Livre blanc BGP RPKI avec XR7 (Cisco8000)

Contenu

Introduction

Ce document décrit la fonctionnalité RPKI (Resource Public Key Infrastructure) du protocole BGP

(Border Gateway Protocol) sur la plate-forme Cisco IOS® XR et comment elle empêche les réseaux activés BGP d'être attaqués/interrompus en raison d'annonces BGP incorrectes et illégitimes.

Informations générales

Préface

Ce document discute de la fonctionnalité RPKI BGP et comment elle protège les routeurs exécutant BGP contre les mises à jour de préfixe BGP fausses/malveillantes.

Portée

Ce document utilise Cisco 8000 exécutant 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'appliqueront à d'autres plates-formes Cisco (exécutant Cisco IOS, Cisco IOS-XE, etc.) avec des conversions 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.

Exclusion de responsabilité

Les adresses IP (Internet Protocol) utilisées dans ce document ne sont pas destinées à être des adresses réelles. Tous les exemples, les résultats de l'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 illustratif n'est pas intentionnelle et coïncidente.

Problèmes BGP dus à une mauvaise annonce de préfixe

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

Cette limitation du protocole BGP en fait un candidat facile pour différents types d'attaques. Une attaque courante est appelée ‘piratage de route’. Cette attaque peut être utilisée pour :

- Volez les adresses IP pour envoyer du spam, ce qui entraîne la mise sur liste noire des adresses IP et donc le refus de service.
- espionner le trafic pour obtenir des informations sensibles comme des mots de passe.
- Interruptions dues à des configurations incorrectes par l'administrateur.
- Empêchez la livraison du trafic en configurant des serveurs falsifiés qui provoquent un déni de service.

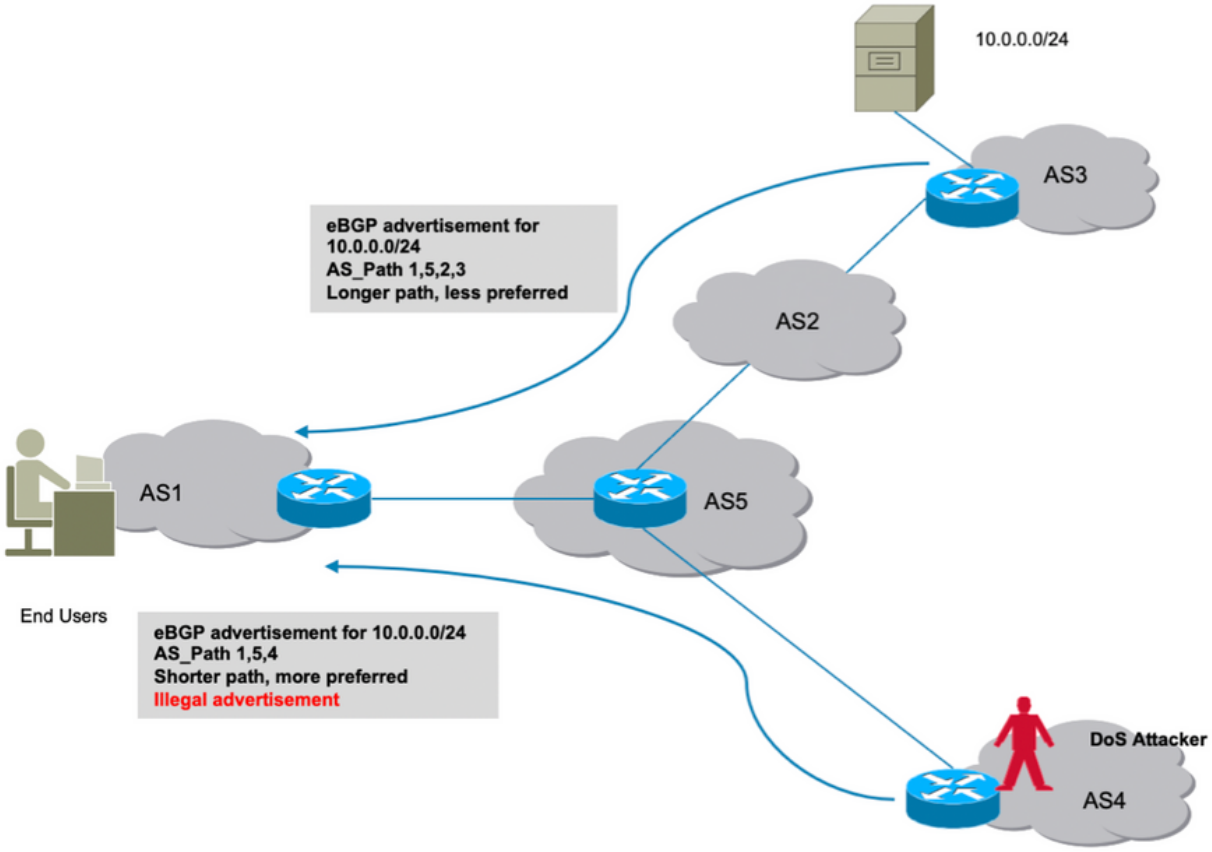
L'attaque par déni de service (communément appelée DoS) est une tentative malveillante de

