Entenda o BGP RPKI com o white paper XR7 Cisco8000

Contents

Introduction

Este documento descreve o recurso de infraestrutura de chave pública (RPKI - Public Key Infrastructure) do protocolo de gateway de borda (BGP - Border Gateway Protocol) na plataforma Cisco IOS® XR.

Informações de Apoio

Prefácio

Este documento discute o recurso BGP RPKI e como ele protege o BGP com roteadores contra atualizações de prefixo BGP falsas/maliciosas.

Escopo

Este documento usa o Cisco 8000 com a versão XR 7.3.1 para demonstração. No entanto, o BGP RPKI é um recurso independente de plataforma, os conceitos discutidos neste documento aplicam-se a outras plataformas Cisco (com Cisco IOS, Cisco IOS-XE .) com conversões de CLI equivalentes apropriadas. Este documento não abrange o procedimento para adicionar as autorizações de origem de rota (ROAs) nos registros regionais da Internet.

Prerequisites

O leitor precisa conhecer o protocolo BGP.

Ressalva

Os endereços IP usados neste documento não devem ser endereços reais. Quaisquer exemplos, saída de exibição de comandos e figuras incluídas no documento são mostrados apenas para fins ilustrativos. Qualquer uso de endereços IP reais no conteúdo ilustrativo não é intencional e é coincidente.

Problemas de BGP devido ao anúncio de prefixo incorreto

O BGP serve como o backbone do tráfego da Internet. Embora seja o componente mais importante do núcleo da Internet, ele não tem a capacidade de verificar se o anúncio de BGP de entrada se originou de um sistema autônomo autorizado ou não.

Essa limitação do BGP o torna um candidato fácil para vários tipos de ataques. Um ataque comum é chamado de "sequestro de rota". Esse ataque pode ser utilizado para:

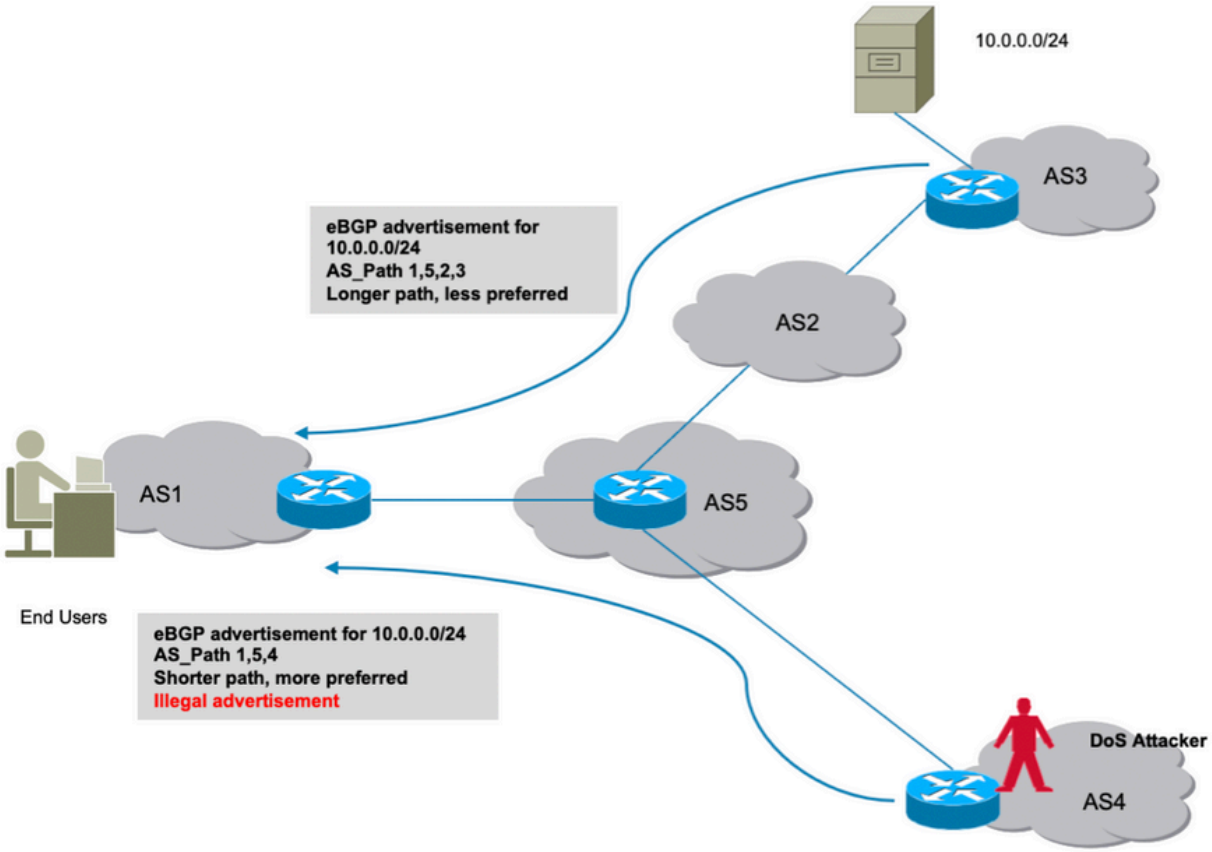
- Roubar IPs para enviar resultados de spam em IP é rejeitado e, portanto, negação de serviço.
- Espionar o tráfego para obter informações confidenciais, como senhas.
- Interrupções causadas por configurações incorretas feitas pelo administrador.
- Evite a entrega de tráfego com servidores falsos, o que resulta em negação de serviço.

