TLOC distant connu, même si le paquet arrive de 203.0.113.10:45678.

L’association de sécurité choisie fournit le contexte de la clé de réception. Le SPI choisit la génération de clé applicable dans ce contexte. Le concentrateur effectue ensuite l'authentification et le déchiffrement IPsec ; seul un paquet qui passe les vérifications requises est disponible pour le traitement BFD. Une queue de bande lisible ou une recherche réussie seule ne prouve pas une authentification, un déchiffrement ou une réception BFD réussis.

La différence entre les ports source observés et transportés peut donc être normale avec la NAT symétrique. Pour que la relation d'homologue fonctionne, le concentrateur doit associer le paquet avec le tunnel et l'état de clé corrects et apprendre le point de terminaison de retour réel. Dans cet exemple, l'apprentissage réussi permet au concentrateur d'envoyer le trafic de retour à 203.0.113.10:45678. Ceci est séparé de l'apprentissage du plan de contrôle via vBond.

VCONTROL ne force pas le récepteur à remplacer une correspondance de point d'extrémité/SPI existante. Si un état de point d'extrémité dupliqué associe un paquet à un homologue ou à une génération de clé incorrects, la queue de bande ne garantit pas la récupération. Il fournit un indice de recherche ; il ne peut pas distinguer deux homologues qui portent le même point d'extrémité sans le tunnel environnant et l'état cryptographique corrects.

VCONTROL est également présent sur les paquets BFD applicables lorsque NAT de type E est signalé. Sa présence n'est pas un indicateur NAT symétrique, et il n'est pas inclus dans chaque paquet de données IPsec ordinaire. Pour le format BFD IPv4 décrit ici, les informations de TLOC public sont omises entre deux couleurs privées.

Le point de terminaison spécifique à l'homologue appris via le plan de données ne signifie pas que le rayon doit remplacer son annonce OMP par un nouveau point de terminaison pour chaque destination. Différents mappages appris sont une partie attendue de la traversée NAT symétrique.

Une fois qu'un tunnel est actif, un expéditeur qui a appris son propre point de terminaison traduit pour ce tunnel peut transporter ce point de terminaison comme informations source. Il s'agit du point de terminaison public par tunnel de l'expéditeur, et non du point de terminaison distant ou d'une nouvelle annonce OMP globale. Comparez toujours le paquet, l'annonce OMP et l'état de session à partir de la même période d'observation.

Conservation des ports, chevauchement des ports et randomisation des ports

Voici les différents comportements :

Comportement	Signification
Conservation du port source	La passerelle NAT tente de conserver le port source UDP d'origine du rayon, par exemple 12346, comme port source traduit.
Chevauchement des ports entre les destinations	La passerelle NAT réutilise une adresse IP traduite : pour différents points d'extrémité internes qui communiquent avec différentes destinations externes.
Hiérarchisation des ports	La passerelle NAT choisit les ports traduits à l'aide d'une méthode d'allocation aléatoire.
Mappage indépendant des terminaux	Un point de terminaison interne conserve le même point de terminaison traduit sur les destinations externes tant que ce mappage existe.

RFC 7857, Section 3 autorise le chevauchement des ports externes pour différents terminaux internes lorsque les destinations diffèrent. La NAT distingue les flux à l'aide de leurs cinq-tuples. Cela n'autorise pas deux flux UDP simultanés et indiscernables vers la même destination. L'allocation et la conservation aléatoires sont des choix d'allocation distincts ; ni établit l'indépendance des terminaux.

