Comprendre le livre blanc BGP RPKI avec XR7 Cisco8000

Table des matières

Introduction

Ce document décrit la fonctionnalité Resource Public Key Infrastructure (RPKI) du protocole BGP (Border Gateway Protocol) sur la plate-forme Cisco IOS® XR.

Informations générales

Préface

Ce document traite de la fonctionnalité BGP RPKI et de la façon dont elle protège BGP avec les routeurs contre les mises à jour de préfixe BGP fausses/malveillantes.

Portée

Ce document utilise la version Cisco 8000 avec XR 7.3.1 pour la démonstration. Cependant, BGP RPKI est une fonctionnalité indépendante de la plate-forme, les concepts abordés dans ce document s'appliquent à d'autres plates-formes Cisco (avec Cisco IOS, Cisco IOS-XE .) avec des conversions de CLI équivalentes appropriées. Ce document ne couvre pas la procédure d'ajout d'autorisations d'origine de route (ROA) sur les registres Internet régionaux.

Conditions préalables

Le lecteur doit connaître le protocole BGP.

Avertissement

Les adresses IP (Internet Protocol) utilisées dans ce document ne sont pas censées être des adresses réelles. Tous les exemples, les résultats d'affichage des commandes et les figures inclus dans le document sont présentés à titre d'illustration uniquement. Toute utilisation d’adresses IP réelles dans un contenu d’illustration est involontaire et fortuite.

Problèmes BGP dus à une annonce de préfixe incorrecte

Le protocole BGP sert de réseau fédérateur du trafic Internet. Bien qu'il s'agisse du composant le plus important d'Internet Core, il n'a pas la capacité de vérifier si l'annonce BGP entrante provient d'un système autonome autorisé ou non.

Cette limitation du protocole BGP en fait un candidat facile pour divers types d'attaques. Une attaque courante est appelée « détournement de route ». Cette attaque peut être exploitée pour :

- Le vol d'adresses IP pour envoyer du courrier indésirable entraîne le rejet de l'adresse IP et donc le déni de service.
- Espionner le trafic pour obtenir des informations sensibles comme les mots de passe.
- Interruptions dues à des configurations incorrectes par l'administrateur.
- Empêchez la livraison du trafic en installant de faux serveurs, ce qui entraîne un déni de service.

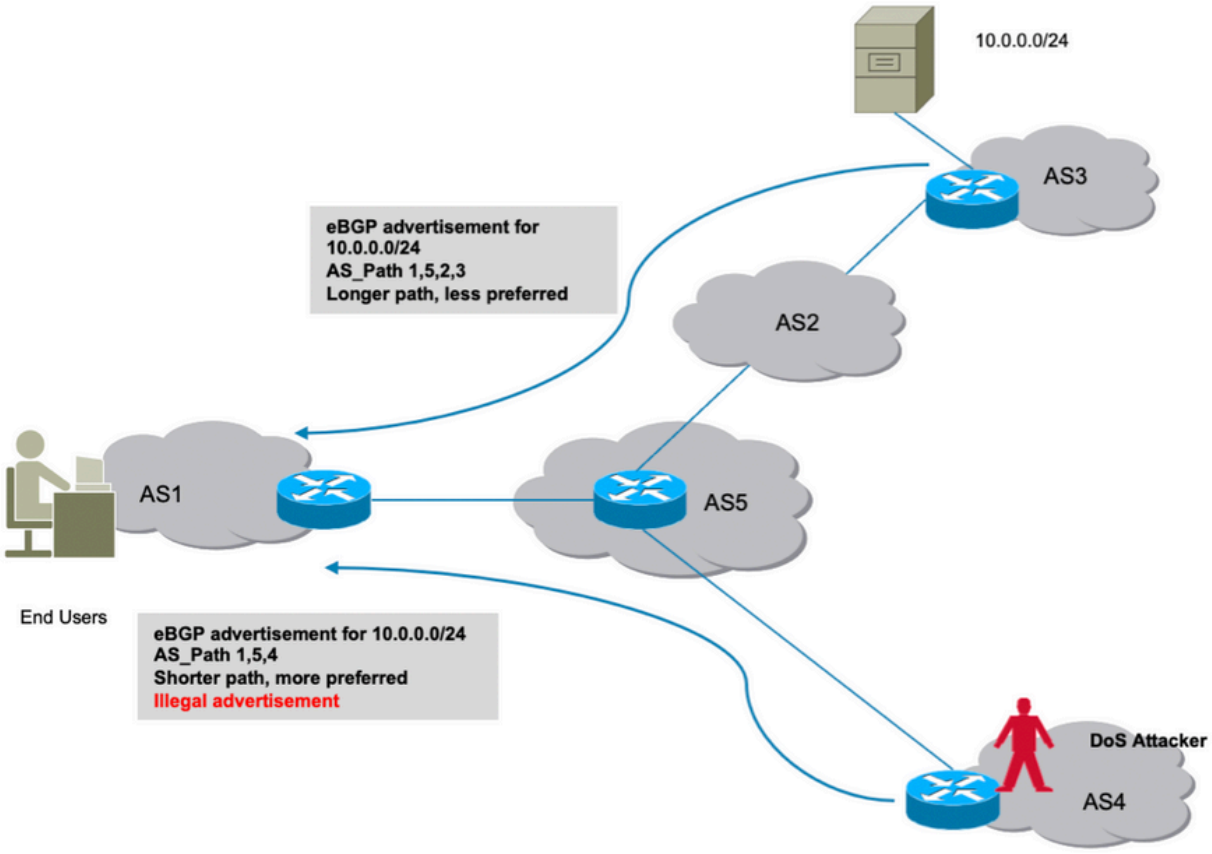
L'attaque par déni de service (DoS) est une tentative malveillante visant à perturber le trafic normal vers un routeur, un commutateur, un serveur, etc. Il existe une variété d'attaques DoS et

peu d'entre elles sont abordées ici.

Piratage de route

Considérez le scénario présenté ici. Le système autonome 3 (AS3) envoie une annonce BGP légale pour son préfixe 10.0.0.0/24. D'après la conception de BGP, rien dans BGP n'empêcherait un attaquant d'annoncer le même préfixe à Internet.

Comme illustré, le pirate de l'AS4 annonce le même préfixe 10.0.0.0/24. L'algorithme du meilleur chemin BGP préfère un chemin avec AS_Path plus court. AS_Path 1,5,4 gagne sur un chemin plus long via AS 1,5,2,3. Par conséquent, le trafic provenant des clients sera désormais redirigé vers l'environnement de l'attaquant et pourra être bloqué, ce qui entraînera un déni de service pour les clients finaux.