O ataque de negação de serviço (comumente conhecido como DoS) é uma tentativa mal-intencionada de interromper o tráfego normal para um roteador, switch, servidor e assim por diante. Há uma variedade de ataques DoS e poucos são discutidos aqui.

Sequestro de rota

Considere o cenário mostrado aqui. O Sistema Autônomo 3 (AS3) envia anúncio legal de BGP para seu prefixo 10.0.0.0/24. Pelo design do BGP, não há nada no BGP que impeça um invasor de anunciar o mesmo prefixo para a Internet.

Como mostrado, o invasor no AS4 anuncia o mesmo prefixo 10.0.0.0/24. O algoritmo de melhor caminho BGP prefere um caminho com AS_Path mais curto. AS_Path 1,5,4 vence sobre caminho mais longo via AS 1,5,2,3. Portanto, o tráfego dos clientes agora será redirecionado para o ambiente do invasor e pode ficar em branco, o que resulta em negação de serviço para os clientes finais.

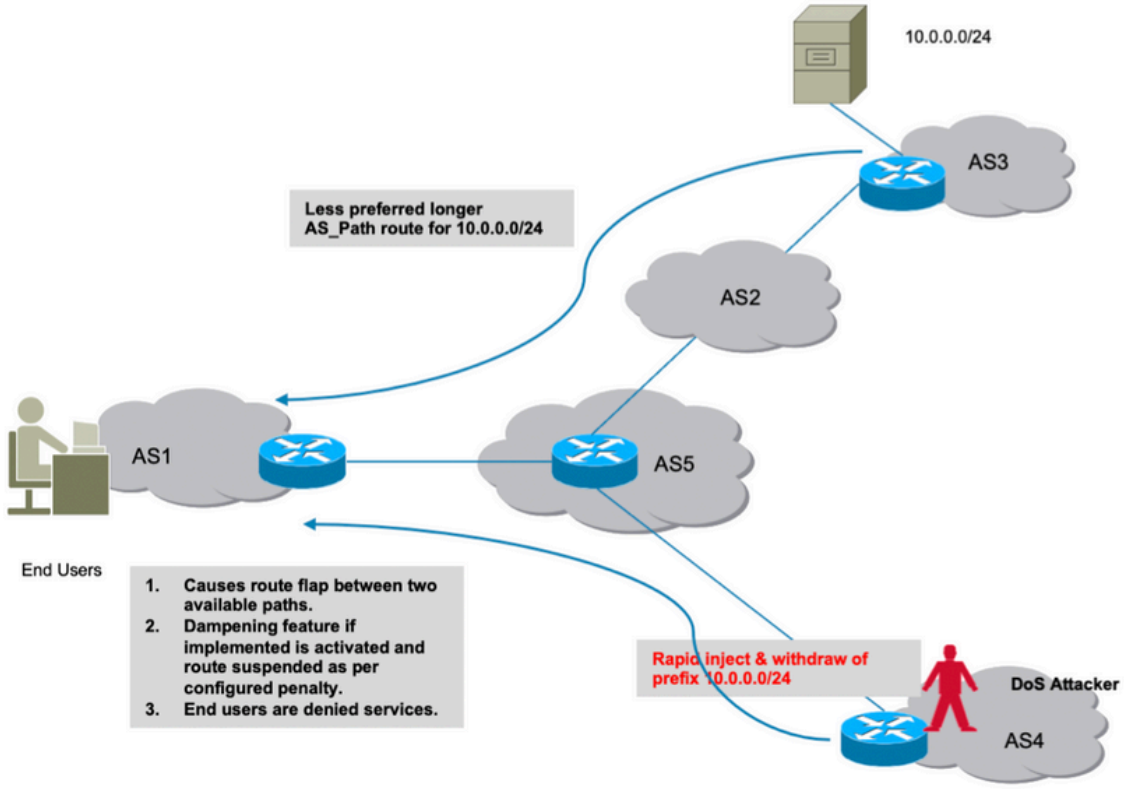


Sequestro de rota

Degradar o desempenho do sistema

Esta seção aborda outra maneira pela qual os serviços podem ser negados. Se o recurso de redução de rota BGP da Cisco estiver configurado, ele poderá ser explorado se o invasor introduzir oscilações de rota rápidas na rede, causando uma rotatividade constante.

O recurso de amortecimento imporá penalidades na rota legítima e a tornará indisponível para o tráfego real. Além disso, esse tipo de oscilação induzida de forma antiética causará sobrecarga nos recursos do roteador, como CPU, memória e assim por diante.