Exemple de topologie

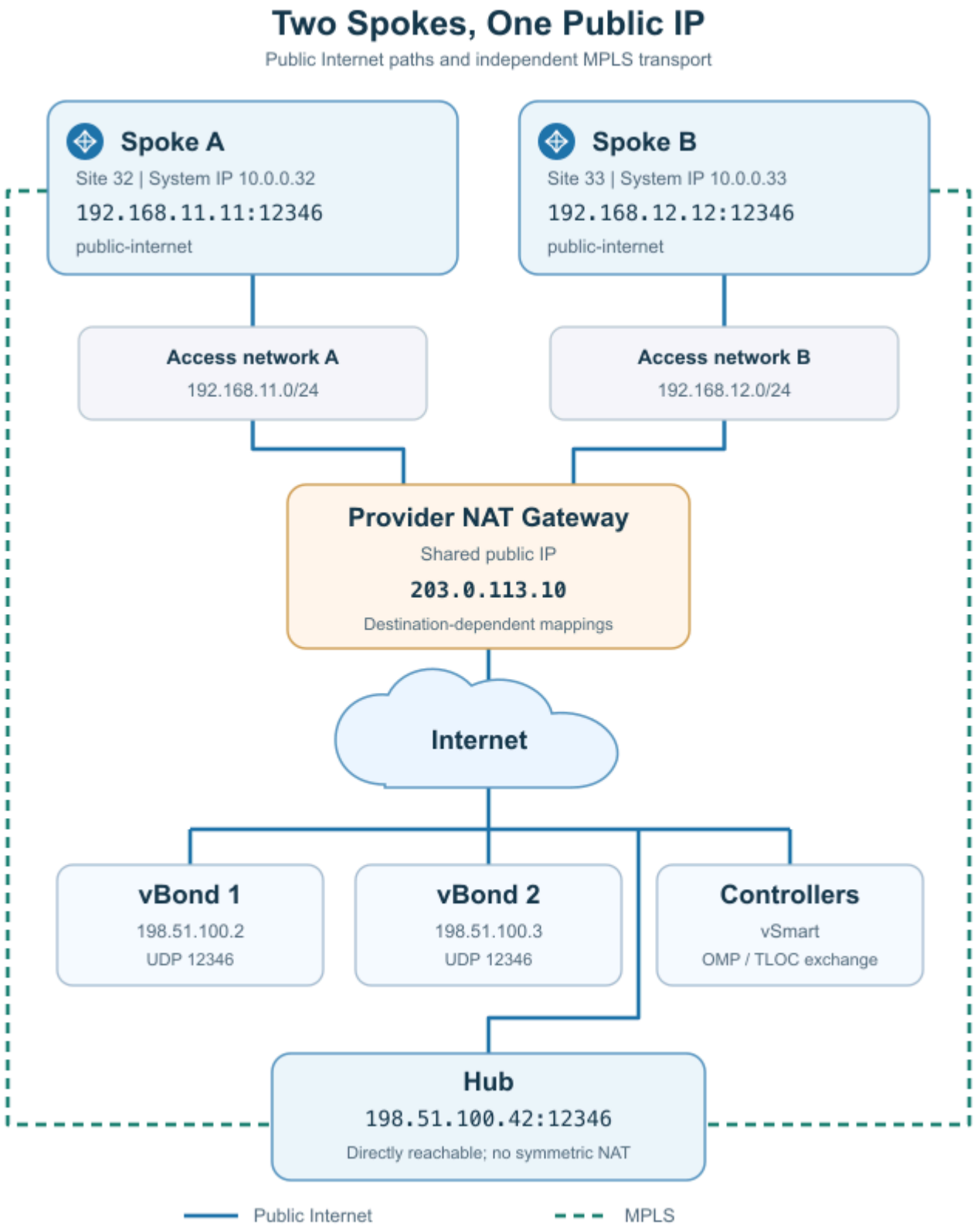

Figure 1. Passerelle NAT partagée, réseaux d’accès en étoile séparés et transport MPLS indépendant.

Le satellite A utilise 192.168.11.11:12346, et le satellite B utilise 192.168.12.12:12346. Tous deux partagent l'adresse publique 203.0.113.10 via la passerelle NAT. Le concentrateur a une adresse directement accessible et n’est pas derrière la fonction NAT symétrique. Les deux rayons peuvent

annoncer le même IP:port public même si leurs points d'extrémité privés diffèrent.

Pourquoi la collision se produit

Différents vBonds peuvent observer le même terminal public

Tenez compte des mappages UDP simultanés suivants :

Pair	Source privée	Source traduite	Destination
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

Les réponses sont renvoyées correctement, car la fonction NAT vérifie également le point de terminaison distant. Une réponse de vBond 1 appartient au mappage pour A ; une réponse de vBond 2 appartient au mappage pour B.

Les deux routeurs peuvent également contacter l'autre vBond et détecter correctement le mappage dépendant de la destination. Par exemple, gardez le tuple source privé pour chaque rayon inchangé et tenez compte de cet ordre d'observation :

Pair	première observation d'obligation acceptée	Comparaison avec l'autre obligation	Résultat
A	vBond 1 renvoie 203.0.113.10:12346	vBond 2 renvoie 203.0.113.10:40001	NAT A ; uplet public conservé 203.0.113.10:12346
B	vBond 2 renvoie 203.0.113.10:12346	vBond 1 renvoie 203.0.113.10:40002	NAT A ; uplet public conservé 203.0.113.10:12346

Les observations vBond diffèrent pour chaque routeur, la détection de la NAT A est donc correcte. Les premières observations conservées correspondent néanmoins entre les deux routeurs. La classification A n'attribue pas un port public unique à tous les rayons, et la visite des deux vBonds ne garantit pas cette unicité.

Chaque rayon reçoit l'adresse publique et le port observés par vBond dans la réponse d'enregistrement et utilise l'observation conservée pour remplir ses propriétés de transport local. Le rayon inclut ce point d'extrémité public, avec son point d'extrémité privé, dans les attributs TLOC qu'il annonce au contrôleur SD-WAN via OMP. Le contrôleur distribue ensuite le TLOC aux

autres arêtes conformément à la politique de contrôle applicable ; vBond n'est pas à l'origine de l'annonce OMP.