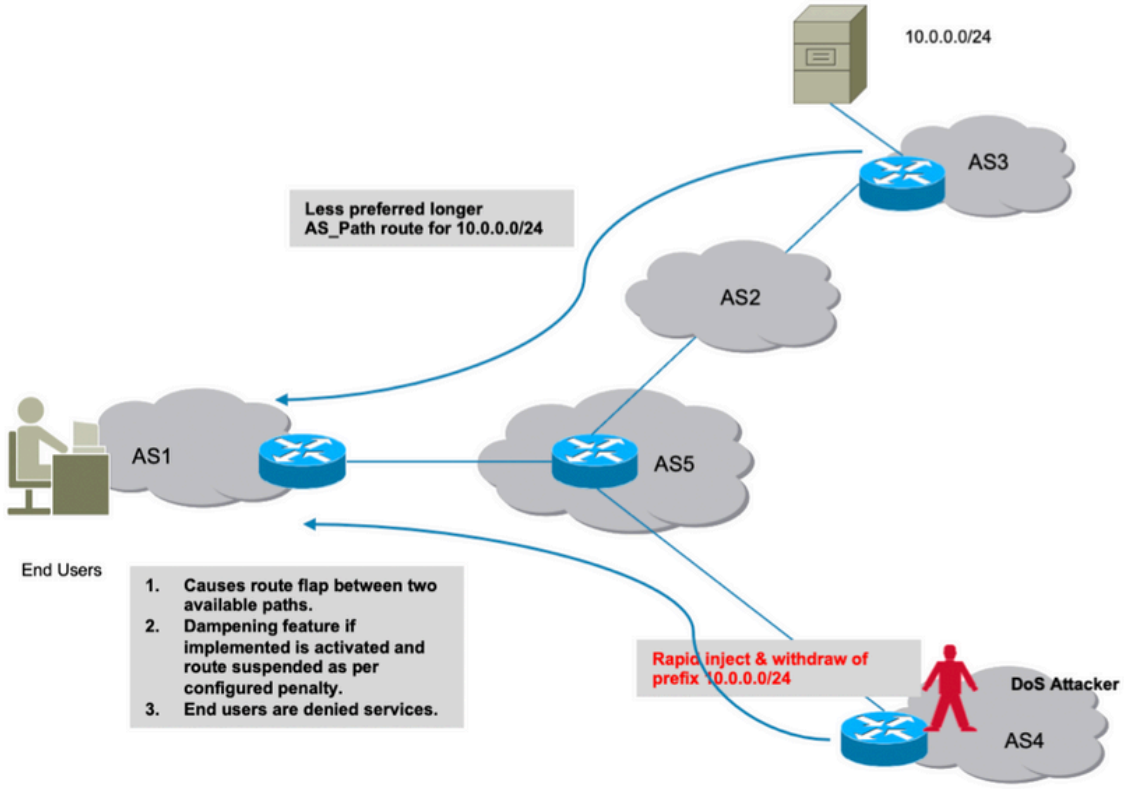


Détournement de route

Dégrader les performances du système

Cette section décrit une autre façon de refuser des services. Si la fonctionnalité d'atténuation de route BGP de Cisco est configurée, elle pourrait être exploitée si le pirate introduit des ailerons de route rapides dans le réseau provoquant un désordre constant.

La fonction d'amortissement impose des pénalités à la route légitime et la rend indisponible pour le trafic réel. En outre, ce type de défaillance non éthique peut surcharger les ressources du routeur, telles que le processeur, la mémoire, etc.

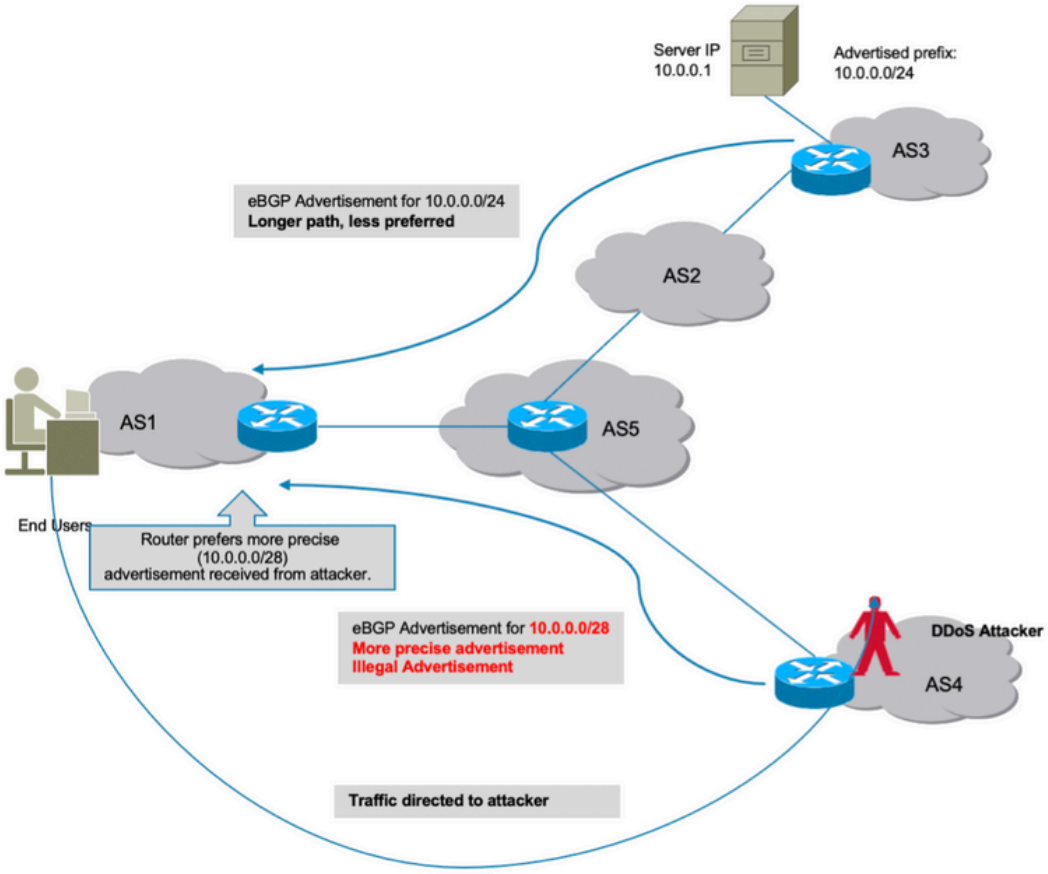


Amortissement D'Itinéraire

Piratage de sous-préfixe

Comme nous l'avons vu dans la section précédente, comment un pirate peut créer un préfixe illégalement et perturber le trafic. Malheureusement, une perturbation n'est pas la seule cause d'inquiétude. Dans de telles attaques, les données réelles peuvent être compromises, un pirate pouvant analyser les données reçues à des fins contraires à l'éthique.

De même, le détournement d'une route pourrait se faire en annonçant illégalement une route plus précise. Le protocole BGP préfère les préfixes qui correspondent plus longtemps et ce comportement peut être mal exploité, comme illustré dans l'image.