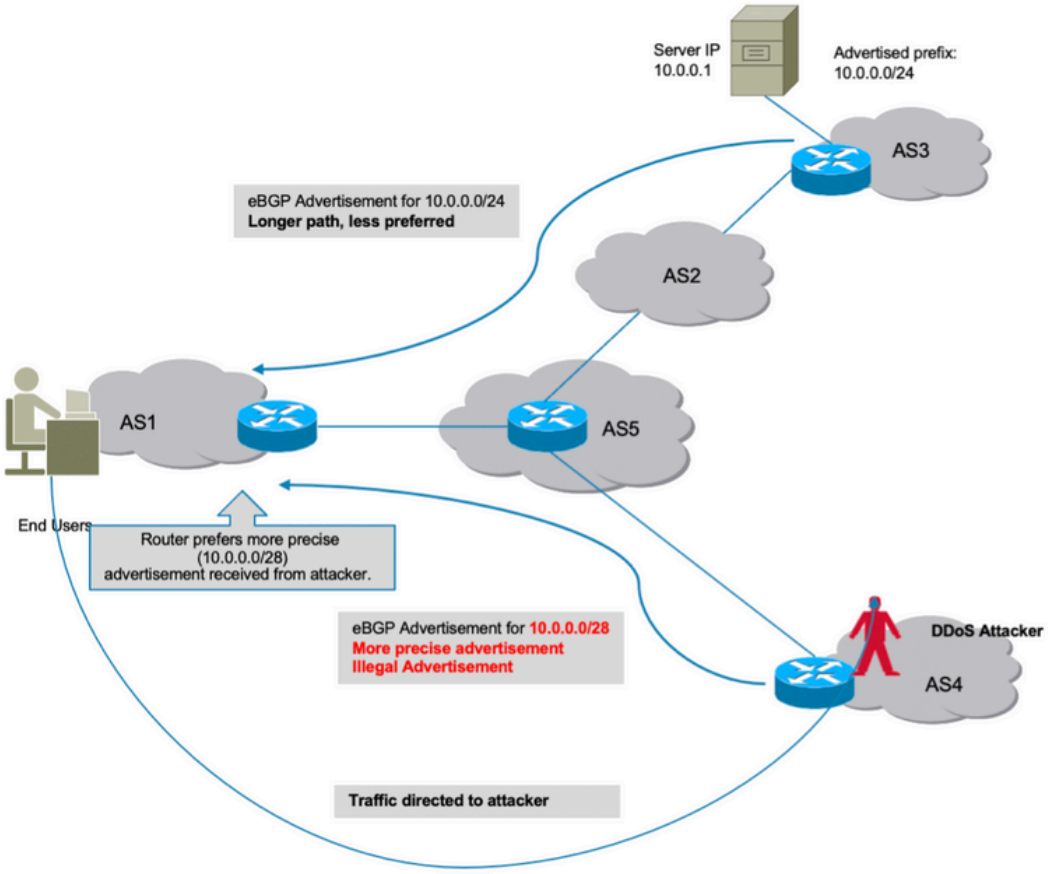


Redução de Rota

Hijacking de subprefixo

Conforme discutido na seção anterior, como um invasor pode originar um prefixo ilegalmente e causar uma interrupção de tráfego. Infelizmente, uma ruptura não é a única causa de preocupação. Nesses ataques, os dados reais podem ser comprometidos, em que um invasor pode verificar os dados recebidos em busca de uso antiético.

Da mesma forma, o sequestro de uma rota poderia ser feito através da publicidade ilegal de uma rota mais precisa. O BGP prefere prefixos que são uma correspondência mais longa e esse comportamento pode ser explorado incorretamente como mostrado na imagem.



Hijack de subprefixo

Todos os ataques discutidos derivam do fato de que o BGP não conseguiu identificar se o AS de origem desses prefixos mal-intencionados era válido ou não. Para corrigir isso, é necessária uma fonte de dados "verdadeira" e "confiável" que um roteador possa manter em seu banco de dados. Em seguida, a cada recebimento de um novo anúncio, o roteador agora se torna capaz de verificar cruzadamente as informações de origem do AS do prefixo recebidas do par BGP com suas informações de banco de dados local do validador.

Assim, o roteador é capaz de distinguir os bons anúncios dos ruins (ilegais) e a capacidade de evitar todos os ataques discutidos anteriormente é inerentemente adicionada ao roteador. O BGP RPKI fornece a fonte confiável necessária de informações.

RPKI

O RPKI faz uso de um repositório que contém ROAs. Um ROA contém informações sobre o prefixo e seu número AS de BGP associado. A autorização de origem de rota é uma instrução assinada criptograficamente.

