Dépannage de la classification de sous-réseau L3Out dans ACI

Table des matières

Introduction

Ce document décrit la classification des sous-réseaux externes dans les groupes de terminaux L3Out de l'ACI Cisco.

Abréviations

- BD : Domaine Bridge
- EPG : Groupe de terminaux
- ExEPG : Groupe de terminaux externes
- RIB : Base d'informations de routage
- VRF : Routage et transfert virtuels
- ID de classe : Balise qui identifie un EPG

Classification EPG externe

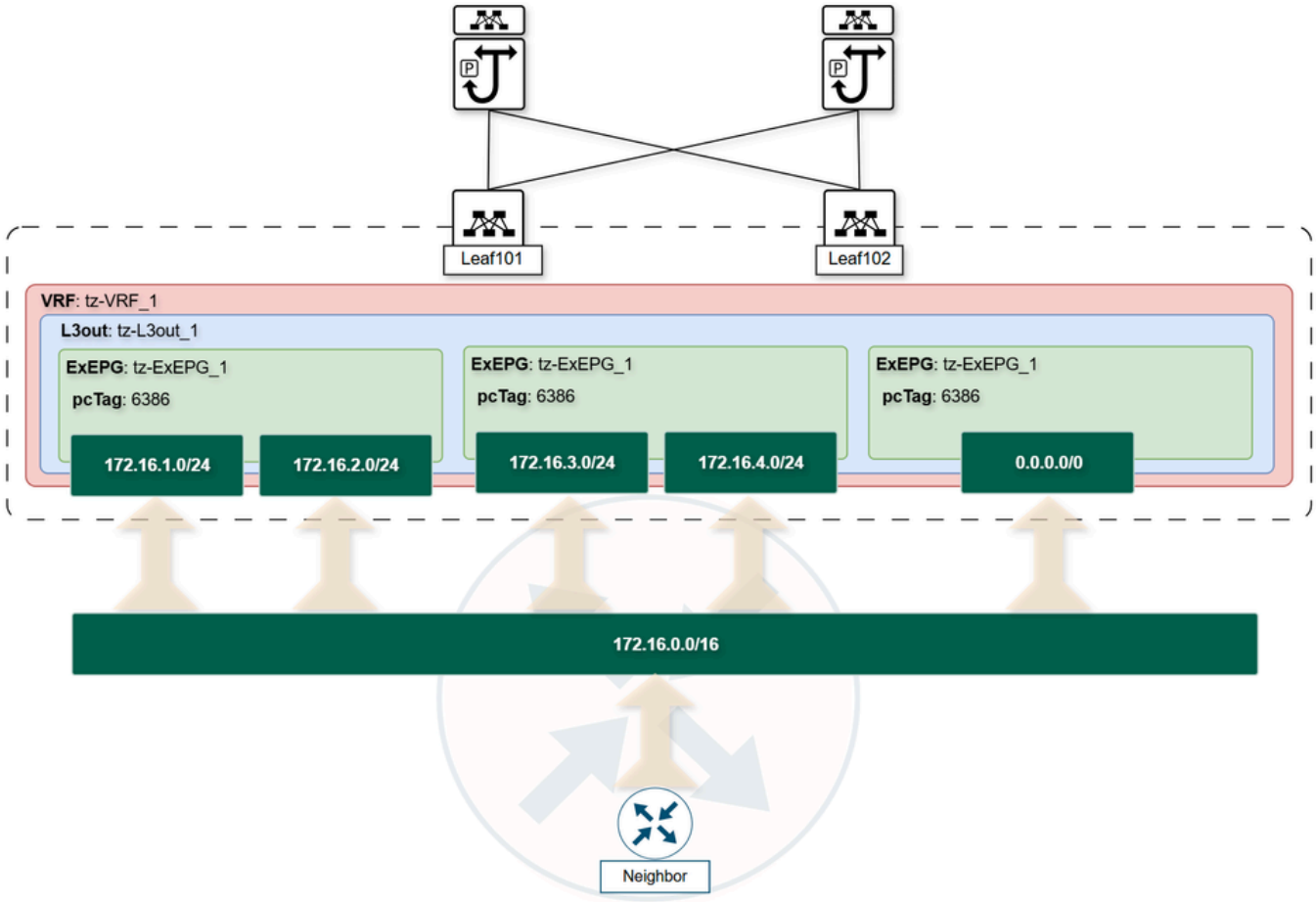
Un EPG externe dans l'ACI Cisco représente des réseaux routés externes connectés via des sorties L3. De la même manière qu'un EPG classique classe les terminaux, un EPG externe classe les sous-réseaux externes par VRF, ce qui signifie que chaque sous-réseau doit être unique dans son contexte VRF.

Une idée fausse courante est qu'un sous-réseau EPG externe inclut uniquement des préfixes acceptés via le protocole de routage dynamique. Cependant, lorsqu'une L3Out est créée, il y a

une route-map par défaut filtrant les annonces entrantes ; ainsi, tous les préfixes annoncés par le protocole de routage dynamique sont acceptés par défaut. L'objectif principal de la définition des sous-réseaux dans un ExEPG est la classification uniquement pour attribuer un pcTag unique aux sous-réseaux inclus dans l'ExEPG pour l'application des contrats et des politiques.