Détournement de sous-préfixe

Toutes les attaques qui sont discutées proviennent du fait que BGP n'a pas pu identifier si l'AS d'origine de ces préfixes annoncés de manière malveillante était valide ou non. Pour résoudre ce problème, une source de données « vraie » et « fiable » est nécessaire, qu'un routeur peut conserver dans sa base de données. Ensuite, à chaque réception d'une nouvelle annonce, le routeur devient maintenant capable de vérifier de manière croisée les informations d'origine AS du préfixe reçues de l'homologue BGP avec les informations de sa base de données locale du validateur.

Ainsi, le routeur est capable de distinguer les bonnes annonces des mauvaises (illégales) et la capacité d'éviter toutes les attaques discutées précédemment est ajoutée de manière inhérente sur le routeur. BGP RPKI fournit la source d'informations fiable requise.

RPKI

RPKI utilise un référentiel qui contient des ROA. Un ROA contient des informations sur le préfixe et leur numéro de système autonome BGP associé. L'autorisation d'origine de route est une instruction cryptographiquement signée.

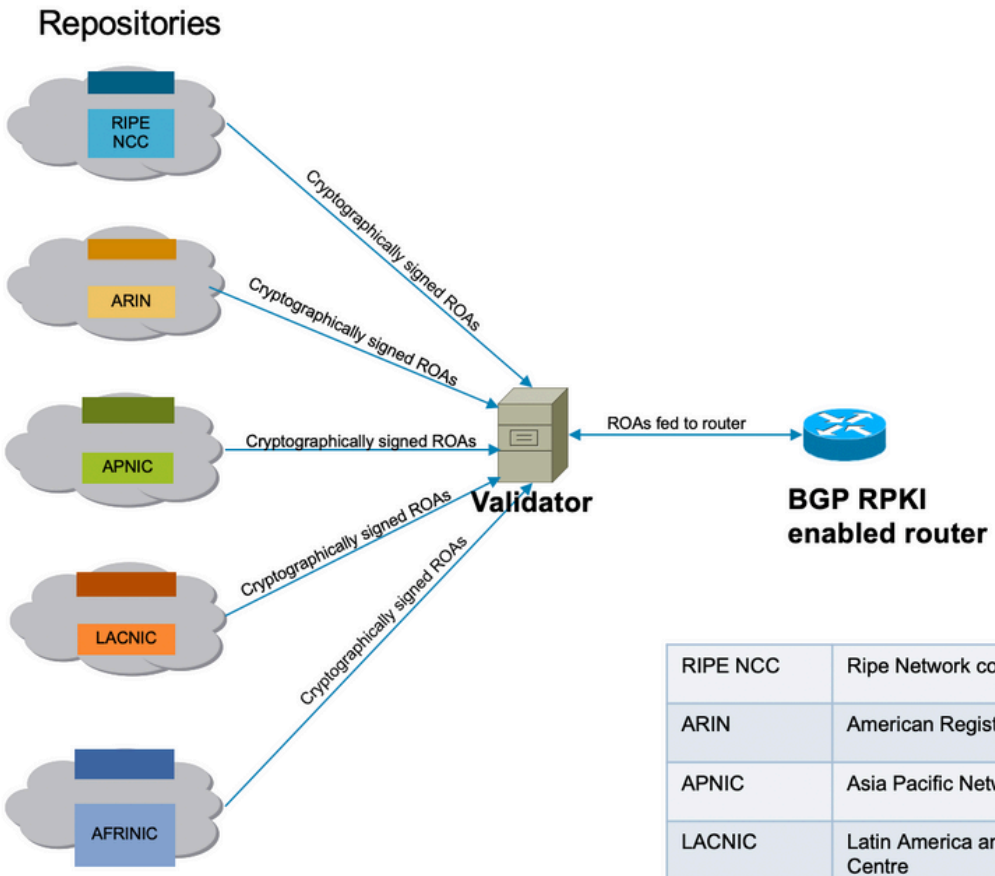
Les 5 RIR (Regional Internet Registries) sont les ancres de confiance de l'ICPR. L'IANA (Internet Assigned Numbers Authority) est la partie supérieure de l'arborescence qui distribue les préfixes IP. Les RIR sont les suivants dans la hiérarchie. Ils attribuent des sous-préfixes aux registres Internet locaux (LIR) et aux grands fournisseurs d'accès Internet (FAI). Ils signent un certificat pour ces préfixes. Le niveau suivant attribue des sous-préfixes de ceux-ci et utilise les certificats ci-dessus pour signer leurs propres certificats afin de certifier leurs propres allocations. Ils utilisent

généralement leurs propres points de publication pour héberger les certificats et les ROA. Chaque certificat répertorie les points de publication des certificats enfants qu'il signe. Ainsi, RPKI forme une arborescence de certificats qui reflète l'arborescence des allocations d'adresses IP. Les validateurs RPKI appartenant aux parties utilisatrices interrogent tous les points de publication pour trouver les certificats et les ROA mis à jour (ainsi que les LCR et les manifestes). Ils commencent au niveau des ancres d'approbation et suivent les liens vers les points de publication des certificats enfants.

Les ROA sont enregistrés dans le référentiel par le biais des RIR, mais la même chose peut être effectuée via d'autres registres (nationaux ou locaux). Cette responsabilité peut également être déléguée aux FSI, avec une supervision et une vérification appropriées par les RIR.

Àl'heure actuelle, il existe cinq dépôts ROA gérés par RIPE NCE, ARIN, APNIC, LACNIC et AFRINIC.

Un validateur présent sur le réseau communique avec ces référentiels et télécharge une base de données ROA approuvée pour créer son cache. Il s'agit d'une copie fusionnée de l'infrastructure RPKI, qui est périodiquement récupérée/actualisée directement ou indirectement à partir de l'infrastructure RPKI globale. Le validateur transmet ensuite ces informations aux routeurs, ce qui leur permet de comparer les annonces BGP entrantes avec la table RPKI afin de prendre une décision en toute sécurité.