perturber le trafic normal vers un routeur, un commutateur, un serveur, etc. Il existe diverses attaques par déni de service et peu sont abordées ici.

Détournement 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. Selon la conception de BGP, il n'y a rien dans BGP qui empêcherait un attaquant d'annoncer le même préfixe à Internet.

Comme indiqué, le pirate dans 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 risque d'être bloqué, ce qui entraînera un déni de service aux clients finaux.

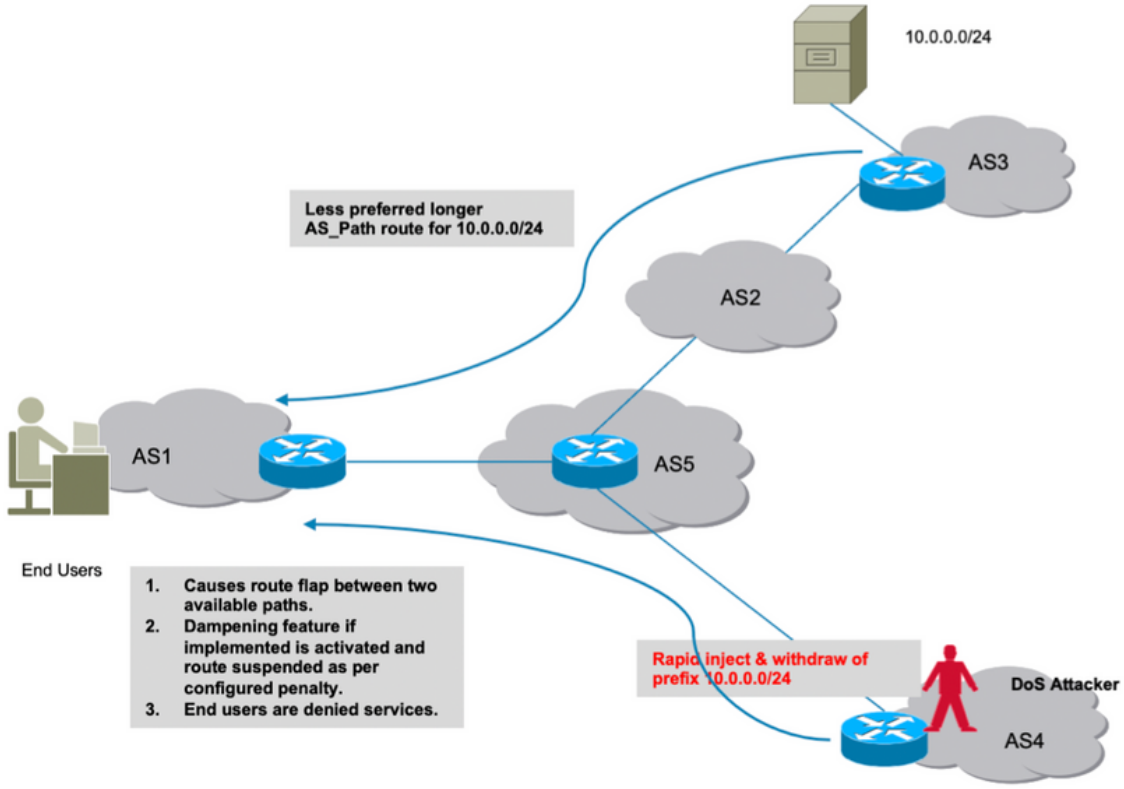


Détection de route

Dégrader les performances du système

Cette section traite d'une autre manière de refuser des services. Si la fonctionnalité de blocage de route BGP de Cisco est configurée, elle peut être exploitée si le pirate introduit des failles de route rapides dans le réseau causant une interruption constante.