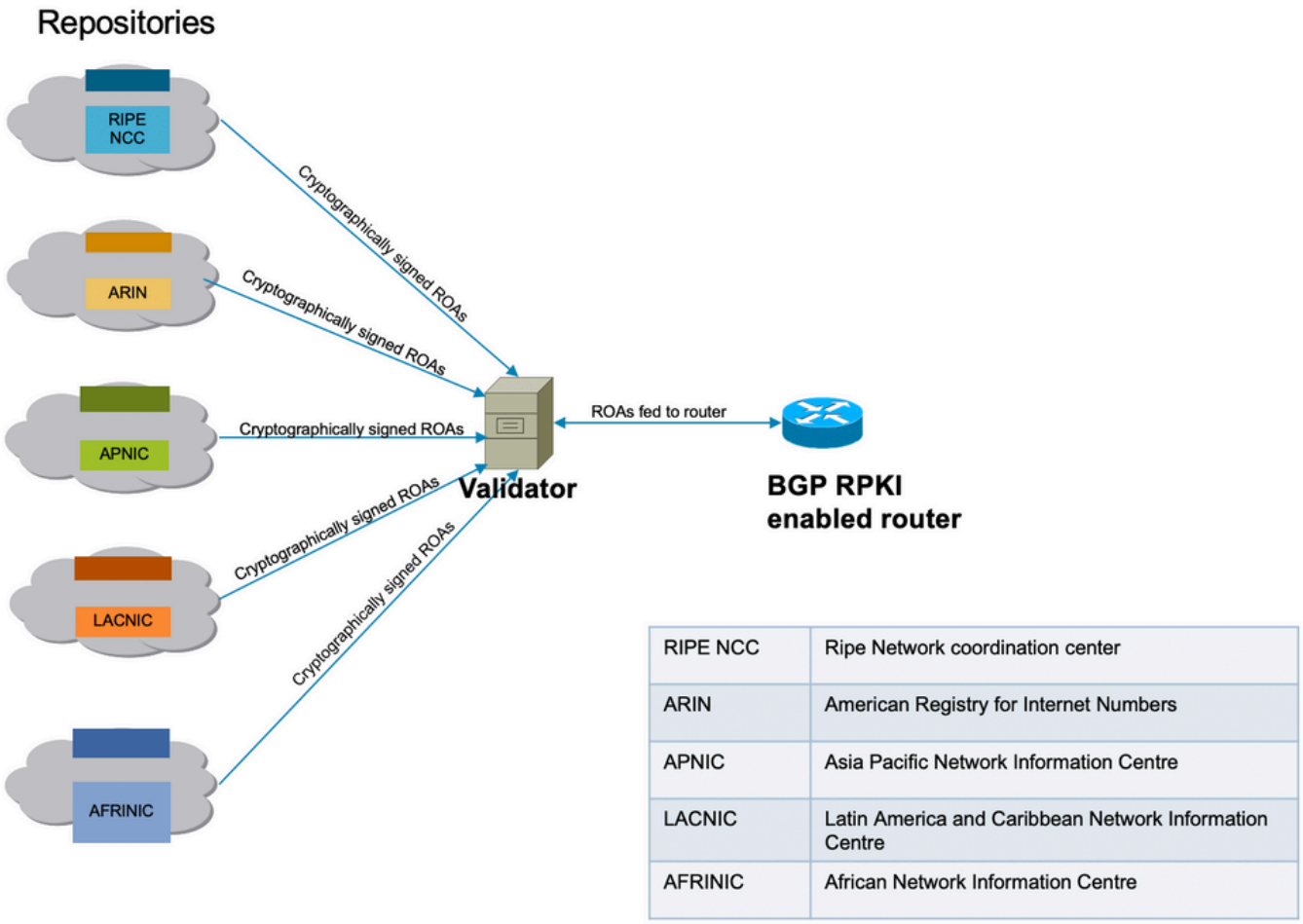
Os 5 Registros Internet Regionais (RIRs) são as âncoras de confiança do RPKI. A Internet Assigned Numbers Authority (IANA) é o topo da árvore que distribui prefixos IP. Os RIRs são os próximos na hierarquia. Eles atribuem subprefixos aos registros locais da Internet (LIRs) e aos grandes provedores de serviços de Internet (ISPs). Eles assinam um certificado para esses prefixos. O próximo nível aloca subprefixos desses e usa os certificados acima para assinar seus próprios certificados para certificar suas próprias alocações. Eles normalmente usam seus próprios pontos de publicação para hospedar os certificados e ROAs. Cada certificado lista os

pontos de publicação dos certificados filhos que ele assina. Assim, o RPKI forma uma árvore de certificados que espelha a árvore de alocações de endereço IP. Os validadores RPKI de propriedade das partes confiáveis sondam todos os pontos de publicação para encontrar certificados e ROAs (e CRLs e manifestos) atualizados. Eles começam nas âncoras de confiança e seguem os links para os pontos de publicação dos certificados filhos.