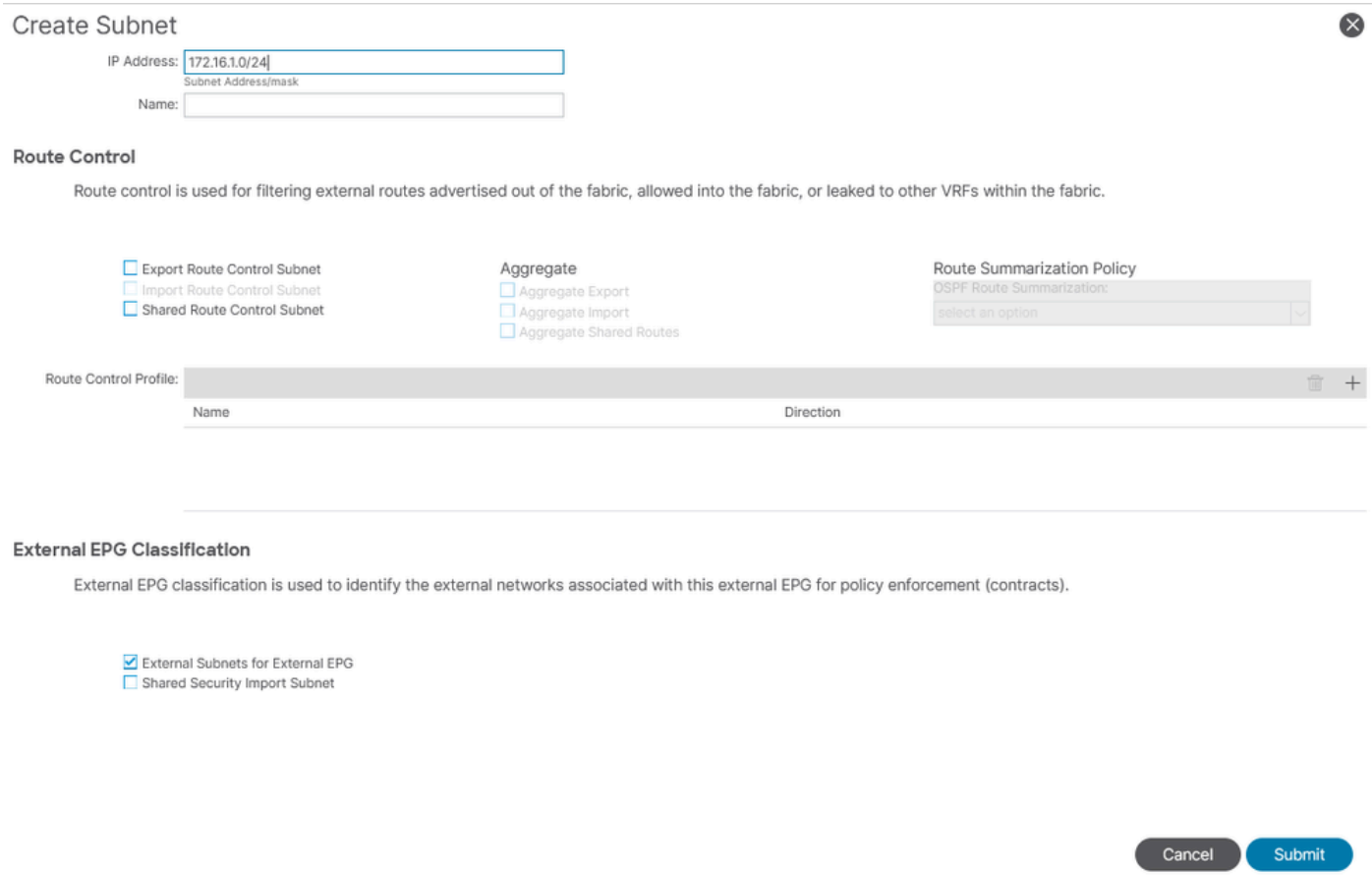
Cette classification permet un contrôle granulaire des politiques. Par exemple, un seul voisin externe peut annoncer un super-réseau à l'ACI, qui peut ensuite être segmenté en plusieurs groupes ExEPG. Cela permet d'appliquer différentes actions de contrat à des sous-réseaux distincts, comme autoriser des EPG internes spécifiques à communiquer uniquement avec des sous-réseaux externes désignés ou rediriger le trafic destiné à certains préfixes vers un noeud PBR avant d'atteindre sa destination finale.

Ce schéma illustre la façon dont l'ACI Cisco classe les sous-réseaux externes en fonction des EPG externes, ce qui permet une segmentation précise du trafic et l'application des contrats.



Indicateurs de sous-réseaux EPG externes

Pour classer et gérer les préfixes externes dans un ExEPG dans l'ACI, des indicateurs de sous-réseau spécifiques sont configurés lors de la création d'un préfixe de sous-réseau sous un ExEPG. Cette section détaille chaque indicateur et son utilisation prévue :



- Sous-réseau externe pour EPG externe :
  Cet indicateur indique que le sous-réseau se trouve en dehors du fabric ACI et n'est configuré dans aucun domaine de pont ou EPG. Il ne doit être utilisé que lorsque le préfixe est annoncé par un voisin de routage ou injecté statiquement dans le RIB. Cet indicateur est activé par défaut.

- Exporter le sous-réseau de contrôle de route :
  Cet indicateur indique que le sous-réseau est annoncé de l'ACI au voisin de routage via le protocole de routage dynamique. Il ne doit pas être activé simultanément avec l'indicateur Sous-réseau externe pour EPG externe, car cela peut provoquer des boucles de routage de couche 3. Puisque l'ACI classe le sous-réseau comme externe et l'annonce en retour, cela peut entraîner des incohérences de routage malgré les mécanismes d'évitement de boucle dans les protocoles de routage.

- Sous-réseau de contrôle de route partagée :
  Cet indicateur est défini lorsque le préfixe de sous-réseau est destiné à être partagé entre plusieurs VRF, ce qui permet la fuite de route entre les contextes.

- Sous-réseau d'importation de sécurité partagée :
  Utilisée conjointement avec l'indicateur de sous-réseau de contrôle de route partagée, cette option permet le partage des pcTags de sécurité pour les sous-réseaux externes sur différents VRF, ce qui facilite l'application cohérente des politiques.

- Importer le sous-réseau de contrôle de route :
  Cet indicateur permet un contrôle granulaire sur les préfixes reçus des voisins de routage.

Par défaut, l'ACI accepte toutes les annonces de routes entrantes ; l'activation de cet indicateur nécessite l'activation de l'application du contrôle de route pour filtrer les préfixes entrants.

- Section agrégée :
  Applicable uniquement au sous-réseau quad-0 (0.0.0.0/0), cette section récapitule tous les préfixes du RIB pour l'exportation ou l'importation d'agrégats. Lorsque des sous-réseaux sont transmis à d'autres VRF, ils sont résumés en tant que routes partagées agrégées pour optimiser les tables de routage.

Commandes de vérification et de dépannage

Routage

Pour commencer, la route doit être présente dans la table de routage du VRF sur les commutateurs de périphérie. Par exemple, cette commande montre une route BGP dans le VRF tz : tz-VRF_1 :

```
<#root>

Leaf101#

show ip route bgp vrf tz:tz-VRF_1


IP Route Table for VRF "tz:tz-VRF_1"


'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>


172.16.1.0/24

, ubest/mbest: 1/0

*via 10.10.1.2

%tz:tz-VRF_1, [20/0], 00:00:04, bgp-65002, external, tag 65003
Leaf101#
```

Cela confirme que la route est installée dans la table de routage VRF et est disponible pour les décisions de transfert.

Classification

Une fois que la route est présente dans la table de routage, la classification détermine la manière dont le trafic est géré en fonction de la stratégie. Dans l'ACI, la classification est liée à l'ExEPG et à ses sous-réseaux associés.