La fonction d'amortissement continuera d'imposer des pénalités sur la route légitime et la rendra indisponible pour le trafic réel. En outre, ce type de volets induits de manière non éthique entraîne des contraintes sur les ressources du routeur telles que le processeur, la mémoire, etc.

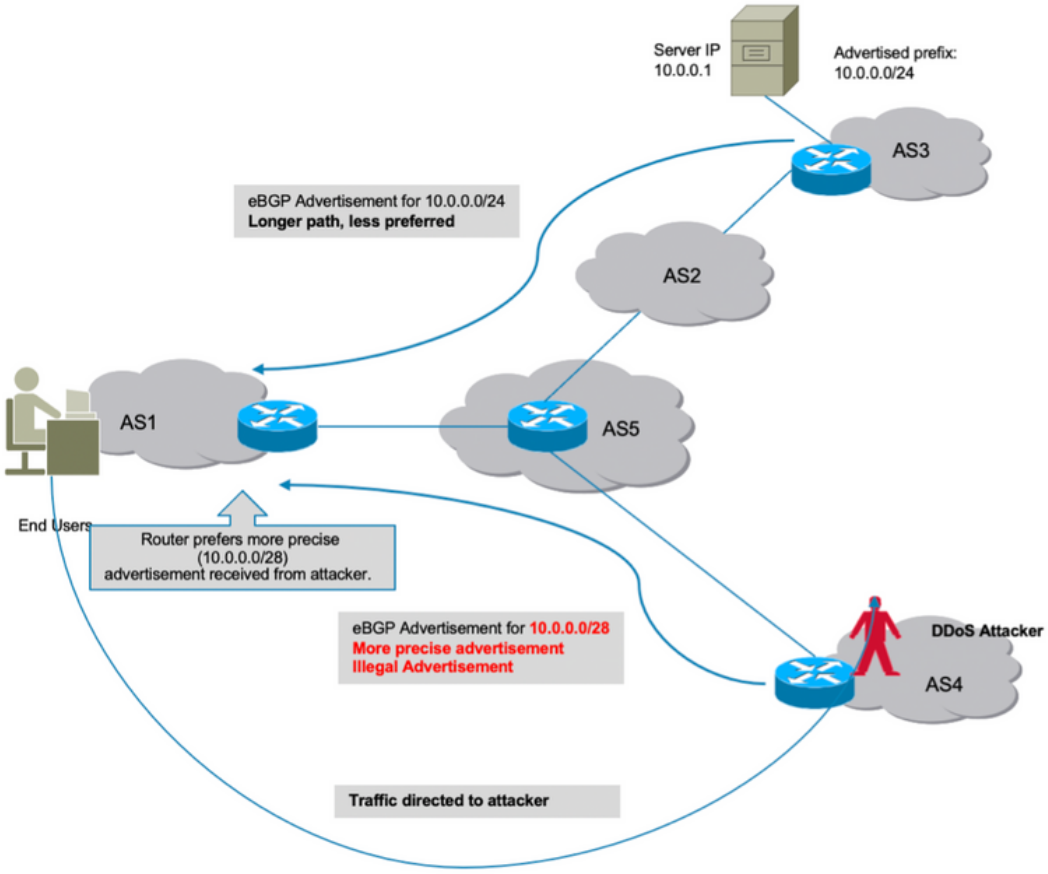


Atténuation de la route

Détournement de sous-préfixe

Comme nous l'avons vu dans la section précédente, comment un pirate peut créer illégalement un préfixe et provoquer un blackholing du trafic. Malheureusement, le blackholing n'est pas la seule cause d'inquiétude. Dans de telles attaques, les données réelles peuvent être compromises par un pirate qui peut analyser les données reçues à des fins non éthiques.