Si les deux rayons conservent le même terminal public observé, OMP peut donc contenir :

```
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
```

Ces TLOC logiques restent différents, car leurs adresses IP système diffèrent.

Les Mêmes Annonces Atteignent Un Concentrateur

La passerelle NAT doit distinguer les deux rayons lorsqu'ils entrent en contact avec le même concentrateur. Une allocation possible est la suivante :

Pair	Source traduite vers concentrateur	Destination
A	203.0.113.10:12346	198.51.100.42:12346
B	203.0.113.10:45678	198.51.100.42:12346

Le concentrateur a deux homologues logiques annonçant le même point de terminaison de destination, tandis que les paquets peuvent arriver de points de terminaison réels différents. Dans un fonctionnement normal de la NAT symétrique, l'apprentissage du plan de données permet d'établir les mappages de retour.

La limitation de la propriété de session du concentrateur se produit lorsque différents homologues partagent l'état de session choisi par cette clé de transport :

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

Si les deux annonces choisissent la même clé, une mise à jour d'un homologue ultérieur peut remplacer les informations de propriété dans l'état partagé. Un retrait, une suppression ou une recréation ultérieurs de la session pour un homologue peut perturber l'état IPsec/BFD pour l'autre homologue. L'apprentissage du plan de données ne garantit pas la récupération lorsque cet état est déjà associé à un homologue ou à une génération de flux incorrects.

La queue de bande visible n'est pas un identificateur d'homologue indépendant : les deux rayons peuvent transporter le même point d'extrémité source. Si le concentrateur choisit un état IPsec existant qui appartient au mauvais homologue ou à la mauvaise génération de clé, l'authentification et l'apprentissage peuvent échouer même si des paquets atteignent le concentrateur. Afin d'établir cette séquence pour un paquet particulier, le centre d'assistance technique de Cisco (TAC) doit mettre en corrélation l'association de sécurité choisie avec son homologue et sa génération. La queue de bande du paquet seule ne peut pas le prouver.

La même adresse IP publique seule n’établit pas de collision. Différents ports distants produisent généralement des clés de transport différentes. Vérifiez également l'adresse locale du concentrateur, le port local et le protocole choisis. Si un troisième homologue annonce le même terminal, il peut également rivaliser pour cette clé ; le modèle d'échec dépend de l'ordre des événements.

Collisions « privé-public » avec adresses de filiale répétées

L'ID de bogue Cisco CSCwx12986 peut amener un périphérique à annoncer un tuple RFC 1918 privé dans ses attributs TLOC publics après un saut de port source automatique pendant la comparaison NAT. Si différentes branches réutilisent la même adresse privée et le même port, ces attributs publics incorrects peuvent correspondre entre les homologues et exposer le problème de propriété de session du concentrateur.

Par exemple, les deux rayons peuvent utiliser 192.168.11.11 derrière des pare-feu de branche séparés. Chaque pare-feu garde son réseau de filiale distinct et traduit l'adresse source. Si une passerelle NAT du fournisseur en amont la traduit à nouveau, le chemin comporte deux étapes NAT, ou NAT double. Des contextes de routage et de transfert virtuels (VRF) séparés constituent un autre moyen d'isoler les adresses de succursale qui se chevauchent.