As ROAs são inseridas no repositório através de RIRs, mas o mesmo pode ser feito através de outros registros (nacionais ou locais). Esta responsabilidade também pode ser delegada nos FSI com supervisão e verificação adequadas por RIR.

Neste momento, existem cinco repositórios ROA mantidos por RIPE NCE, ARIN, APNIC, LACNIC e AFRINIC.

Um validador presente na rede se comunica com esses repositórios e faz o download de um banco de dados ROA confiável para criar seu cache. Esta é uma cópia agrupada do RPKI, que é buscada/atualizada periodicamente, direta ou indiretamente, a partir do RPKI global. O validador, então, alimenta essas informações aos roteadores, permitindo que eles comparem os anúncios de BGP recebidos com a tabela RPKI para tomar uma decisão segura.



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

conectividade de infraestrutura RPKI

Validador

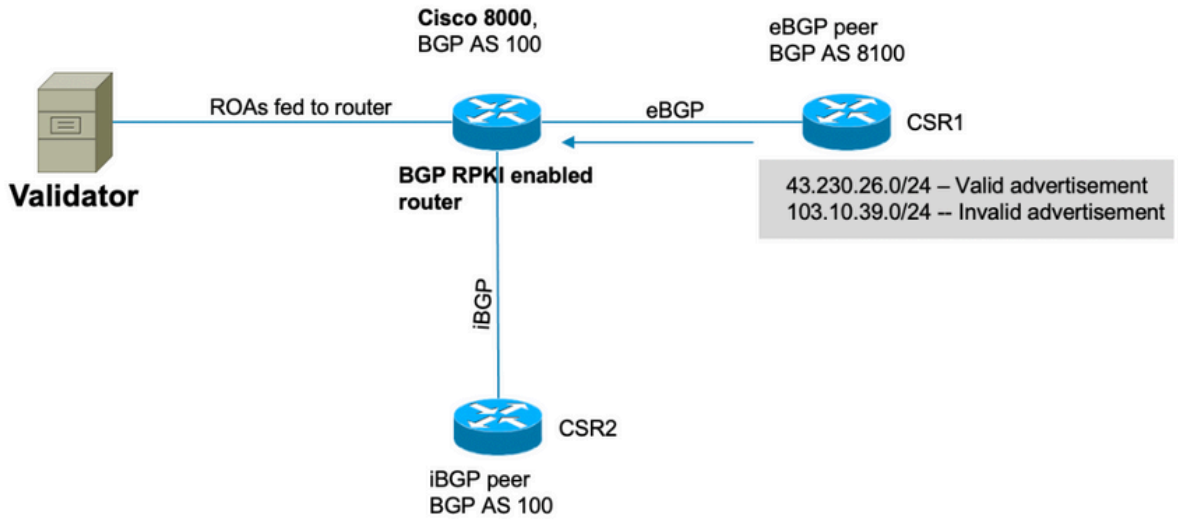
Esta demonstração usa o validador RIPE. O validador se comunicará com o roteador estabelecendo uma sessão TCP. Nesta demonstração, o validador ouve seu IP 192.168.122.120 e a porta 3323.

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

A IANA especificou a porta 3323 para esta comunicação. O temporizador de atualização define o intervalo de tempo após o qual o repositório local será sincronizado e atualizado para permanecer atualizado.

Demonstração de BGP RPKI

Topologia



Topologia

> **Observação**: esta demonstração usa o número aleatório de AS público e prefixos simplesmente para explicar a mecânica de RPKI do BGP. Os IPs públicos são usados devido ao RPKI, principalmente para proteção de prefixo público e todos os ROAs criados em RIRs são prefixos públicos. Por fim, nenhuma das ações, configurações, etc. descritas neste documento afetam esses IPs e AS públicos de nenhuma forma.