De même, le piratage d'une route pourrait se faire en annonçant illégalement une route plus précise. BGP préfère les préfixes qui correspondent plus longtemps et ce comportement peut être exploité à tort comme le montre l'image.



Détection de sous-préfixe

Toutes les attaques abordées 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. À chaque réception d'une nouvelle annonce, le routeur devient alors capable de vérifier les informations d'origine AS du préfixe reçues de l'homologue BGP avec ses informations de base de données locales du validateur.

Ainsi, le routeur est capable de distinguer les bonnes annonces des mauvaises (illégales) et la capacité d'éviter toutes les attaques évoquées ci-dessus est intrinsèquement ajoutée sur le routeur. BGP RPKI fournit la source d'informations de confiance requise.

RPKI

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

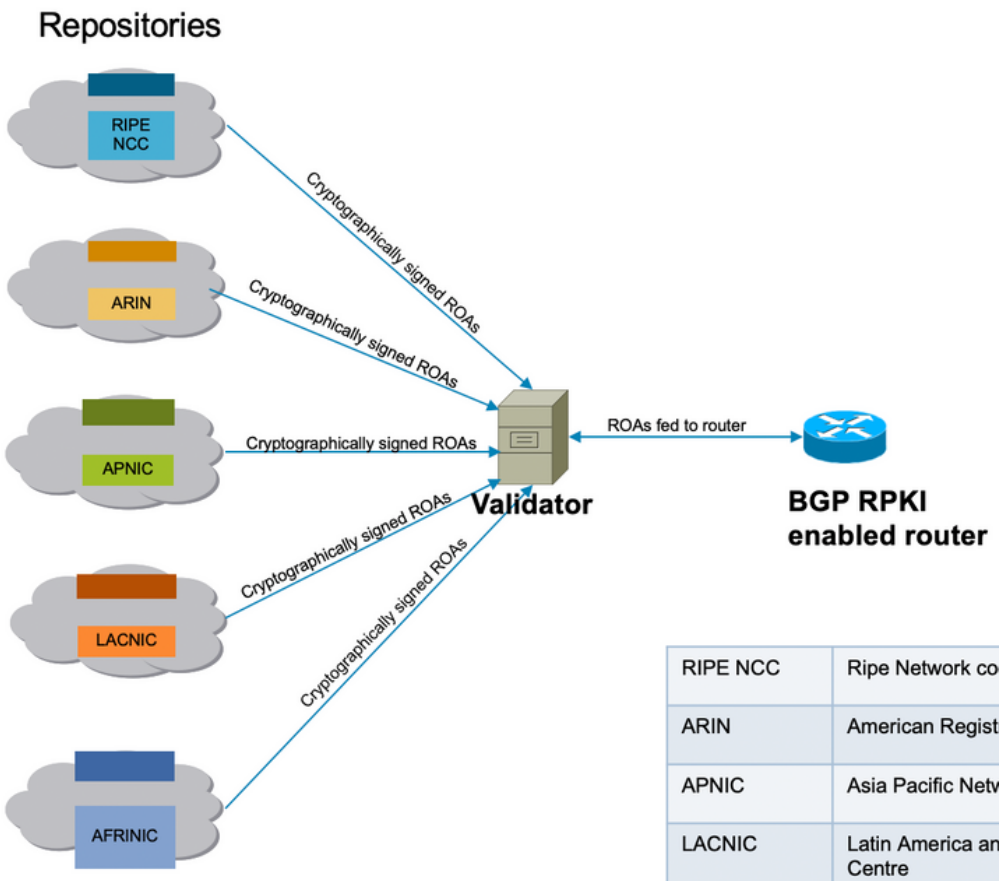
Les cinq registres Internet régionaux (RIR) sont les points d'ancrage de l'ICPR. L'IANA (Internet Assigned Numbers Authority) est la partie supérieure de l'arborescence qui distribue les préfixes IP. Les RIR sont les suivants dans la hiérarchie. Ils attribuent des sous-préfixes aux registres Internet locaux (LIR) et aux grands fournisseurs d'accès à Internet (FAI). Ils signent un certificat pour ces préfixes. Le niveau suivant alloue des sous-préfixes de ces derniers 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, l'ICP forme une arborescence de certificats qui reflète l'arborescence des allocations d'adresses IP. Les validateurs RPKI appartenant aux parties de confiance interrogent tous les points de publication pour trouver des certificats et des ROA mis à jour (ainsi que des LCR et des manifestes). Ils commencent aux points d'ancrage de la fiducie et suivent les liens vers les points de publication des certificats enfants.