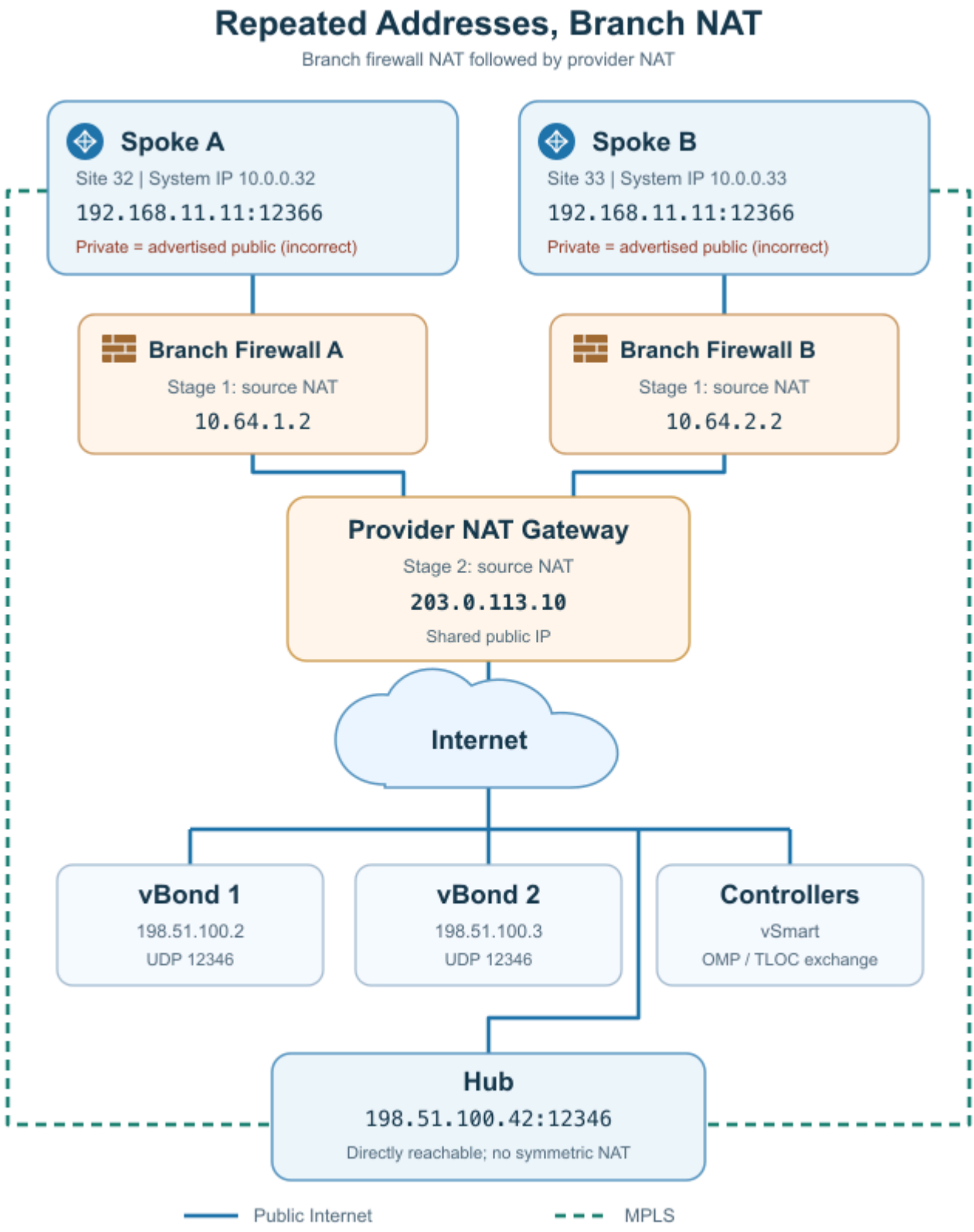

Figure 2. Traductions NAT du pare-feu de filiale et du fournisseur Les attributs TLOC publics incorrects conservent le point de terminaison privé d'origine, tandis que les paquets utilisent les adresses source traduites.

Si les deux rayons rencontrent le défaut sur leurs TLOC Internet publics et utilisent le même port source post-saut, le concentrateur peut recevoir ces annonces :

Pair	Point de terminaison privé après saut de port	Point de terminaison TLOC public incorrect
A	192.168.11.11:12366	192.168.11.11:12366
B	192.168.11.11:12366	192.168.11.11:12366

Remarque : Les adresses privées répétées et la double NAT ne sont pas des conditions préalables pour l'ID de bogue Cisco CSCwx12986 lui-même. La collision inter-homologues supplémentaire nécessite la correspondance des valeurs IP:port annoncées et les mêmes adresse, port et protocole concentrateur-local. La même adresse IP privée avec des ports annoncés différents n'établit pas cette collision.

L'égalité des adresses publiques et privées est normale sur un transport sans NAT. Une adresse RFC 1918 dans un champ public n'est suspecte que lorsqu'elle contredit le chemin prévu et la traduction observée. Les adresses privées de branche uniques évitent le chevauchement répétée-private-tuple dans cet exemple ; elles n'empêchent pas le scénario séparé de réutilisation de port public traduit décrit précédemment.

Réaffectation des terminaux NAT

Un point de terminaison peut également être réaffecté à un autre satellite après l'expiration d'un mappage NAT alors que le concentrateur conserve l'ancien état homologue. Comparez l'historique de mappage avec le cycle de vie de la session du concentrateur afin d'identifier le chevauchement entre les anciens et les nouveaux propriétaires.

Symptômes et impact

Les indicateurs types sont les suivants :

- Les connexions de contrôle restent actives tant que les chemins BFD affectés sont désactivés ou ne parviennent pas à récupérer.
- Deux adresses IP système différentes annoncent la même adresse IP publique : port.
- Le concentrateur conserve les deux TLOC logiques, mais il manque un homologue dans une vue de session de plate-forme filtrée ou affiche des informations système, IP, couleur distante, discriminateur, tunnel ou SA incohérentes.
- Un retrait de transport ou une récréation pour un homologue perturbe un autre homologue avec le même terminal.
- Les pertes d'entrée IPsec peuvent augmenter sur le chemin affecté.