RIPE NCC	Ripe Network coordination center
ARIN	American Registry for Internet Numbers
APNIC	Asia Pacific Network Information Centre
LACNIC	Latin America and Caribbean Network Information Centre
AFRINIC	African Network Information Centre

Connectivité de l'infrastructure RPKI

Validateur

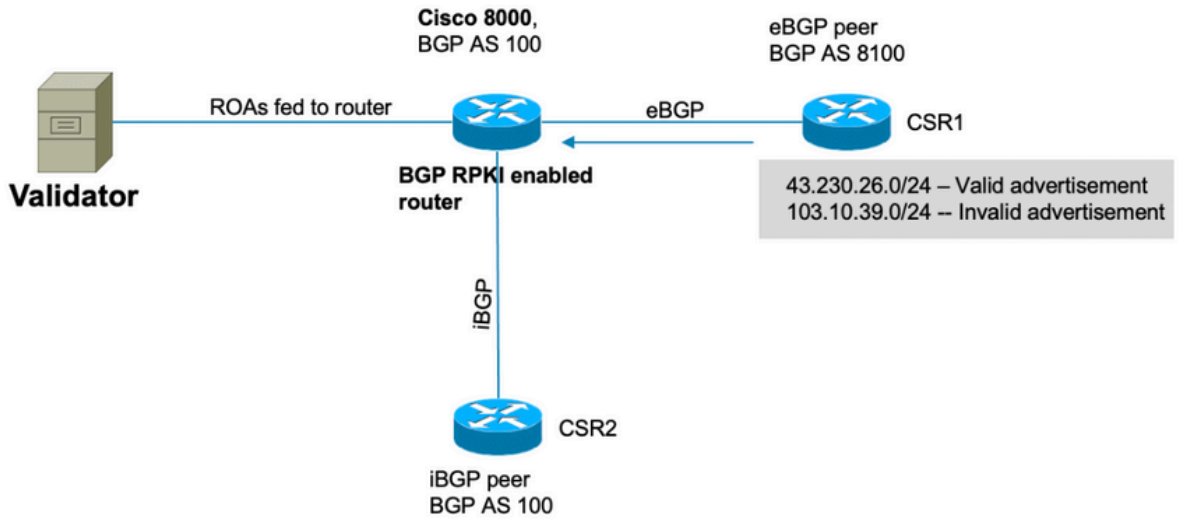
Cette démonstration utilise le validateur RIPE. Le validateur communique avec le routeur en établissant une session TCP. Dans cette démonstration, le validateur écoute son adresse IP 192.168.122.120 et son port 3323.

```
routinator server --rtr 192.168.122.120:3323 --refresh=900
```

L'IANA a spécifié le port 3323 pour cette communication. Le compteur d'actualisation définit l'intervalle de temps après lequel le référentiel local sera synchronisé et mis à jour pour rester à jour.

Démonstration de BGP RPKI

Topologie



Topologie

> **Remarque** : cette démonstration utilise un numéro et des préfixes AS publics aléatoires simplement pour expliquer la mécanique RPKI de BGP. Les adresses IP publiques sont utilisées en raison de la RPKI, qui est principalement destinée à la protection des préfixes publics, et tous les ROA créés sur les RIR sont des préfixes publics. Enfin, aucune des actions, configurations, etc. décrites dans ce document n'affecte ces IP et AS publics de quelque manière que ce soit.

Configurer

```
router bgp 100

bgp router-id 10.1.1.1

rpki server 192.168.122.120
```

```
 transport tcp port 3323
 refresh-time 900


address-family ipv4 unicast
 !
 neighbor 10.0.12.2
 remote-as 8100
 address-family ipv4 unicast
  route-policy Pass in
  route-policy Pass out
 !
 !
 neighbor 10.0.13.3
 remote-as 100
 address-family ipv4 unicast
 !
 !
// ‘Pass’ is a permit all route-policy.
```

Session RPKI BGP

Le routeur établit une session TCP avec un validateur (IP : 192.168.122.120, port 3323) afin de télécharger le cache ROA dans la mémoire du routeur.

```
RP/0/RP0/CPU0:Cisco8000#show bgp rpki server 192.168.122.120
Wed Jan 20 22:54:15.763 UTC

RPKI Cache-Server 192.168.122.120
 Transport: TCP port 3323
 Bind source: (not configured)
 Connect state: ESTAB
 Conn attempts: 1
 Total byte RX: 4428792
 Total byte TX: 1400
```

```
 Last reset

   Timest: Jan 20 05:59:58 (16:54:17 ago)

   Reason: protocol error
```

Téléchargements ROA sur le routeur

Le validateur transmet les informations ROA au routeur. Ce cache est actualisé à intervalles réguliers afin de minimiser la possibilité que le routeur contienne des informations périmées. Dans cette démonstration, un temps d'actualisation de 900 secondes a été configuré. Comme indiqué ici, le routeur Cisco 8000 a téléchargé 172632 ROA IPv4 et 28350 IPv6 à partir du validateur.

```
RP/0/RP0/CPU0:Cisco8000#show bgp rpki server summary

Wed Jan 20 23:01:59.432 UTC

Hostname/Address        Transport       State           Time                ROAs (IPv4/IPv6)

192.168.122.120         TCP:3323        ESTAB           17:00:21        172632/28350

RP/0/RP0/CPU0:Cisco8000#show bgp rpki table ipv4

Wed Jan 20 23:09:26.899 UTC

>>>Snipped output<<<

 Network               Maxlen      Origin-AS         Server

 10.0.0.0/24            24              13335             192.168.122.120

 10.0.4.0/22            22              38803             192.168.122.120

 10.0.4.0/24            24              38803             192.168.122.120

 10.0.5.0/24            24              38803             192.168.122.120

 10.0.6.0/24            24              38803             192.168.122.120

 10.0.7.0/24            24              38803             192.168.122.120

 10.1.1.0/24            24              13335             192.168.122.120

 10.1.4.0/22            22              4134              192.168.122.120

 10.1.16.0/20           20              4134              192.168.122.120

 10.2.9.0/24            24              4134              192.168.122.120

 10.2.10.0/24           24              4134              192.168.122.120

 10.2.11.0/24           24              4134              192.168.122.120

 10.2.12.0/22           22              4134              192.168.122.120
```

```
 10.3.0.0/16            16              4134              192.168.122.120

 10.6.0.0/22            24              9583              192.168.122.120
```

Vérifier

Cette section montre comment le protocole BGP RPKI est actif et comment il empêche le routeur d'afficher des annonces incorrectes/illégales.

Activation de la validité Origin-As

Par défaut, le routeur extrait les ROA du validateur, mais ne commence pas à les utiliser tant qu'il n'a pas été configuré à cet effet. Par conséquent, ces préfixes sont marqués comme « D » ou désactivés.

```
RP/0/RP0/CPU0:Cisco8000#show bgp origin-as validity

Wed Jan 20 23:27:37.268 UTC

BGP router identifier 10.1.1.1, local AS number 100

BGP generic scan interval 60 secs

Non-stop routing is enabled

BGP table state: Active

Table ID: 0xe0000000   RD version: 30

BGP main routing table version 30

BGP NSR Initial initsync version 2 (Reached)

BGP NSR/ISSU Sync-Group versions 0/0

BGP scan interval 60 secs


Status codes: s suppressed, d damped, h history, * valid, > best

              i - internal, r RIB-failure, S stale, N Nexthop-discard

Origin codes: i - IGP, e - EGP, ? - incomplete

Origin-AS validation codes: V valid, I invalid, N not-found, D disabled

   Network            Next Hop            Metric LocPrf Weight Path

D*> 203.0.113.0/24     10.0.12.2                0             0 8100 ?

D*> 203.0.113.1/24     10.0.12.2                0             0 8100 ?

D*> 192.168.122.1/32   10.0.12.2                0             0 8100 ?
```