Les ROA sont entrées dans le référentiel par le biais des RIR, mais il en est de même pour d'autres registres (nationaux ou locaux). Cette responsabilité peut également être déléguée aux FAI avec une supervision et une vérification appropriées par les RIR.

Àl'heure actuelle, il existe cinq référentiels 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 de confiance pour créer son cache. Il s'agit d'une copie combinée de l'ICP, qui est régulièrement récupérée/actualisée directement ou indirectement à partir de l'ICP globale. Le validateur transmet ensuite ces informations aux routeurs, leur permettant 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 IP

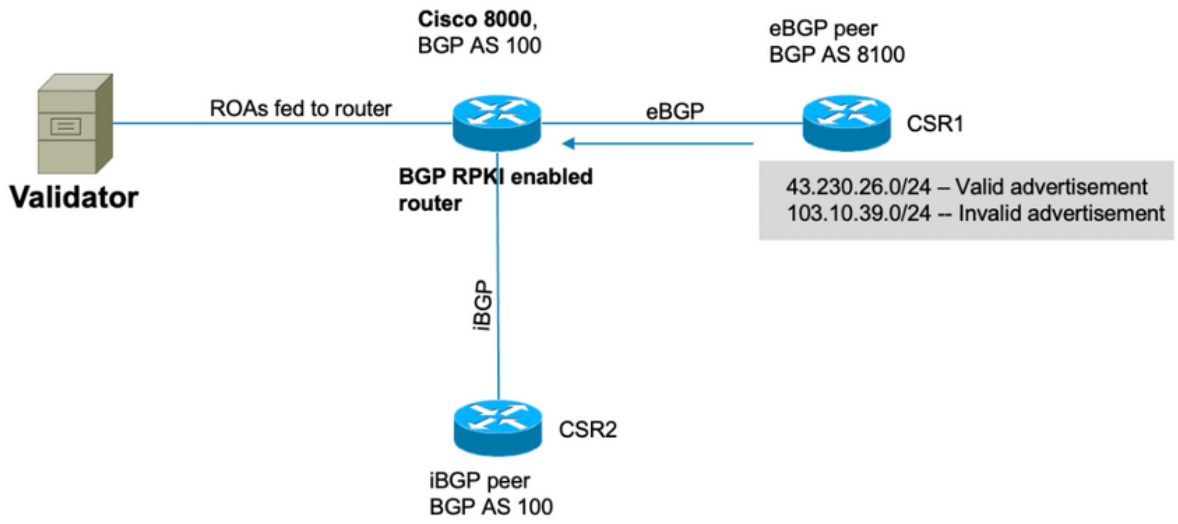
192.168.122.120 et son port 3323.

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

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 mis à jour.

Démonstration RPKI BGP

Topologie



Topologie

> **Remarque:** Cette démonstration utilise un numéro de système autonome public aléatoire et des préfixes simplement pour expliquer les mécanismes RPKI BGP. Les adresses IP publiques sont utilisées en raison de l'ICP principalement pour la protection des préfixes publics et toutes les adresses de retour d'accès créées sur les RIR sont des préfixes publics. Enfin, aucune des actions, configurations, etc. décrites dans ce document n'affecte ces adresses IP publiques et AS de quelque manière que ce soit.