La défaillance peut rendre le chemin de transport affecté inutilisable. S'il s'agit du seul chemin de

transport utilisable pour un rayon, le rayon est isolé de la superposition. Sur un grand concentrateur, de nombreux rayons partageant une adresse NAT augmentent la population nécessitant une corrélation, mais le nombre de chemins affectés doit être mesuré ; elle ne peut pas être déduite du nombre de sites derrière cette adresse.

Diagnostiquer le problème

Étape 1 : enregistrement de l'état de transport du logiciel et du satellite

Sur chaque périphérie affectée de Cisco IOS XE, remplacez **<**involved-color**>** par la couleur de transport réelle, par exemple 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

Notez l'adresse et le port privés/publics, l'interface, la couleur, le type NAT, l'ID de site et l'état du contrôle. Après le premier exemple complet de propriétés locales, utilisez ce filtre compact pour conserver les en-têtes d'interface et la ligne de transport choisie. Remplacez TenGigabitEthernet0/1/0.10 par l'interface concernée :

<#root>

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

Pour le scénario principal, les champs appropriés peuvent être lus sur les deux rayons. Il s'agit d'un résumé de champ et non de la disposition de colonne native :

```
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
```

Une connexion de contrôle Up vérifie cette relation de contrôleur. Il ne vérifie pas le chemin BFD entre les rayons et le concentrateur ni l'unicité du point d'extrémité public.

Étape 2. Recherche des terminaux en double dans OMP

Sur le concentrateur Cisco IOS XE :

<#root>

show clock detail

show sdwan omp tlocs table | include 203.0.113.10

Sur le contrôleur SD-WAN (vSmart) :

<#root>

show clock

show omp tlocs | tab | include 203.0.113.10

Cette sortie du contrôleur utilise des identificateurs normalisés. Les colonnes IPv6 et de région sont omises :

<#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
```

Comparez l'adresse IP, la couleur, l'encapsulation, l'adresse IP publique et le port public du système TLOC. Plusieurs chemins vers le même TLOC logique à partir de différents contrôleurs sont normaux. Le candidat de collision est deux identités d'homologue différentes avec le même point de terminaison effectif.

Répétez la commande filtrée après 30 à 60 secondes et conservez les deux horodatages. Le chevauchement persistant est une preuve utile ; des commandes parfaitement simultanées ne sont pas requises. Pour une courte période transitoire, collectez les événements OMP et les enregistrements de cycle de vie de session correspondants avec le TAC.

Pour deux homologues suspects, conservez leurs entrées de table TLOC individuelles :

<#root>

show sdwan omp tlocs table | include 10.0.0.32

show sdwan omp tlocs table | include 10.0.0.33

Pour obtenir une vue compacte des détails du TLOC IPv4 sur Cisco IOS XE, utilisez ce filtre. Remplacez l'adresse IP du système par l'homologue étudié et conservez les guillemets autour du sélecteur 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|::

Comparez les attributs d'adresse/port public et privé pour les deux homologues. Le filtre omet les lignes non définies, de valeur zéro, par défaut et IPv6 correspondantes dans cette vue.

Étape 3. Comparaison des TLOC logiques avec la propriété de session Hub

Inspectez l'état TLOC fourni au FTM. Comparez les adresses IP système, les couleurs, les adresses IP publiques et les ports UDP publics des deux homologues.

<#root>

show platform software sdwan ttmd tlocs

Exécutez ces commandes afin de comparer les enregistrements de session FTM (Forwarding Table Manager), BFD et de plate-forme :

<#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

Cet exemple montre le modèle avec des identificateurs normalisés. Des étiquettes de colonne sont ajoutées aux lignes FTM filtrées pour plus de clarté; ils ne font pas partie du résultat de la commande filtrée :

<#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 et FTM conservent les deux identités d'homologue logique, tandis que cette recherche de session affiche uniquement Spoke A pour la destination partagée. Spoke A a également un état Down FTM ; la présence dans la table de session de la plate-forme n'indique pas un chemin sain. L'index TLOC FTM est distinct du discriminateur local BFD (BFD-LD) dans la table de session. Ce schéma est une raison d'étudier la propriété entre les couches. Une ligne CLI manquante seule n'établit pas la perte de paquets.

Dans ce tableau, Color représente le transport local ; utilisez la couleur Remote dans la vue détaillée afin d'identifier le TLOC distant. Comparez également DstIp et DPubIp : une destination réelle apprise peut différer du point de terminaison public associé à la session. Utilisez la vue spécifique à l'homologue si un filtre de destination uniquement ne correspond plus à l'adresse apprise.

Un saut suivant non valide, un identificateur SA de 0 ou Dans le matériel : FALSE fournit des preuves supplémentaires plutôt qu'une signature unique. Ces champs peuvent également se produire lors d'une défaillance ordinaire de la configuration ou de la sous-couche.

Étape 4. Vérification de la fonction réelle de la passerelle NAT

Demandez à l'opérateur NAT ou du fournisseur de services des enregistrements de traduction UDP simultanés avec ces champs :