Afin d'activer le routeur pour le contrôle de validité as-origin, activez cette commande pour la famille d'adresses concernée.

```
router bgp 100

 address-family ipv4 unicast

 bgp origin-as validation enable

 !
```

Lorsque vous activez cette commande, le routeur analyse les préfixes présents dans sa table BGP par rapport aux informations ROA reçues du validateur et l'un des trois états est attribué aux préfixes .

```
RP/0/RP0/CPU0:Cisco8000#show bgp origin-as validity

Thu Jan 21 00:04:58.136 UTC

Status codes: s suppressed, d damped, h history, * valid, > best

              i - internal, r RIB-failure, S stale, N Nexthop-discard

Origin codes: i - IGP, e - EGP, ? - incomplete

Origin-AS validation codes: V valid, I invalid, N not-found, D disabled

   Network            Next Hop            Metric LocPrf Weight Path

V*> 203.0.113.0/24     10.0.12.2                0             0 8100 ?

I*  203.0.113.1/24     10.0.12.2                0             0 8100 ?

N*> 192.168.122.1/32   10.0.12.2                0             0 8100 ?
```

Afin de permettre au routeur d'utiliser les informations d'état de validation de préfixe tout en effectuant le meilleur calcul de chemin, cette commande est nécessaire. Cette option n'est pas activée par défaut, car elle vous permet de ne pas utiliser les informations de validité pour le meilleur calcul de chemin, mais de les utiliser dans les politiques de routage qui sont traitées plus loin dans ce document.

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as use validity

 !
```

États de validité du préfixe

Il existe trois états dans lesquels un préfixe peut être trouvé.

```
RP/0/RP0/CPU0:Cisco8000#show bgp origin-as validity

Thu Jan 21 00:04:58.136 UTC
```

```
Status codes: s suppressed, d damped, h history, * valid, > best

              i - internal, r RIB-failure, S stale, N Nexthop-discard

Origin codes: i - IGP, e - EGP, ? - incomplete

Origin-AS validation codes: V valid, I invalid, N not-found, D disabled

   Network            Next Hop            Metric LocPrf Weight Path

V*> 203.0.113.0/24     10.0.12.2                0             0 8100 ?

I*  203.0.113.1/24     10.0.12.2                0             0 8100 ?

N*> 192.168.122.1/32   10.0.12.2                0             0 8100 ?
```

- Invalid : indique que le préfixe remplit l'une des deux conditions suivantes : 1. Il correspond à une ou plusieurs **Autorisations d'origine de route** (ROA), mais il n'y a pas de correspondance ROA où le AS d'origine correspond au AS d'origine sur l'AS-PATH. 2. Il correspond à un ou plusieurs ROA à la longueur minimale spécifiée dans le ROA, mais pour tous les ROA où il correspond à la longueur minimale, il est plus long que la longueur maximale spécifiée. L'AS d'origine n'importe pas pour la condition #2.
- Valid : indique que le préfixe et la paire AS se trouvent dans la table de cache RPKI.
- Not Found : indique que le préfixe ne figure pas parmi les préfixes valides ou non valides.

Cette section traite en détail de chaque préfixe et de son état.

1. 203.0.113.0/24 - Valide

L'homologue eBGP dans AS 8100 a émis cette route et annoncé au noeud Cisco8000. Puisque le système autonome d'origine (8100) correspond au système autonome d'origine dans ROA (reçu du validateur), ce préfixe est marqué comme valide et est installé dans la table de routage du routeur.

```
RP/0/RP0/CPU0:Cisco8000#show bgp rpki table | in "203.0.113.0|Max"

Thu Jan 21 00:21:26.026 UTC

 Network               Maxlen          Origin-AS         Server

 203.0.113.0/24               24       8100              192.168.122.120
```

La route est installée dans la table BGP.

```
RP/0/RP0/CPU0:Cisco8000#show bgp 203.0.113.0/24

Thu Jan 21 05:30:13.858 UTC

BGP routing table entry for 203.0.113.0/24

Versions:

 Process           bRIB/RIB  SendTblVer

 Speaker                 31          31

Last Modified: Jan 21 00:03:33.344 for 05:26:40
```

```
Paths: (1 available, best #1)

 Not advertised to any peer

 Path #1: Received by speaker 0

 Not advertised to any peer

 8100

   10.0.12.2 from 10.0.12.2 (192.168.122.105)

     Origin incomplete, metric 0, localpref 100, valid, external, best, group-best

     Received Path ID 0, Local Path ID 1, version 31

     Origin-AS validity: valid
```

Étant donné qu'il s'agit du meilleur préfixe BGP et qu'il est également valide par RPKI, il est correctement installé dans la table de routage.

```
RP/0/RP0/CPU0:Cisco8000#show route 203.0.113.0/24

Thu Jan 21 00:29:43.667 UTC



Routing entry for 203.0.113.0/24

 Known via "bgp 100", distance 20, metric 0

 Tag 8100, type external

 Installed Jan 21 00:03:33.731 for 00:26:10

 Routing Descriptor Blocks

   10.0.12.2, from 10.0.12.2, BGP external

     Route metric is 0

 No advertising protos.
```

2. 203.0.113.1/24 - Non valide

Ce préfixe n'est pas valide car il y a un conflit dans les informations AS d'origine contenues dans ROA et les informations origine-as reçues via le message BGP de l'homologue eBGP. 203.0.113.1/24 est reçu via BGP avec l'origine AS 8100.

```
RP/0/RP0/CPU0:Cisco8000#show bgp origin-as validity invalid

Thu Jan 21 00:34:38.171 UTC

BGP router identifier 10.1.1.1, local AS number 100

BGP generic scan interval 60 secs

Non-stop routing is enabled
```

```
BGP table state: Active

Table ID: 0xe0000000   RD version: 33

BGP main routing table version 33

BGP NSR Initial initsync version 2 (Reached)

BGP NSR/ISSU Sync-Group versions 0/0

BGP scan interval 60 secs


Status codes: s suppressed, d damped, h history, * valid, > best

              i - internal, r RIB-failure, S stale, N Nexthop-discard

Origin codes: i - IGP, e - EGP, ? - incomplete

  Network            Next Hop            Metric LocPrf Weight Path

*  203.0.113.1/24     10.0.12.2                0             0 8100 ?
```

Cependant, le ROA reçu du validateur montre que ce préfixe appartient à AS 10021.

```
RP/0/RP0/CPU0:Cisco8000#show bgp rpki table 203.0.113.1/24 max 24

Thu Jan 21 00:37:05.615 UTC


RPKI ROA entry for 203.0.113.1/24-24

 Origin-AS: 10021 from 192.168.122.120

Version: 124211
```

Étant donné que les informations d'origine AS dans l'annonce BGP reçue (AS 8100) ne correspondaient pas à l'origine AS réelle reçue dans ROA (AS 10021), le préfixe est marqué comme non valide et n'est pas installé dans la table de routage.