Configuration

```
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 de retour sur investissement au routeur. Ce cache est actualisé à intervalles réguliers afin de minimiser la possibilité que le routeur conserve des informations périmées. Dans cette démonstration, une durée de rafraîchissement de 900 secondes a été configurée. Comme indiqué ici, le routeur Cisco 8000 a téléchargé les ROA IPv4 et 28350 IPv6 172632 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

 1.0.0.0/24            24              13335             192.168.122.120

 1.0.4.0/22            22              38803             192.168.122.120

 1.0.4.0/24            24              38803             192.168.122.120

 1.0.5.0/24            24              38803             192.168.122.120

 1.0.6.0/24            24              38803             192.168.122.120

 1.0.7.0/24            24              38803             192.168.122.120

 1.1.1.0/24            24              13335             192.168.122.120

 1.1.4.0/22            22              4134              192.168.122.120

 1.1.16.0/20           20              4134              192.168.122.120

 1.2.9.0/24            24              4134              192.168.122.120

 1.2.10.0/24           24              4134              192.168.122.120

 1.2.11.0/24           24              4134              192.168.122.120

 1.2.12.0/22           22              4134              192.168.122.120

 1.3.0.0/16            16              4134              192.168.122.120
```

```
 1.6.0.0/22            24              9583              192.168.122.120
```

Vérification

Cette section explique comment BGP RPKI est actif et comment il empêche le routeur de publier des annonces erronées ou illégales.

Activation de la valeur Origine-As

Par défaut, le routeur récupère les ROA du validateur, mais ne commence pas à les utiliser avant d'être configuré pour le faire. 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*> 43.230.26.0/24     10.0.12.2                0             0 8100 ?

D*> 103.10.39.0/24     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-original, 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 de retour d'accès 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*> 43.230.26.0/24     10.0.12.2                0             0 8100 ?

I*  103.10.39.0/24     10.0.12.2                0             0 8100 ?

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

Pour permettre au routeur d'utiliser les informations d'état de validation du 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 vous permet néanmoins de les utiliser dans les stratégies de routage qui sont abordées 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*> 43.230.26.0/24     10.0.12.2                0             0 8100 ?

I*  103.10.39.0/24     10.0.12.2                0             0 8100 ?

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

- Non valide - 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 ROA correspondant où l'AS d'origine correspond à l'AS d'origine sur l'AS-PATH. 2. Il correspond à une ou plusieurs des valeurs de la longueur minimale spécifiée dans la valeur de référence, mais pour toutes les valeurs de la longueur minimale, elle est supérieure à la longueur maximale spécifiée. L'AS d'origine n'a pas d'importance pour la condition n° 2.
- Valide : indique que le préfixe et la paire AS se trouvent dans la table de cache RPKI.
- Non trouvé - 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. 43.230.26.0/24 - Valide

L'homologue eBGP de l'AS 8100 a lancé cette route et annoncé au noeud Cisco8000. Puisque l'AS d'origine (8100) correspond à l'AS 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 "43.230.26.0|Max"

Thu Jan 21 00:21:26.026 UTC

 Network               Maxlen          Origin-AS         Server

 43.230.26.0/24                24       8100              192.168.122.120
```

La route est installée dans la table BGP.

```
RP/0/RP0/CPU0:Cisco8000#show bgp 43.230.26.0/24

Thu Jan 21 05:30:13.858 UTC

BGP routing table entry for 43.230.26.0/24

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
```

Comme 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 43.230.26.0/24

Thu Jan 21 00:29:43.667 UTC



Routing entry for 43.230.26.0/24

 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. 103.10.39.0/24 - 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. 103.10.39.0/24 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

*  103.10.39.0/24     10.0.12.2                0             0 8100 ?
```

Cependant, la ROA reçue du validateur indique que ce préfixe appartient à AS 10021.

```
RP/0/RP0/CPU0:Cisco8000#show bgp rpki table 103.10.39.0/24 max 24

Thu Jan 21 00:37:05.615 UTC


RPKI ROA entry for 103.10.39.0/24-24

 Origin-AS: 10021 from 192.168.122.120

Version: 124211
```