```
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
```

Ceci est un tableau de traduction illustratif. Utilisez la commande de diagnostic équivalente pour votre plateforme NAT. Sur une passerelle NAT Cisco IOS XE, commencez par :

<#root>

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

Conservez les colonnes source et de destination internes. Une liste contenant uniquement la source traduite ne peut pas montrer pourquoi la réutilisation est valide vers différentes vBonds ou quel rayon possède le mappage de concentrateur.

Capturez les flux UDP affectés avant et après NAT pour les vBonds et le concentrateur. Conserver la longueur totale des paquets afin que la queue de bande VCONTROL ne soit pas tronquée. Utiliser une collection SPAN (Switched Port Analyzer) ou EPC (Embedded Packet Capture) limitée, en accord avec le TAC. Comparer le même flux entre les points de capture ; ne déduisez pas la latence des horloges de capture non synchronisées.

Étape 5. Inspection de VCONTROL dans Wireshark

Utilisez le décodeur Lua de l'annexe pour le format IPv4 BFD/VCONTROL décrit ici. Enregistrez le code sous le nom sdwan_vctrl.lua dans le même dossier que votre capture. Ce décodeur de diagnostic affiche des champs visibles ; il ne déchiffre pas ESP et ne vérifie pas l'authentification. Sa portée est complète, paquets IPv4/UDP non imbriqués avec l'indicateur BFD/VCONTROL attendu et la structure de fin.

Exécutez la commande Wireshark à partir du dossier qui contient le décodeur et votre capture :

<#root>

```
wireshark -n \

  -X lua_script:sdwan_vctrl.lua \

  -d udp.port==12346,sdwan_vctrl \

  -d udp.port==45678,sdwan_vctrl \

  -r hub.pcap
```

Remplacez les exemples de ports par ceux de la capture. Sous Windows ou macOS, utilisez le chemin d'accès complet à l'exécutable Wireshark s'il ne figure pas dans PATH.

Vous pouvez également charger le script Lua au démarrage de Wireshark, choisir un paquet UDP approprié et accéder à Analyze > Decode As. Attribuez les entrées de port UDP source et de destination applicables à SDWAN_VCTRL / SD-WAN VCTRL (IPv4). Une entrée UDPENCAP explicite sur l'autre port peut être prioritaire sur ce décodeur. Appliquez la modification à cette

session d'analyse.

Appliquer ce filtre d'affichage :

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

Développez SD-WAN VCTRL > Visible VCONTROL trailer et comparez :

```
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
```

Ces valeurs illustratives décrivent l'établissement initial du tunnel avant la fin de l'apprentissage du plan de données. Ils montrent le mappage de concentrateur réel pour B à côté de son point d'extrémité transporté. NAT modifie les en-têtes externes ; il ne réécrit pas les informations de point de terminaison dans cette queue de bande. Cette non-concordance ne démontre pas à elle seule une défaillance.

Exécutez la commande TShark pour une exportation compacte :

<#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
```

Sans décodeur personnalisé, le format IPv4 décrit ici a ce code de fin TLV (type-length-value) de 17 octets à la fin de la charge utile UDP. Le remplissage de la couche liaison se trouve en dehors de la charge utile 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
```

N'appliquez pas cette disposition fixe à IPv6, à d'autres formats de contrôle, à des paquets tronqués ou à des données chiffrées arbitraires. Un filtre décodeur vide peut signifier que le paquet se trouve en dehors de son format pris en charge. Un point d'extrémité source lisible n'est pas non plus une étiquette d'IP système authentifiée.

Remarque : Les avertissements de séquence ESP génériques dans Wireshark ne sont pas des diagnostics anti-relecture de routeur. Corrélez le paquet avec les compteurs et les journaux du routeur avant de tirer une conclusion sur l’authentification ou la propriété.

Étape 6. Confirmation du chemin affecté et conservation des preuves

Sur le concentrateur, prélevez deux échantillons de compteur QuantumFlow Processor (QFP) autour de la défaillance observée :

<#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
```

Utilisez les différences de compteur plutôt que les totaux absolus. Les abandons IPsec agrégés peuvent avoir d'autres causes. Confirmez l'arrivée du paquet sur le transport affecté et établissez une corrélation avec l'homologue, le point d'extrémité, le discriminateur BFD, le tunnel, l'association de sécurité et l'heure de l'événement.

Fournir au TAC les deux instantanés OMP, les propriétés et les contrôles locaux en étoile, les détails de la session du concentrateur pour les deux homologues, l'état BFD, les enregistrements de traduction, les captures limitées et les versions logicielles. S'il est déjà disponible, inclure les admin-techs et l'historique de connexion/retrait appropriés. Évitez la collecte répétée de tables entières sur un grand concentrateur.