Pour valider la classification de sous-réseau sous un ExEPG, l'APIC peut être interrogé pour la classe l3extInstP, qui représente l'instance EPG externe. Sa classe enfant l3extSubnet répertorie les sous-réseaux configurés sous cet ExEPG. Exemple :

<#root>

```
moquery -c l3extInstP -f 'l3ext.InstP.dn*"[ tenant name ].*[ l3out name ]"' -x rsp-subtree=children rsp-
```

<#root>

```
APIC#

moquery -c l3extInstP -f 'l3ext.InstP.dn*"tz.*l3out"' -x rsp-subtree=children rsp-subtree-class=l3extSub


Total Objects shown: 1

# l3ext.InstP

name : tz-ExEPG_1


!-- cut for brevity --!
configSt : applied
descr :
dn : uni/tn-tz/out-l3out/instP-tz-ExEPG_1
!-- cut for brevity --!
floodOnEncap : disabled
isSharedSrvMsiteEPg : no
lcOwn : local
matchT : AtleastOne
mcast : no
modTs : 2025-09-10T00:36:49.239+00:00
monPolDn : uni/tn-common/monepg-default
nameAlias :
pcEnfPref : unenforced

pcTag : 32771


pcTagAllocSrc : idmanager
prefGrMemb : exclude
prio : unspecified
rn : instP-tz-ExEPG_1
scope : 3047430
status : modified
targetDscp : unspecified
triggerSt : triggerable
txId : 11529215046123188 28
uid : 15374
userdom : :all:

# l3ext.Subnet

ip : 172.16.1.0/24


!-- cut for brevity --!
dn : uni/tn-tz/out-l3out/instP-tz-ExEPG_1/extsubnet-[172.16.1.0/24]
```

```
extMngdBy :
lcOwn : local
modTs : 2025-09-10T01:05:13.249+00:00
monPolDn : uni/tn-common/monepg-default
!-- cut for brevity --!
rn : extsubnet-[172.16.1.0/24]

scope : import-security


status :
uid : 15374
userdom : :all:

APIC#
```

Si aucun résultat n'est renvoyé pour la classe l3extSubnet, cela indique qu'aucun sous-réseau n'est configuré sous l'EPG externe. Sans sous-réseaux configurés, l'ACI ne peut pas associer un pcTag au sous-réseau du trafic entrant, ce qui entraîne l'abandon du trafic malgré la route existant dans la table de routage.

Un autre aspect important à noter est la portée du sous-réseau, qui représente les indicateurs définis pour le sous-réseau en question :

- Import-security

Le sous-réseau a été marqué avec Sous-réseau externe pour EPG externe.

- export-rtctrl

Le sous-réseau a été marqué avec le contrôle de routage d'exportation.

- import-rtctrl

Le sous-réseau a été marqué avec le contrôle de routage d'importation.

- sécurité partagée

Le sous-réseau a été marqué avec le sous-réseau d'importation de sécurité partagée.

- shared-rtctrl

Le sous-réseau a été marqué avec le contrôle de routage partagé.

Les protocoles de routage et les processus de plan de contrôle mettent à jour les tables de routage lors de la réception d'un préfixe provenant d'un voisin mentionné, qui sont ensuite programmés dans les tables de transfert HAL L3. Les routes de couche 3 HAL représentent les routes de couche 3 réelles programmées dans les tables de transfert matériel (ASIC) sur les commutateurs Leaf. Ces routes sont dérivées des protocoles de routage et des calculs de la table de routage et sont utilisées pour les décisions de transmission.

<#root>

```
<-- When the prefix is not configured under the External EPG, a classification of 0xf is seen -->
Leaf101#
vsh_lc -c 'show platform internal hal l3 routes vrf tz:tz-VRF_1' | egrep "Prefix/Len|172.16.1.0" | cut -
VRF |   Prefix/Len  | RT|CLSS| Flags
4675| 172.16.1.0/ 24| UC| f|spi,dpi
Leaf101#
<-- When the prefix is configured under the External EPG, a classification of the pcTag in hexadecimal
Leaf101#
vsh_lc -c 'show platform internal hal l3 routes vrf tz:tz-VRF_1' | egrep "Prefix/Len|172.16.1.0" | cut -
VRF |   Prefix/Len  | RT|CLSS| Flags
4675| 172.16.1.0/ 24| UC|8003|spi,dpi
Leaf101#
Leaf101#
vsh_lc -c '
dec 0x8003'
32771
Leaf101#
```

Par la suite, lorsqu'un sous-réseau est configuré avec l'indicateur Sous-réseau externe pour EPG externe dans un ExEPG, un processus interne appelé Policy Manager (policy-mgr) met à jour sa table de mappage préfixe vers pcTag avec cette entrée de sous-réseau et le pcTag associé. Policy Manager sert de moteur d'orchestration centralisée des politiques de fabric, traduisant les définitions de politiques de haut niveau en configurations exploitables sur l'ensemble du fabric ACI. Cela garantit une connectivité des applications et un comportement réseau cohérents et sécurisés en appliquant les pcTags appropriés pour la classification du trafic et les décisions de transfert en fonction des sous-réseaux externes configurés.

<#root>

```
Leaf101#
vsh -c 'show system internal policy-mgr prefix' | egrep "tz:tz-VRF_1"
3047430 36 0x80000024 Up tz:tz-VRF_1 ::/0 15 True True False False
3047430 36 0x24 Up tz:tz-VRF_1 0.0.0.0/0 15 True True False False
3047430 36 0x24 Up tz:tz-VRF_1 172.16.1.0/24 32771 True True False False
Leaf101#
```

Cela confirme que le préfixe 172.16.1.0/24 est annoncé par le voisin au commutateur leaf de périphérie ACI, et l'ACI a classé le préfixe sous pcTag 32771

Contrats

Une règle de zonage est le processus sous-jacent qui applique les politiques de contrat entre les groupes de terminaux (y compris les groupes de terminaux ExEPG) au sein du fabric. Le VNID VRF (portée) et le pcTag de l'EPG externe peuvent être utilisés pour définir et valider les règles de communication appliquées entre les EPG source et de destination. Les règles de zonage traduisent essentiellement les relations contractuelles de haut niveau en règles spécifiques et applicables, programmées sur les commutateurs Leaf.

L'emplacement d'installation du contrat dans le fabric est un aspect important à prendre en compte. Par défaut, le VRF est configuré avec la direction d'application du contrôle des stratégies définie sur Entrée. Ce paramètre détermine que la règle de zonage d'un contrat donné est installée sur le commutateur Leaf où réside le point d'extrémité source.

Pour cet exercice, le trafic est entrant à partir d'une L3Out, la règle de zonage est installée sur le noeud leaf périphérique qui se connecte à cette L3Out, car ce noeud leaf agit comme noeud leaf source pour le trafic entrant dans le fabric.