Puisque 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 103.10.39.0/24

Thu Jan 21 05:37:26.714 UTC

BGP routing table entry for 103.10.39.0/24

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 et n'est pas présent dans le cache ROA. Le protocole BGP a déclaré ce préfixe « introuvable ».

```
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
```

Comme l'ICR est toujours en cours d'adoption, les préfixes 'introuvable' sont installés dans la table de routage. Si vous faites autrement, BGP ignorera ces préfixes légitimes qui ne sont pas enregistrés dans la base de données RPKI.

Autoriser un préfixe non valide

Bien qu'il 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 calcul du meilleur chemin tout en les conservant marqués comme non valides. Ce résultat montre « 103.10.39.0/24" 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

*> 43.230.26.0/24     10.0.12.2                0             0 8100 ?

*> 103.10.39.0/24     10.0.12.2                0             0 8100 ?

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

Comme le montre cette sortie, le préfixe est marqué comme étant le meilleur, même s'il n'est pas valide.

```
RP/0/RP0/CPU0:Cisco8000#show bgp 103.10.39.0/24

Thu Jan 21 06:23:26.994 UTC

BGP routing table entry for 103.10.39.0/24

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 ROA manuelle 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 « Non trouvé », comme indiqué 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*> 43.230.26.0/24     10.0.12.2                0             0 8100 ?

I*> 103.10.39.0/24     10.0.12.2                0             0 8100 ?

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

Il est possible de configurer un ROA manuel comme indiqué ici. Cette commande associe le préfixe '192.168.122.1/32' à AS 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*> 43.230.26.0/24     10.0.12.2                0             0 8100 ?

I*> 103.10.39.0/24     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 route. Ces états peuvent être utilisés dans une instruction match et les actions souhaitées par l'administrateur

peuvent être prises. L'exemple suivant montre comment associer tous les préfixes à un état non valide et définir une valeur de poids de 12345 pour eux.

```
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 103.10.39.0/24
Thu Jan 21 06:57:33.816 UTC
BGP routing table entry for 103.10.39.0/24
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
```

Partager les informations de validation du préfixe via la communauté étendue

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

Ceci est rendu possible en utilisant 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 des homologues contiennent les informations de validation du préfixe. Dans ce cas, le routeur voisin iBGP est un routeur IOS-XE.

```
csr2#show ip bgp 103.10.39.0/24

BGP routing table entry for 103.10.39.0/24, 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 de 32 bits ayant l'une des valeurs suivantes :

- 0 - Valide
- 1 - Non trouvé

- 2 - Non valide

La communauté du préfixe 103.10.39.0/24 est 0x4300:0:2 qui correspond au préfixe 'Non valide'. De cette manière, le routeur csr2, bien qu'il n'ait pas de cache local propre, peut toujours 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 RPKI BGP

Bonnes pratiques pour la création de ROA

Voici quelques recommandations basées sur des réseaux inaccessibles observés à l'Observatoire RPKI. L'Observatoire de l'ICP analyse plusieurs aspects du paysage de l'ICP déployé.

- Si un ROA est créé pour un préfixe, il est recommandé d'annoncer ce préfixe dans BGP. En l'absence de cette réponse, quelqu'un d'autre peut l'annoncer simplement en prétendant être ASN contenu dans cette ROA et en utilisant le préfixe.
- Si ROA est créé avec un maxlen supérieur à la longueur du préfixe, cela équivaut à créer des 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 une ROA est créée pour un préfixe et que le propriétaire du préfixe annonce un sous-préfixe du préfixe d'origine, la ROA invalide ce sous-préfixe. Un ROA pour le sous-préfixe ou le maxlen de la ROA d'origine doit être étendu pour couvrir le sous-préfixe.
- Si une organisation possède un préfixe, mais prévoit de ne pas l'annoncer dans BGP, alors une ROA pour le préfixe pour AS0 doit être créée. Cela invalidera toute annonce de préfixe, car AS0 ne peut pas apparaître dans un chemin AS.
- Si plusieurs ASN sont originaires du même préfixe, des ROA pour ce préfixe doivent être créés 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'eux sera valide. Plusieurs ROA pour le même préfixe ne sont pas en conflit.
- Si « A » est à l'origine d'un préfixe pour son client « B » et que vous créez un ROA pour ce préfixe au nom de « B », alors « A » doit remplacer « B » ASN à l'annonce ou faire en sorte que le 'B' soit à l'origine du préfixe lui-même.