Testez l'accessibilité du service séparément et enregistrez le transport réellement choisi. L'indication d'une augmentation BFD de zéro seconde n'est pas une récupération soutenue.

Solutions de contournement et limitations

Évaluez ces options d'atténuation avec le TAC Cisco et l'opérateur NAT. La limitation concerne les points d'extrémité annoncés en double atteignant un concentrateur commun ; cela ne signifie pas que la NAT symétrique, les adresses publiques partagées ou l'adressage privé répété ne sont généralement pas pris en charge. Les procédures de rétablissement et l'efficacité des mesures d'atténuation dépendent du déploiement.

Utiliser des mappages stables indépendants des terminaux avec des allocations uniques

Une option de prévention côté NAT pour la séquence illustrée consiste à donner à chaque point d'extrémité de transport satellite distinct un point d'extrémité UDP traduit unique et stable sur les destinations vBond, contrôleur et concentrateur. Demandez l'indépendance et l'unicité des terminaux sur les rayons. Vérifiez le résultat dans les enregistrements et les captures de traduction, également après le basculement de la passerelle NAT ou la recréation de l'état.

Par exemple, A peut utiliser systématiquement 203.0.113.10:40001, tandis que B utilise systématiquement 203.0.113.10:40002. Des traductions statiques ou une stratégie d'allocation appropriée peuvent fournir ce comportement, sous réserve de la plate-forme NAT et de la conception d'accès.

Le mappage indépendant du point de terminaison ne nécessite pas un accès entrant illimité. Conservez les contrôles de pare-feu requis. Une correction de stratégie NAT ne supprime pas automatiquement l'état de session précédemment annoncé ou programmé ; coordonner un nouveau cycle de découverte et de rétablissement avec le TAC.

Désactiver le chevauchement de ports sur différents rayons

Lorsque cela est pris en charge, réservez les terminaux UDP traduits uniquement sur les rayons SD-WAN appropriés. Ne les réutilisez pas uniquement parce que leurs destinations sont différentes. Cela supprime directement le mappage vBond en double simultané indiqué précédemment, même si la passerelle NAT conserve un comportement dépendant de la destination.

Le paramètre peut consommer plus de ports ou nécessiter des adresses publiques supplémentaires. Il ne rend pas le mappage pour chaque point d'extrémité satellite indépendant et ne résout pas l'état de récepteur obsolète après une réaffectation ultérieure du point d'extrémité.

Avec plusieurs étapes NAT ou des passerelles NAT redondantes, vérifiez l'allocation visible à la limite Internet finale. L'unicité au sein d'une instance NAT n'établit pas l'unicité entre d'autres instances qui partagent la même adresse publique.