```
RP/0/RP0/CPU0:Cisco8000#show bgp 203.0.113.1/24

Thu Jan 21 05:37:26.714 UTC

BGP routing table entry for 203.0.113.1/24

Versions:

 Process           bRIB/RIB  SendTblVer

 Speaker                 32          32

Last Modified: Jan 21 00:03:33.344 for 05:33:53

Paths: (1 available, no best path)

 Not advertised to any peer

 Path #1: Received by speaker 0

 Not advertised to any peer
```

```
 8100

   10.0.12.2 from 10.0.12.2 (192.168.122.105)

     Origin incomplete, metric 0, localpref 100, valid, external

     Received Path ID 0, Local Path ID 0, version 0

     Origin-AS validity: invalid
```

3. 192.168.122.1/32 Introuvable

Il s'agit d'un préfixe privé qui n'est pas présent dans le cache ROA. BGP a déclaré ce préfixe comme « Not found ».

```
RP/0/RP0/CPU0:Cisco8000#show bgp 192.168.122.1/32

Thu Jan 21 05:44:39.861 UTC

BGP routing table entry for 192.168.122.1/32

Versions:

 Process           bRIB/RIB  SendTblVer

 Speaker                 33          33

Last Modified: Jan 21 00:03:33.344 for 05:41:06

Paths: (1 available, best #1)

 Not advertised to any peer

 Path #1: Received by speaker 0

 Not advertised to any peer

 8100

   10.0.12.2 from 10.0.12.2 (192.168.122.105)

     Origin incomplete, metric 0, localpref 100, valid, external, best, group-best

     Received Path ID 0, Local Path ID 1, version 33

     Origin-AS validity: not-found
```

Puisque RPKI est toujours adopté, les préfixes « introuvable » sont installés dans la table de routage. Sinon, BGP ignorera ces préfixes légitimes qui ne sont pas enregistrés dans la base de données RPKI.

Autoriser le préfixe non valide

Bien que cela ne soit pas recommandé, le logiciel fournit un bouton permettant aux préfixes non valides de participer à l'algorithme de calcul du meilleur chemin.

```
router bgp 100
```

```
 address-family ipv4 unicast

 bgp bestpath origin-as allow invalid

 !
```

Avec cette configuration, le routeur ne considère pas les préfixes non valides pour le meilleur calcul de chemin tandis que Ceci est marqué comme « non valide ». Ce résultat indique « 203.0.113.1/24 » marqué comme le meilleur chemin.

```
RP/0/RP0/CPU0:Cisco8000#show bgp

Thu Jan 21 06:21:34.294 UTC

BGP router identifier 10.1.1.1, local AS number 100

BGP generic scan interval 60 secs

Non-stop routing is enabled

BGP table state: Active

Table ID: 0xe0000000   RD version: 34

BGP main routing table version 34

BGP NSR Initial initsync version 2 (Reached)

BGP NSR/ISSU Sync-Group versions 0/0

BGP scan interval 60 secs



Status codes: s suppressed, d damped, h history, * valid, > best

              i - internal, r RIB-failure, S stale, N Nexthop-discard

Origin codes: i - IGP, e - EGP, ? - incomplete

  Network            Next Hop            Metric LocPrf Weight Path

*> 203.0.113.0/24     10.0.12.2                0             0 8100 ?

*> 203.0.113.1/24     10.0.12.2                0             0 8100 ?

*> 192.168.122.1/32   10.0.12.2                0             0 8100 ?
```

Comme l'illustre ce résultat, le préfixe est marqué comme étant le meilleur, même s'il est maintenu non valide.

```
RP/0/RP0/CPU0:Cisco8000#show bgp 203.0.113.1/24

Thu Jan 21 06:23:26.994 UTC

BGP routing table entry for 203.0.113.1/24

Versions:

 Process           bRIB/RIB  SendTblVer
```

```
 Speaker                 34          34

Last Modified: Jan 21 06:05:31.344 for 00:17:55

Paths: (1 available, best #1)

 Not advertised to any peer

 Path #1: Received by speaker 0

 Not advertised to any peer

 8100

   10.0.12.2 from 10.0.12.2 (192.168.122.105)

     Origin incomplete, metric 0, localpref 100, valid, external, best, group-best

     Received Path ID 0, Local Path ID 1, version 34

     Origin-AS validity: invalid
```

Il est à noter qu'un routeur traite toujours le préfixe non valide comme la dernière option et préfère toujours un préfixe valide à un préfixe non valide s'il est disponible.

Configuration manuelle du ROA sur le routeur

Si, pour une raison quelconque, un ROA pour un préfixe donné n'est pas encore créé, reçu ou est retardé, un ROA manuel peut être configuré sur le routeur. Par exemple, le préfixe « 192.168.122.1/32 » est marqué comme « introuvable », comme illustré ici.

```
RP/0/RP0/CPU0:Cisco8000#show bgp origin-as validity

Thu Jan 21 06:36:31.041 UTC

BGP router identifier 10.1.1.1, local AS number 100

BGP generic scan interval 60 secs

Non-stop routing is enabled

BGP table state: Active

Table ID: 0xe0000000   RD version: 34

BGP main routing table version 34

BGP NSR Initial initsync version 2 (Reached)

BGP NSR/ISSU Sync-Group versions 0/0

BGP scan interval 60 secs



Status codes: s suppressed, d damped, h history, * valid, > best

              i - internal, r RIB-failure, S stale, N Nexthop-discard
```

```
Origin codes: i - IGP, e - EGP, ? - incomplete

Origin-AS validation codes: V valid, I invalid, N not-found, D disabled

   Network            Next Hop            Metric LocPrf Weight Path

V*> 203.0.113.0/24     10.0.12.2                0             0 8100 ?

I*> 203.0.113.1/24     10.0.12.2                0             0 8100 ?

N*> 192.168.122.1/32   10.0.12.2                0             0 8100 ?
```

Un ROA manuel peut être configuré comme indiqué ici. Cette commande associe le préfixe « 192.168.122.1/32 » au système autonome 8100.

```
router bgp 100

 rpki route 192.168.122.1/32 max 32 origin 8100
```

Avec cette configuration, l'état du préfixe passe de « N » à « V ».

```
RP/0/RP0/CPU0:Cisco8000#show bgp origin-as validity

Thu Jan 21 06:36:34.151 UTC

BGP router identifier 10.1.1.1, local AS number 100

BGP generic scan interval 60 secs

Non-stop routing is enabled

BGP table state: Active

Table ID: 0xe0000000   RD version: 35

BGP main routing table version 35

BGP NSR Initial initsync version 2 (Reached)



Status codes: s suppressed, d damped, h history, * valid, > best

              i - internal, r RIB-failure, S stale, N Nexthop-discard

Origin codes: i - IGP, e - EGP, ? - incomplete

Origin-AS validation codes: V valid, I invalid, N not-found, D disabled

   Network            Next Hop            Metric LocPrf Weight Path

V*> 203.0.113.0/24     10.0.12.2                0             0 8100 ?