<#root>

```
Leaf101#
```

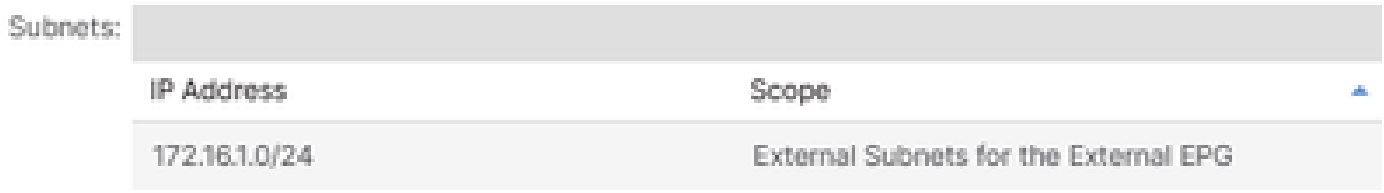
show zoning-rule scope 3047430 | egrep "Rule|---|32771"

```
+---------+--------+--------+----------+----------------+---------+---------+------------------+------
| Rule ID | SrcEPG | DstEPG | FilterID |      Dir       | operSt  |  Scope  |        Name      |
+---------+--------+--------+----------+----------------+---------+---------+------------------+------
|  4441   | 49153  | 32771  |     5    |     bi-dir     | enabled | 3047430 |     tz:Contract  |
|  4500   | 32771  | 49153  |     5    | uni-dir-ignore | enabled | 3047430 |     tz:Contract  |
+---------+--------+--------+----------+----------------+---------+---------+------------------+------
Leaf101#
```

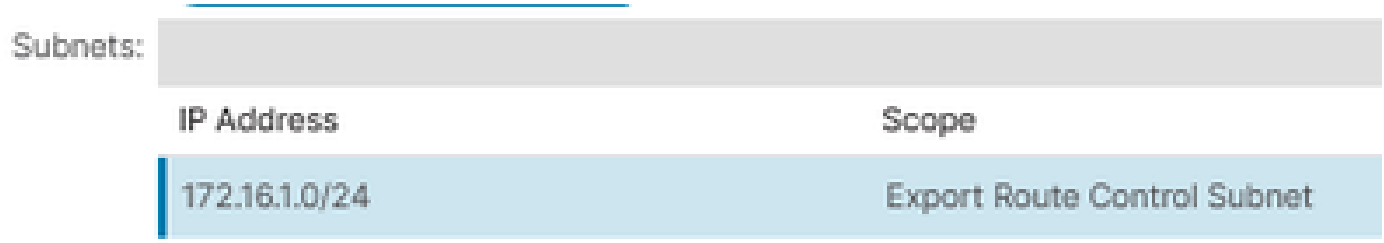
Routage de transit

Le routage de transit permet au fabric d'agir en tant que réseau de transit en annonçant des

routes externes apprises d'une L3Out à une autre. Pour configurer correctement le routage de transit, le sous-réseau entrant doit être marqué avec l'indicateur Sous-réseau externe pour EPG externe.



Simultanément, l'indicateur de sous-réseau de contrôle de route d'exportation doit être activé sur le sous-réseau correspondant de l'interface L3Out qui annonce ce sous-réseau à d'autres homologues externes. Cet indicateur permet de redistribuer le sous-réseau et de l'annoncer hors du fabric via le protocole de routage configuré sur ce L3Out.



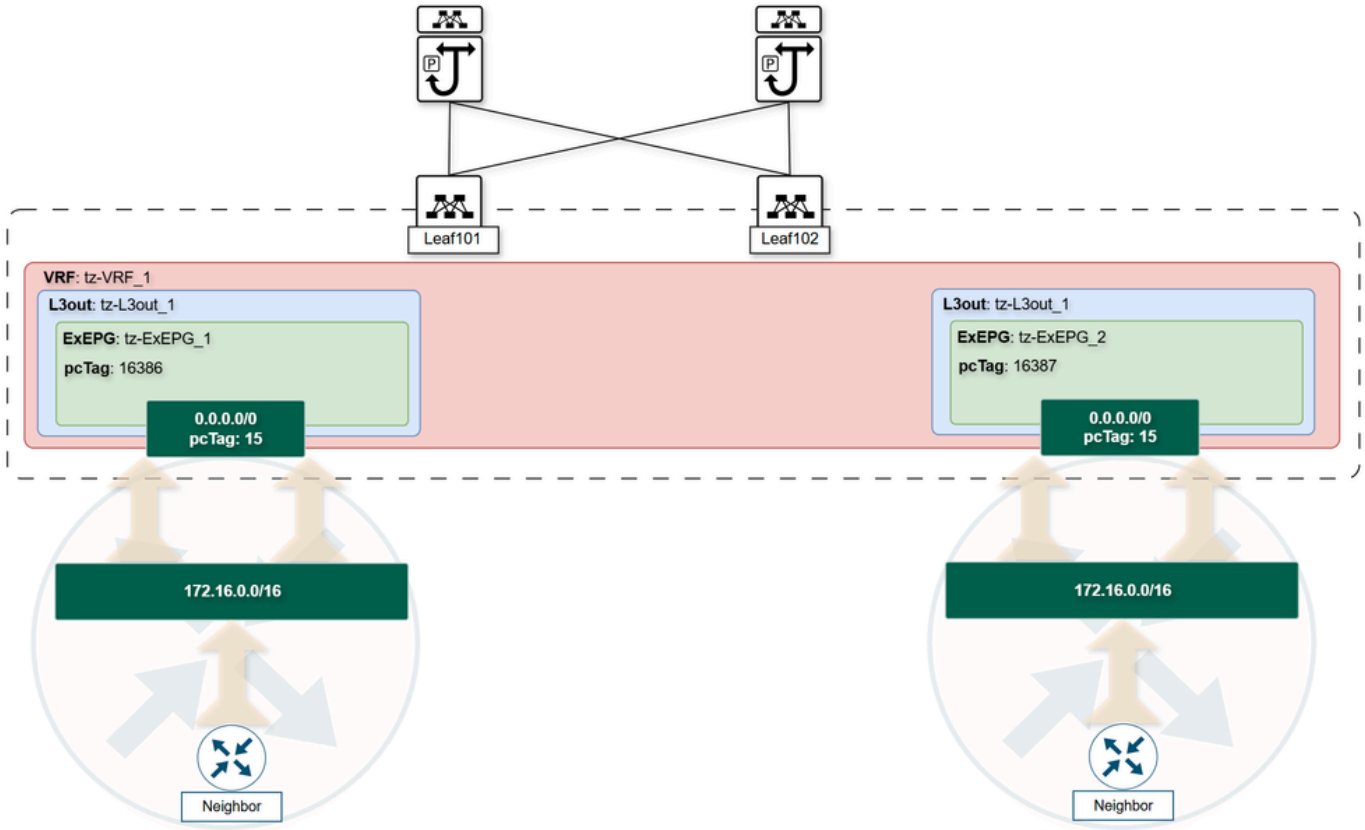
Enfin, un contrat entre le L3out reçu et le L3out d'exportation doit être configuré pour terminer le processus de distribution de la route.