Utiliser le décalage de port pour un petit groupe contrôlé

La commande port-offset fournit les valeurs de 0 à 19. Par exemple, Spoke A peut conserver le décalage 0 et Spoke B peut utiliser le décalage 1. Exécutez ces commandes de configuration sur 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
```

Vérifiez la source active et les ports traduits après la modification. Une modification de ce paramètre système peut recréer le contrôle et les connexions de données sur les transports. Il ne

s'agit pas d'une étape de diagnostic sans interruption.

Les 20 décalages fournissent 20 groupes de ports de base distincts avec le décalage 0 comme choix. Il ne s’agit pas d’une limite générale de 20 routeurs SD-WAN derrière une adresse IP publique. Cette solution de contournement particulière ne propose que 20 choix de décalage uniques et nécessite que la passerelle NAT les préserve et les distingue. Il échoue en tant que schéma d'unicité lorsque les décalages sont réutilisés. Le saut de port n'ajoute pas d'autres identités de décalage.

Par exemple, 124 routeurs ne peuvent pas tous recevoir des décalages uniques derrière une adresse partagée. Un schéma de décalage uniquement avec au plus 20 routeurs par groupe d'adresses publiques nécessite au moins sept groupes. Cette arithmétique n'est pas une recommandation de capacité ou un substitut à la vérification du comportement NAT. Une passerelle NAT qui réécrit ou chevauche des ports traduits peut annuler le schéma même avec moins de routeurs.

Désactiver la conservation du port source ou activer l'allocation aléatoire

Une politique sans préférence de conservation peut réduire la tendance de nombreux rayons à recevoir :12346 vers différentes vBonds. L'allocation aléatoire peut distribuer ces mappages sur un pool de ports plus important.

Cependant, aucun changement ne garantit à lui seul la prévention des collisions. Un allocateur aléatoire peut toujours réutiliser un port entre plusieurs destinations, modifier un mappage vers le concentrateur ou réaffecter un point d'extrémité expiré alors que l'ancien état du concentrateur reste inchangé. Vérifiez si l'allocation est unique entre les abonnés et les destinations. Les ports aléatoires seuls n'établissent pas l'unicité.

Examinez les paramètres de conservation et d'allocation du port source sur la passerelle NAT avec son opérateur. La configuration applicable dépend de la plate-forme NAT ; vérifiez les mappages résultants vers vBonds et le concentrateur avec les enregistrements de traduction et les captures de paquets.

Les ports traduits peuvent se trouver en dehors de la plage de ports de base SD-WAN d'origine. Examiner les règles de pare-feu afin qu'elles ne rejettent pas le trafic traduit légitime uniquement parce que le port source a changé.

Adresses publiques distinctes ou blocs de ports de réserve

Des adresses traduites dédiées, ou des blocs de ports sans chevauchement réservés par rayon, peuvent fournir des points d'extrémité annoncés distincts pour différents rayons actifs. La politique

doit être appliquée de manière cohérente sur les chemins de découverte et de concentrateur et sur les noeuds NAT redondants. Un espace d'adressage supplémentaire et une prise en charge du fournisseur peuvent être nécessaires. La réutilisation dynamique après le retrait du mappage nécessite toujours une gestion correcte du cycle de vie de la session du concentrateur.

Conserver un autre transport et coordonner la récupération

Un MPLS vérifié ou un autre chemin admissible peut préserver le service pendant que le chemin Internet affecté est examiné. Une liste de contrôle d'accès sous-jacente qui bloque l'adresse NAT partagée fournit uniquement la contention : il peut supprimer la redondance Internet pour chaque rayon derrière cette adresse et ne retire pas automatiquement leurs annonces OMP.

Un contrôle clear, une réinitialisation d'interface, un NAT-state clear ou un rechargement de concentrateur peuvent recréer temporairement l'état. Aucun ne garantit la prévention de la récurrence. N'effectuez pas d'opérations de nettoyage à grande échelle comme première étape de diagnostic. Conservez les preuves, corrigez les conditions d'allocation si possible et convenez de la plus petite mesure de rétablissement avec le TAC. Validez le trafic BFD up et bidirectionnel soutenu sur le chemin restauré.

Défauts logiciels associés

ID de bogue Cisco	Champ d'application pertinent au présent article
CSCww45867	Documentation de la limitation de propriété de session de concentrateur lorsque des homologues différents annoncent le même point de terminaison TLOC public. Cet enregistrement effectue le suivi de la mise à jour de la documentation, plutôt que d'une correction logicielle.
CSCwx12986	Une séquence d'apprentissage NAT/saut de port automatique peut laisser un tuple privé dans les champs TLOC publics. Si les filiales réutilisent le même IP:port privé, les points de terminaison publics incorrects peuvent correspondre entre les homologues. La correction de ce défaut du producteur n'empêche pas une réutilisation valide du port public traduit.
CSCwx48929	La classification NAT peut devenir E à tort lorsque la détection revient au même vBond. Une classification correcte n'empêche pas le scénario de réutilisation de port différent-vBond valide.

Les corrections pour l'ID de bogue Cisco CSCwx12986 et l'ID de bogue Cisco CSCwx48929 ne garantissent pas des points d'extrémité publics annoncés uniques et ne permettent pas, par elles-

mêmes, d'éliminer ce scénario de collision. La fermeture de l'enregistrement de documentation de l'ID de bogue Cisco CSCww45867 n'établit pas de changement dans le comportement du logiciel. Confirmez les limitations et les mesures d'atténuation applicables avec le TAC Cisco.

Annexe : Décodage de la queue de bande IPv4 VCONTROL

Enregistrez ce code sous le nom sdwan_vctrl.lua. Le décodeur de diagnostic n'authentifie pas les paquets et ne déchiffre pas leurs données utiles. Il utilise le filtre d'affichage sdwan_vctrl et enregistre un choix Decode As explicite. La capture de paquets reste inchangée.

```
-- 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)
```

Informations connexes

- RFC 7857, Section 3 : Comportement de chevauchement de ports
- RFC 4787 : Exigences comportementales NAT pour le protocole UDP monodiffusion
- RFC 5737 : Blocs d'adresses IPv4 réservés à la documentation
- Guide de conception SD-WAN de Cisco Catalyst
- Commandes système SD-WAN de Cisco IOS XE Catalyst : port-offset
- Configuration du décalage du port SD-WAN Cisco
- Référence de la ligne de commande Wireshark

- Référence de la ligne de commande TShark
- Assistance et documentation techniques - Cisco Systems

À propos de cette traduction
Cisco a traduit ce document en traduction automatisée vérifiée par une personne dans le cadre d'un service mondial permettant à nos utilisateurs d'obtenir le contenu d'assistance dans leur propre langue.

Il convient cependant de noter que même la meilleure traduction automatisée ne sera pas aussi précise que celle fournie par un traducteur professionnel.