I*> 203.0.113.1/24     10.0.12.2                0             0 8100 ?

V*> 192.168.122.1/32   10.0.12.2                0             0 8100 ?
```

État de validation de la stratégie de routage et du préfixe

Le résultat de l'état du préfixe peut être utilisé pour créer des stratégies de routage. Ces états peuvent être utilisés dans une instruction de correspondance et des actions souhaitées par l'administrateur peuvent être entreprises. Cet exemple montre comment faire correspondre tous les préfixes avec un état non valide et leur attribuer la valeur de pondération 12345.

```
route-policy Invalid
 if validation-state is invalid then
   set weight 12345
 endif
end-policy
!



router bgp 100
 remote-as 8100
 address-family ipv4 unicast
  route-policy Invalid in
 !
 !
!
```

Ce résultat montre un poids appliqué de préfixe non valide de 12345.

```
RP/0/RP0/CPU0:Cisco8000#show bgp 203.0.113.1/24
Thu Jan 21 06:57:33.816 UTC
BGP routing table entry for 203.0.113.1/24
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                 38          38
Last Modified: Jan 21 06:54:04.344 for 00:03:29
Paths: (1 available, best #1)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Not advertised to any peer
 8100
   10.0.12.2 from 10.0.12.2 (192.168.122.105)
```

```
    Origin incomplete, metric 0, localpref 100, weight 12345, valid, external, best, group-best

    Received Path ID 0, Local Path ID 1, version 38

    Origin-AS validity: invalid
```

Partage des informations de validation de préfixe via la communauté étendue

En tant que routeur BGP peut également partager l'état de validation de préfixe avec d'autres routeurs (sans cache local du validateur) via la communauté étendue BGP. Cela permet d'économiser la charge de chaque routeur du réseau en établissant une session avec le validateur et en téléchargeant tous les ROA.

Ceci est rendu possible par la communauté étendue BGP.

Cette commande permet au routeur de partager des informations de « validation de préfixe » avec des homologues iBGP.

```
router bgp 100

 address-family ipv4 unicast

 bgp origin-as validation signal ibgp
```

Une fois le routeur Cisco 8000 configuré comme indiqué, les mises à jour BGP aux homologues contiennent les informations de validation du préfixe. Dans ce cas, le routeur iBGP voisin est un routeur IOS-XE.

```
csr2#show ip bgp 203.0.113.1/24

BGP routing table entry for 203.0.113.1/24, version 14

Paths: (1 available, best #1, table default)

 Not advertised to any peer

 Refresh Epoch 1

 8100

   10.0.12.2 from 10.0.13.1 (10.1.1.1)

     Origin IGP, metric 0, localpref 100, valid, internal, best

     Extended Community: 0x4300:0:2

     rx pathid: 0, tx pathid: 0x0

     Updated on Jan 21 2021 18:16:56 UTC
```

Ce mappage de communauté étendu peut être compris avec l'utilisation de 0x4300 0x0000 (4 octets indiquant l'état).

Les quatre octets indiquant l'état sont traités comme un entier non signé 32 bits ayant l'une des valeurs suivantes :

- 0 - Valide

- 1 - Introuvable
- 2 - Non valide

La communauté du préfixe 203.0.113.1/24 est 0x4300:0:2, qui correspond au préfixe « Invalid ». De cette manière, le routeur csr2, bien qu'il ne dispose pas de cache local, est toujours en mesure de prendre des décisions en fonction de l'état de validation du préfixe.

L'état de validation du préfixe peut maintenant être utilisé pour correspondre dans une route-map ou dans l'algorithme du meilleur chemin BGP.

Recommandations pour la mise en oeuvre de BGP RPKI

Bonnes pratiques pour la création de ROA

Voici quelques recommandations basées sur les réseaux inaccessibles observés à RPKI-Observatory. L'observatoire RPKI analyse plusieurs aspects du paysage RPKI déployé.

- Si un ROA est créé pour n'importe quel préfixe, alors il est recommandé d'annoncer ce préfixe dans BGP. En l'absence de ROA, quelqu'un d'autre peut l'annoncer en prétendant simplement être un ASN contenu dans ce ROA et en utilisant le préfixe.
- Si ROA est créé avec un maxlen supérieur à la longueur de préfixe, alors il est équivalent à la création de ROA pour tous les préfixes possibles sous le préfixe d'origine jusqu'au maxlen. Il est fortement recommandé d'annoncer tous ces préfixes dans BGP.
- Si un ROA est créé pour un préfixe et que le propriétaire du préfixe annonce un sous-préfixe du préfixe d'origine, le ROA invalide ce sous-préfixe. Un ROA pour le sous-préfixe également ou le maximum du ROA d'origine doit être étendu pour couvrir le sous-préfixe.
- Si une organisation possède un préfixe, mais ne prévoit pas de l'annoncer dans BGP, alors un ROA pour le préfixe pour AS0 doit être créé. Cela invalidera toute annonce de préfixe car AS0 ne peut pas apparaître dans un chemin d'accès AS.
- Si plusieurs ASN proviennent du même préfixe, des ROA pour ce préfixe doivent être créées pour chacun des ASN. Par conséquent, si un routeur a plusieurs ROA pour le même préfixe, une annonce BGP qui correspond à l'un d'entre eux sera valide. Plusieurs ROA pour le même préfixe ne sont pas en conflit.
- Si « A » crée un préfixe pour son client « B » et crée un ROA pour ce préfixe pour le compte de « B », alors « A » doit ajouter le préfixe ASN « B » à l'annonce ou faire en sorte que « B » crée le préfixe lui-même.

Impact de RPKI sur les performances des routeurs XR BGP

Effet de la mise à jour ROA sur le processeur avec la politique de routage

Lorsque les ROA sont mis à jour et si le routeur dispose d'une stratégie de route d'entrée locale pour un voisin qui contient un « état de validation est », il devient important de revalider l'état des préfixes en fonction des nouvelles ROA mises à jour. Pour ce faire, le routeur envoie une requête BGP REFRESH à son homologue.

Lorsque les voisins BGP reçoivent ce message comme indiqué, les voisins envoient à nouveau

leurs préfixes et la route-policy entrante peut revalider les préfixes entrants .