Problèmes courants dans la classification EPG externe de sous-réseau

pcTag 15

Auparavant, dans ce document, il était indiqué que le sous-réseau ExEPG vous aide à classer les sous-réseaux dans le pcTag correct pour des raisons d'application de la stratégie. Une exception importante à cette classification est le sous-réseau quad-0 (0.0.0.0/0) lorsqu'il est configuré avec l'indicateur Sous-réseau externe pour EPG externe. Ce sous-réseau se voit toujours attribuer le pcTag 15 réservé, qui agit en fait comme un caractère générique pour tout le trafic externe au sein d'un VRF.

Ce schéma représente le problème de configuration de quad-0 avec un sous-réseau externe pour un EPG externe sur plusieurs ExEPG au sein du même VRF :



- Le sous-réseau quad-0 est souvent confondu avec la route par défaut. Bien que cela soit parfois vrai, par exemple lorsqu'un voisin de routage dynamique annonce uniquement la route par défaut à l'ACI L3Out, le rôle du sous-réseau quad-0 dans l'ACI est plus large en tant que classification globale.
- Il est courant de configurer plusieurs groupes de terminaux ExEPG avec le sous-réseau quad-0 pour accepter tous les préfixes annoncés par un voisin. Bien que cela atteigne l'objectif d'une large acceptation, cela peut conduire à un routage asymétrique inattendu lorsque plusieurs ExEPG avec quad-0 sont configurés dans le même VRF. Lorsque plusieurs ExEPG dans le même VRF sont configurés avec quad-0 en tant que sous-réseau externe, l'ACI ne peut pas sélectionner de manière déterministe la L3Out à utiliser pour un sous-réseau de destination spécifique. Au lieu de cela, il sélectionne un L3Out arbitrairement.
- Ce comportement peut entraîner un routage asymétrique, un trafic intermittent ou même des pertes de trafic si L3Out sélectionné aléatoirement ne dispose pas des contrats nécessaires pour autoriser la communication.

Chevauchement de sous-réseaux

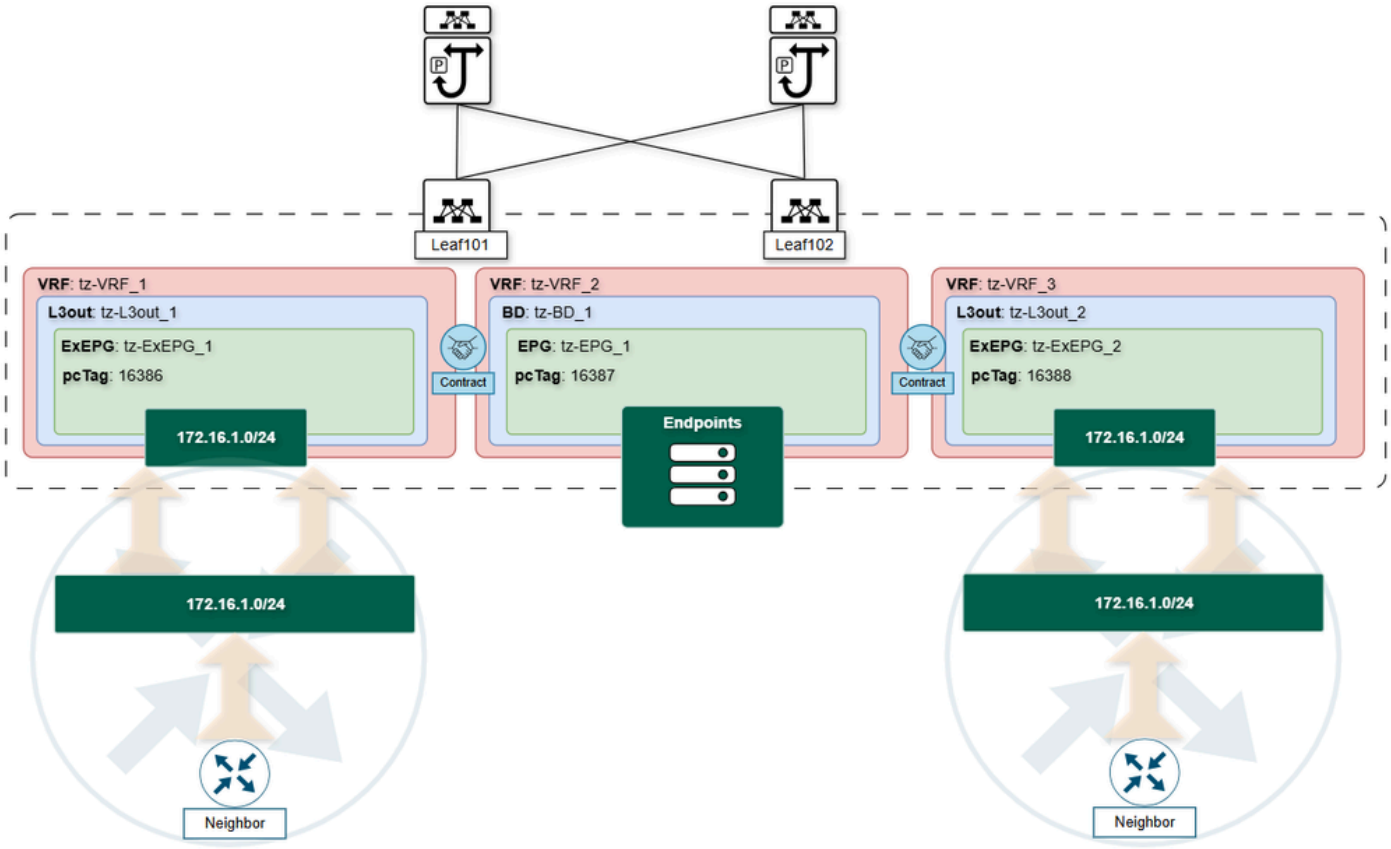
La configuration de sous-réseaux identiques sur différents groupes de terminaux ExEPG n'est pas autorisée. Si vous tentez de le faire, le message « F0467 : Préfixe déjà utilisé dans un autre EPG, ce qui empêche la duplication de sous-réseau dans un VRF.