Impact sur les performances de RPKI sur les routeurs XR BGP

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

Lorsque les ROA sont mis à jour et que le routeur dispose d'une stratégie de route d'entrée locale pour un voisin qui contient un état de validation, 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 politique de route 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 actualisent en même temps chaque fois que les ROA sont mis à jour. Si la politique de routage entrante du voisin est complexe et nécessite beaucoup de traitement, alors les résultats élevés du CPU 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 “ reconfiguration logicielle entrante toujours ” est configurée pour un voisin, alors les messages BGP REFRESH ne seront pas envoyés, mais les mêmes stratégies de route seront toujours exécutées au même rythme et l'utilisation du CPU peut être attendue.

Il est recommandé de préférer l'approche « bgp bestpath origine-as use validité » plutôt que de configurer une stratégie de route pour les raisons expliquées au paragraphe 6.2.2 ci-dessous.

Minimiser l'impact du processeur causé par la mise à jour ROA

La meilleure façon d'éviter le problème expliqué ici est d'utiliser **bestpath origine-as use validité** sans **validation-state est** 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 la prochaine mise à jour ROA devient valide, aucune REFRESH n'est requise et elle deviendra automatiquement éligible au meilleur chemin sans exécution de stratégie nécessaire.

Si l'utilisateur préfère autoriser les préfixes « non valides » et ne pas les utiliser, alors en plus de **bestpath origine-as use value**, utilisez la configuration **best path origination-as allow Invalid**.

Dans ce cas, lorsqu'une ROA change, le meilleur chemin est automatiquement mis à jour sans nécessiter de message REFRESH. Afin de désapprouver, une route signifie que, lors de la sélection de la route BGP, le chemin non valide RPKI est considéré comme moins préférable que tout autre chemin vers la même destination. Il est similaire à l'affectation de son poids ou préférence locale inférieure à 0.

Le nombre d'invalides de l'ICP est relativement faible et le fait de les garder dans le tableau n'a pas d'incidence importante sur les ressources.

> **Remarque:** Pour utiliser la « validité d'utilisation du meilleur chemin d'origine », 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 validation-state dans route-policy peut toujours être utilisé.

Les routes IBGP ne sont pas validées par le routeur par rapport à 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 introuvable.

Empreinte mémoire RPKI BGP

Chaque ROA consomme de la mémoire pour l'index et les données. Si deux ROA correspondent au 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 distinctes. Les besoins en mémoire peuvent varier, car la surcharge de mémoire n'est pas constante. Un surbudget de 10 % est recommandé. Les plates-formes 64 bits nécessitent plus de mémoire pour chaque objet mémoire que les plates-formes 32 bits. Utilisation de la mémoire IOS-XR en octets pour un objet d'index et un objet de données se trouve dans le tableau suivant. Certains frais généraux 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

Considérez un routeur utilisant 3 serveurs RPKI, chacun fournissant 200 000 ROA IPv4 et 20 000 ROA IPv6 sur un processeur de routage 64 bits nécessitant cette mémoire :

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

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

Scénario 2. Serveurs RPKI uniques configurés sur le routeur

Mémoire de processus 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 un système d'exploitation 64 bits. Il a reçu 172796 ROA IPv4 et 28411 ROA.

Mémoire (octets) = 172 796 x [111 (index) + 53 (données)] + 28 411 x [125 (index) + 53 (données)].

Ces calculs donnent ~27 Mo, soit approximativement l'incrément observé sur la mémoire du routeur ci-dessus.