```
Jan 22 18:28:41.360: BGP: 10.0.12.1 rcv message type 5, length (excl. header) 4

Jan 22 18:28:41.360: BGP: 10.0.12.1 rcvd REFRESH_REQ for afi/safi: 1/1, refresh code is 0
```

Le problème s'amplifie lorsque de nombreux voisins sont actualisés en même temps chaque fois que les ROA sont mis à jour. Si la politique de routage entrant de voisinage est complexe et nécessite beaucoup de traitement, le CPU est élevé pendant quelques minutes après une mise à jour ROA. Ces messages REFRESH ne se produisent pas si la route-policy entrante voisine ne contient pas de commande « validation-state is ».

Si « soft-reconfiguration inbound always » est configuré pour un voisin, alors les messages BGP REFRESH ne seront pas envoyés, mais les mêmes politiques de route seront toujours exécutées au même débit et la même utilisation du CPU peut être attendue.

Il est recommandé de préférer l'approche « bgp bestpath origin-as use valid » à la configuration d'une politique de routage pour les raisons expliquées dans la section 6.2.2 ci-dessous.

Minimiser l'impact CPU provoqué par la mise à jour ROA

La meilleure façon d'éviter le problème expliqué ici est d'utiliser **bestpath origin-as use valid** sans **validation-state** dans la politique.

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as use validity

 !
```

Cette commande conserve une route non valide reçue sur le routeur, mais l'empêche de devenir un meilleur chemin. Il ne sera pas installé ni annoncé. C'est aussi bien que de le laisser tomber. Si, avec la prochaine mise à jour ROA, elle devient valide, aucune ACTUALISATION n'est requise et elle devient automatiquement éligible pour le meilleur chemin sans aucune exécution de stratégie nécessaire.

Si l'utilisateur préfère autoriser les préfixes « invalides » et ne pas les utiliser, alors en plus de **bestpath origin-as use valid**, utilisez la configuration **best path origin-as allow invalid**.

Dans ce cas, lorsqu'un ROA change, le meilleur chemin est automatiquement mis à jour sans nécessiter de message REFRESH. Afin de dépréférer, une route signifie que pendant la sélection de route BGP, le chemin non valide RPKI est considéré comme moins préférable que tout autre chemin vers la même destination. Elle est similaire à l'attribution d'une pondération ou d'une préférence locale inférieure à 0.

Le nombre d'invalides RPKI est relativement faible et le fait qu'ils soient conservés dans le tableau n'a pas d'impact significatif sur les ressources.

> **Remarque** : pour utiliser « bestpath origin-as use valid », tous les chemins d'une route, y compris les chemins IBGP, doivent avoir la validité RPKI correcte. Si ce n'est pas le cas, le

test de l'état de validation dans la politique de routage peut encore être utilisé.

Les routes IBGP ne sont pas validées par le routeur sur la base de données ROA. Les routes IBGP obtiennent une validité RPKI de la communauté étendue RPKI. Si la route IBGP est reçue sans cette communauté étendue, son état de validation est défini sur not-found.

Empreinte mémoire RPKI BGP

Chaque ROA consomme de la mémoire pour l'index et les données. Si deux ROA sont pour le même préfixe IP, mais ont un max_len différent ou sont reçus de serveurs RPKI différents, alors ils partagent le même index mais ont des données séparées. Les besoins en mémoire peuvent varier car la surcharge mémoire n'est pas constante. Un budget excédentaire de 10 % est recommandé. Les plates-formes 64 bits nécessitent plus de mémoire pour chaque objet mémoire que les plates-formes 32 bits. L'utilisation de la mémoire IOS-XR en octets pour un objet d'index et un objet de données figure dans la table. Certains frais généraux, pour la plupart constants, sont inclus dans les chiffres.

	Plate-forme 32 bits (octets)	Plate-forme 64 bits (octets)
index IPv4	74	111
Index IPv6	86	125
données	34	53

Cette section présente deux scénarios pour expliquer comment les ROA consomment la mémoire.

Scénario 1. Trois serveurs RPKI configurés sur le routeur

Prenons l'exemple d'un routeur utilisant 3 serveurs RPKI, chacun fournissant 200 000 ROA IPv4 et 20 000 ROA IPv6 sur un processeur de routage 64 bits. Cette mémoire est nécessaire :

20000 * (125 + 3*53) + 200000 * (111 + 3*53) octets = 59,68 millions d'octets

Lors du calcul de la mémoire, le ROA pour le même préfixe de trois validateurs différents partageait la même valeur d'index.

Scénario 2. Serveurs RPKI uniques configurés sur le routeur

Mémoire de traitement BGP sans ROA :

```
RP/0/RP0/CPU0:Cisco8000#show processes memory detail location 0/RP0/CPU0 | in $

Fri Jan 22 17:19:57.945 UTC

JID        Text       Data       Stack      Dynamic    Dyn-Limit  Shm-Tot    Phy-Tot
 Process

1069          2M        71M       132K        25M      7447M        50M        74M        bgp
```

```
RP/0/RP0/CPU0:Cisco8000#show bgp rpki server summary

Fri Jan 22 17:12:09.073 UTC

Hostname/Address    Transport     State         Time          ROAs (IPv4/IPv6)

192.168.122.120     TCP:3323      NONE          00:00:25      N/A
```

Le processus BGP consomme 25 Mo de mémoire sans ROA.

Mémoire de processus BGP avec ROA :

```
RP/0/RP0/CPU0:Cisco8000#show bgp rpki server summary

Fri Jan 22 17:23:46.769 UTC

Hostname/Address    Transport     State         Time          ROAs (IPv4/IPv6)

192.168.122.120     TCP:3323      ESTAB         00:02:42      172796/28411

RP/0/RP0/CPU0:Cisco8000#show processes memory detail location 0/RP0/CPU0 | in $

Fri Jan 22 17:24:14.659 UTC

JID        Text       Data       Stack      Dynamic    Dyn-Limit  Shm-Tot    Phy-Tot     Process

1069          2M        99M       132K        53M      7447M        50M       102M       bgp
```

Le processus BGP consomme 25 Mo de mémoire sans ROA.

Mémoire de processus BGP avec ROA :

```
RP/0/RP0/CPU0:Cisco8000#show bgp rpki server summary

Fri Jan 22 17:23:46.769 UTC

Hostname/Address    Transport     State         Time          ROAs (IPv4/IPv6)

192.168.122.120     TCP:3323      ESTAB         00:02:42      172796/28411

RP/0/RP0/CPU0:Cisco8000#show processes memory detail location 0/RP0/CPU0 | in $

Fri Jan 22 17:24:14.659 UTC

JID        Text       Data       Stack      Dynamic    Dyn-Limit  Shm-Tot    Phy-Tot     Process

1069          2M        99M       132K        53M      7447M        50M       102M       bgp
```

Le routeur Cisco 8000 exécute le système d'exploitation 64 bits. Il a reçu 172796 ROA IPv4 et 28411 ROA.

Mémoire (octets) = 172 796 x [111 (index) + 53 (données)] + 28411 x [125 (index) + 53 (données)].

Ces calculs donnent ~27 Mo, ce qui correspond approximativement à l'incrément noté sur la mémoire du routeur ci-dessus.

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.