Cependant, les sous-réseaux qui se chevauchent peuvent exister entre différents VRF, car chaque VRF conserve un contexte de table de routage indépendant. Cette séparation permet de configurer le même sous-réseau dans des groupes de terminaux virtuels (ExEPG) appartenant à des VRF différents. Malgré cela, la prudence est essentielle lors de l'exécution de fuites de route

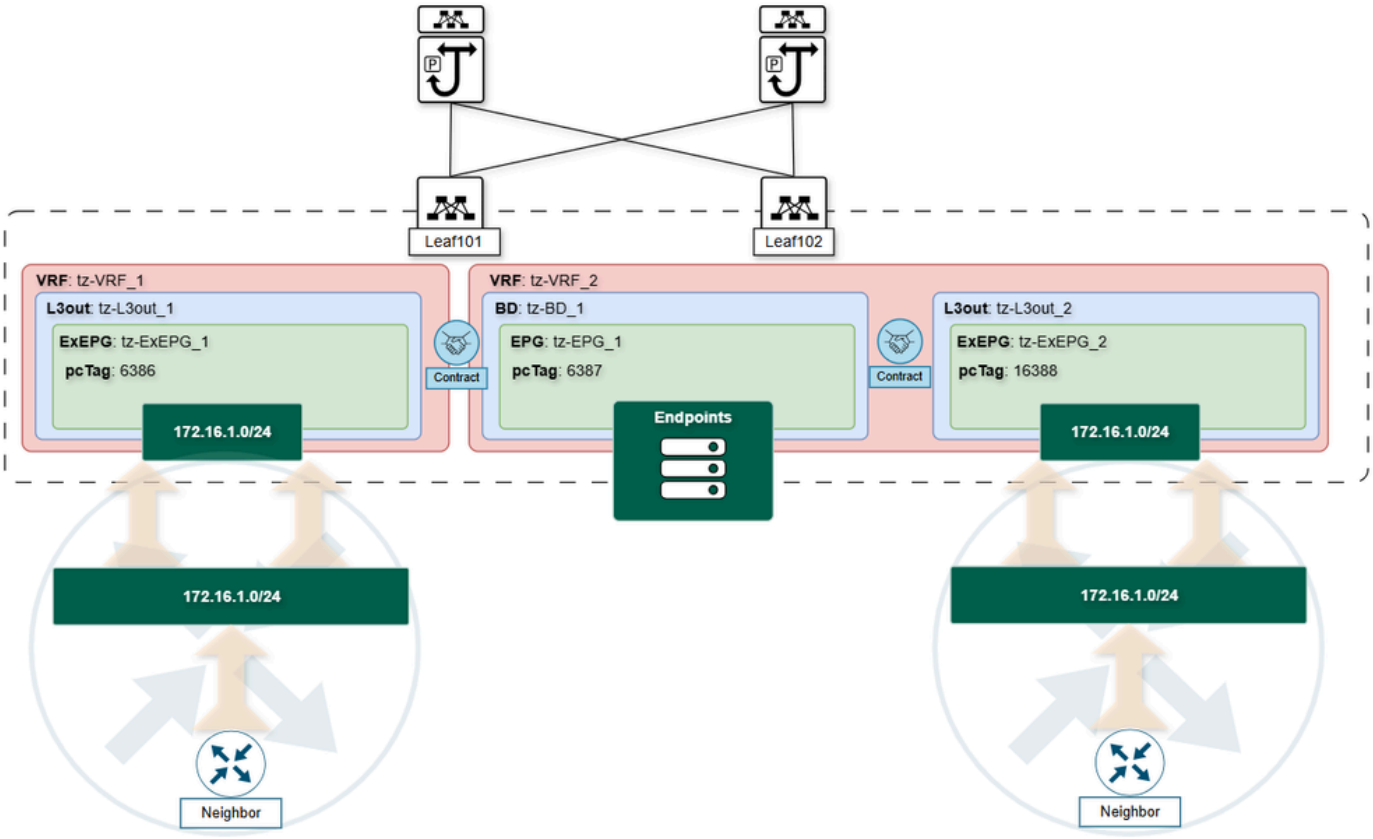
VRF impliquant ces sous-réseaux qui se chevauchent, car cela peut conduire à des décisions de transfert asymétriques en raison de conflits dans la classification de sous-réseau (pcTag) par rapport aux informations de routage (RIB).

Les scénarios clés incluent :

- Fuite de route de deux VRF vers un troisième VRF :
  Lorsque deux VRF font passer le même sous-réseau dans un troisième VRF, le VRF récepteur installe le premier sous-réseau qu'il reçoit en fonction de la stratégie partagée du contrôleur APIC. Si le commutateur leaf qui gère ce VRF redémarre ou si la sélection de routage change, la table de routage peut être mise à jour vers une autre L3Out, ce qui entraîne un comportement de transfert incohérent.



- Chevauchement ExEPG L3Out local vers VRF avec fuites de sous-réseaux :
  Dans les conceptions où la fuite de route est utilisée, si un ExEPG L3Out local est configuré avec le même sous-réseau qu'un sous-réseau ayant fui, l'entrée de routage local a toujours priorité sur les routes ayant fui.



Ces situations mettent en évidence que les problèmes de transfert asymétrique proviennent de la couche de classification et de décision de transfert, et non de la table de routage elle-même. Alors que la classification de sous-réseau associe un sous-réseau à une L3Out et à un ExEPG spécifiques pour l'application de la stratégie, la table de routage peut pointer vers une destination L3Out différente. Cette disparité peut entraîner un transfert incohérent du trafic, ce qui peut entraîner des problèmes de connectivité ou des lacunes dans l'application des politiques.

Importer le changement de comportement par défaut du contrôle de route

Par défaut, l'ACI accepte toutes les annonces de routes entrantes des voisins. Pour contrôler quels préfixes sont acceptés, vous devez activer l'application du contrôle de route : entrant sur l'objet racine L3Out :

Accédez à Locants > [nom du locataire] > Networking > L3out > [nom L3out].

Cette action crée une route-map sous le protocole de routage sélectionné.

<#root>

Border Leaf#

show ip bgp neighbors vrf tz:tz-VRF1 | egrep route-map

Outbound route-map configured is exp-l3out-ExEPG-peer-2981888, handle obtained

Inbound route-map configured is imp-l3out-ExEPG-peer-2981888, handle obtained

Border Leaf#

show route-map imp-l3out-ExEPG-peer-2981888

route-map imp-l3out-ExEPG-peer-2981888,

permit

, sequence 15801
Match clauses:

ip address prefix-lists: IPv4-peer49155-2981888-exc-ext-inferred-import-dst

ipv6 address prefix-lists: IPv6-deny-all
Set clauses:

Border Leaf#

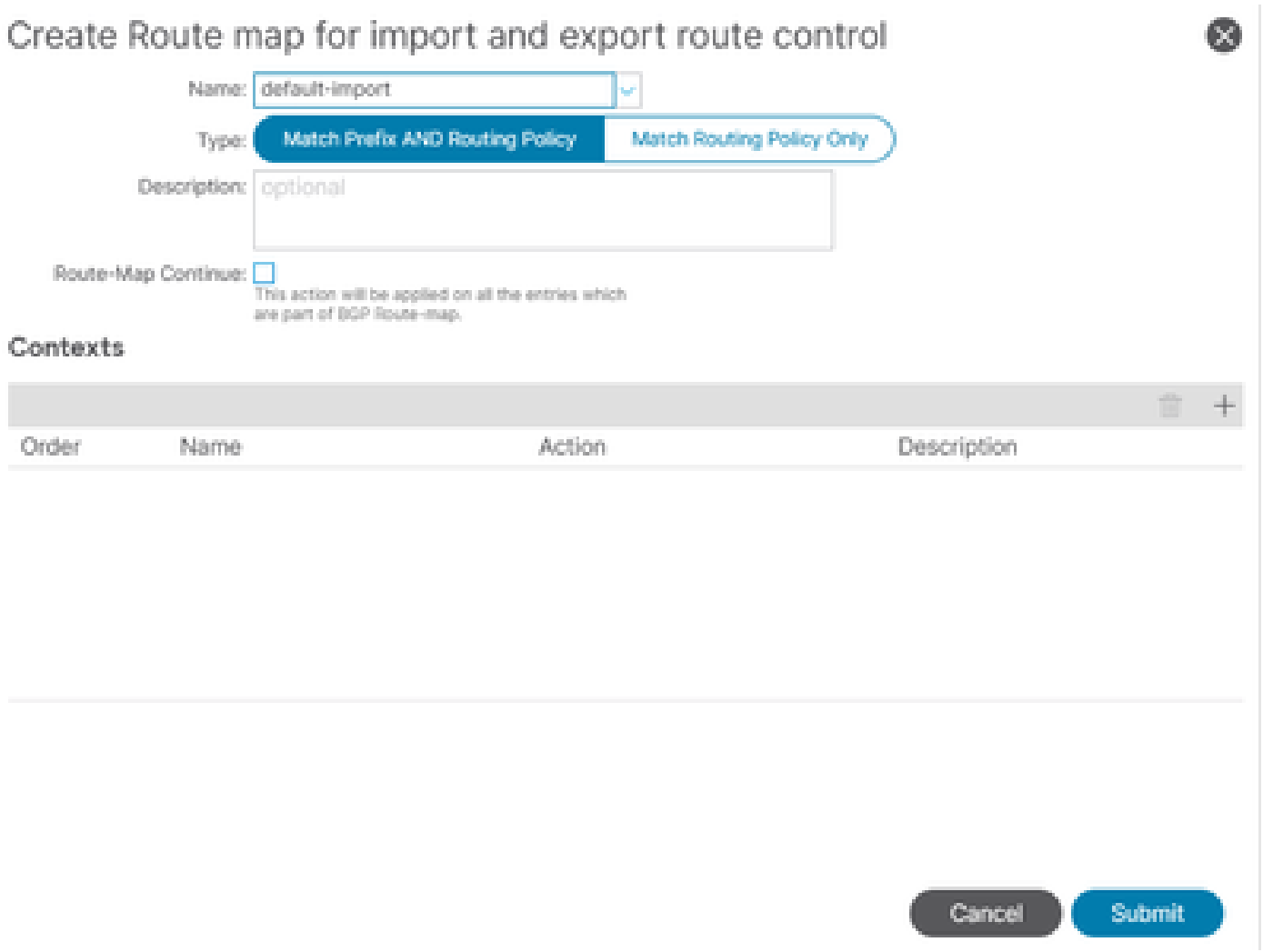
show ip prefix-list IPv4-peer49155-2981888-exc-ext-inferred-import-dst

ip prefix-list IPv4-peer49155-2981888-exc-ext-inferred-import-dst: 1 entries
seq 1 permit 172.16.1.0/24
Border Leaf#

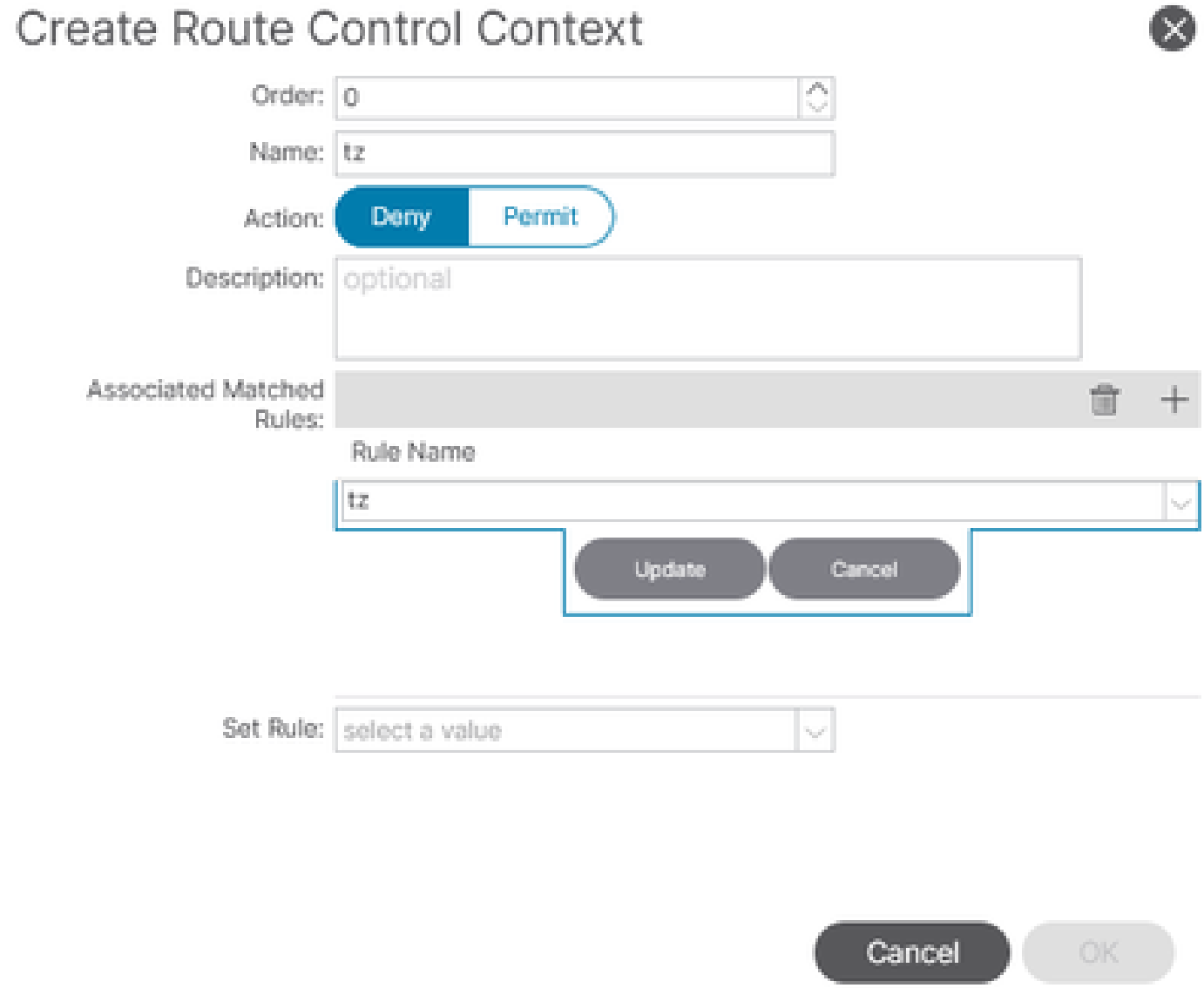
Par défaut, cette route-map d'importation autorise tous les préfixes entrants. Pour modifier ce comportement :

Accédez à Tenants > [nom du locataire] > Networking > L3out > [nom L3out] > Route map pour le contrôle de route d'importation et d'exportation

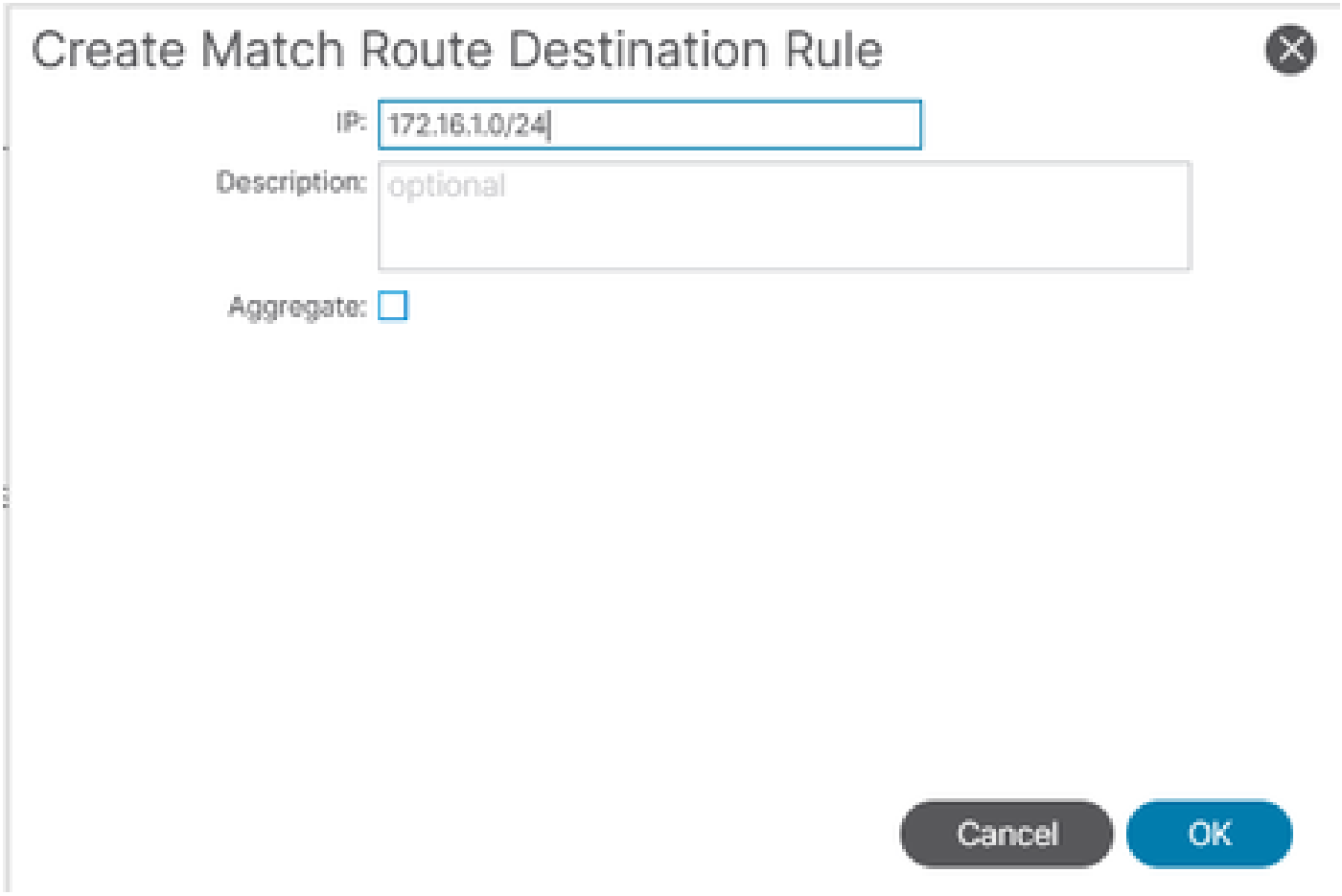
Sélectionnez la carte de routage d'importation par défaut ou créez-en une nouvelle à l'aide de l'icône d'engrenage en haut à droite.



Dans la section Contexte, créez une nouvelle règle associée correspondante.



Dans la section Règles de correspondance, faites défiler jusqu'à Préfixe de correspondance et ajoutez le ou les sous-réseaux spécifiques que vous souhaitez contrôler.



Après l'envoi des stratégies, l'action d'importation de route-map change en conséquence, appliquant le filtrage de préfixe souhaité.

<#root>

```
Border Leaf#
```

show route-map imp-l3out-ExEPG-peer-2981888

```
route-map imp-l3out-ExEPG-peer-2981888,
```

deny

```
, sequence 8001
Match clauses:
ip address prefix-lists: IPv4-peer49155-2981888-exc-ext-in-default-import2tzOtz-dst
ipv6 address prefix-lists: IPv6-deny-all
Set clauses:

Border Leaf#
```

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.