Configurar

```
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.
```

Sessão RPKI BGP

O roteador estabelece uma sessão TCP com um validador (IP: 192.168.122.120, porta 3323) para baixar o cache ROA para a memória do roteador.

```
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
```

Downloads de ROA no roteador

O validador alimenta as informações de ROA para o roteador. Esse cache é atualizado em intervalos periódicos para minimizar a possibilidade de o roteador manter informações obsoletas. Nesta demonstração, um tempo de atualização de 900 segundos foi configurado. Como mostrado aqui, o roteador Cisco 8000 fez download 172632 ROAs IPv4 e 28350 IPv6 do validador.

```
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
```

Verificar

Esta seção demonstra como o BGP RPKI está em ação e como ele evita que o roteador faça anúncios errados/ilegais.

Ativação da Validade da Origem Como

Por padrão, o roteador busca ROAs do validador, mas não começa a usá-los até que seja configurado para fazê-lo. Como resultado, esses prefixos são marcados como "D" ou desativados.

```
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 ?
```

Para habilitar o roteador para verificação de validade de as-origem, ative esse comando para a família de endereços em questão.

```
router bgp 100

 address-family ipv4 unicast
```

```
 bgp origin-as validation enable

 !
```

Quando você ativa esse comando, ele faz com que o roteador examine os prefixos presentes em sua tabela BGP em relação às informações ROA recebidas do validador e um dos três estados é atribuído aos prefixos .

```
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 ?
```

Para permitir que o roteador use as informações de estado de validação de prefixo ao fazer o cálculo do melhor caminho, este comando é necessário. Isso não é ativado por padrão, pois oferece a opção de não usar as informações de validade para o cálculo do melhor caminho, mas ainda permitir que você as use em políticas de rota que serão discutidas mais adiante neste documento.

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as use validity

 !
```

Estados de Validade do Prefixo

Há três estados nos quais um prefixo pode ser encontrado.

```
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 ?
```

- Inválido - Indica que o prefixo atende a uma destas duas condições: 1. Ele corresponde a uma ou mais **Route Origin Authorizations** (ROAs), mas não há nenhuma correspondência de ROA onde o AS de origem corresponda ao AS de origem no AS-PATH. 2. Ele corresponde a um ou mais ROAs no comprimento mínimo especificado no ROA, mas para todos os ROAs em que ele corresponde ao comprimento mínimo, é maior que o comprimento máximo especificado. A origem AS não importa para a condição #2.
- Válido - Indica que o prefixo e o par AS são encontrados na tabela de cache RPKI.
- Não encontrado - Indica que o prefixo não está entre os prefixos válidos ou inválidos.

Esta seção discute cada prefixo e seu estado em detalhes.

1. 203.0.113.0/24 - Válido

O peer eBGP no AS 8100 originou esta rota e anunciou ao nó Cisco8000. Como o AS de origem (8100) corresponde ao AS de origem no ROA (recebido do validador), esse prefixo é marcado como válido e instalado na tabela de roteamento do roteador.

```
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
```

A rota é instalada na tabela 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
```

Como esse é o melhor prefixo de BGP e também é válido por RPKI, ele é instalado com êxito na tabela de roteamento.

```
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 - Inválido

Esse prefixo é inválido porque há um conflito na informação AS de origem contida no ROA e na informação de origem como recebida através da mensagem BGP do peer eBGP. 203.0.113.1/24 é recebido via BGP com origem 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 ?
```

No entanto, o ROA recebido do validador mostra que esse prefixo pertence ao 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
```

Como as informações de origem AS no anúncio BGP recebido (AS 8100) não corresponderam à origem AS real recebida no ROA (AS 10021), o prefixo é marcado como Inválido e não é instalado na tabela de roteamento.

```
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 Não Encontrado

Este é um prefixo privado e não está presente no cache ROA. O BGP declarou este prefixo como "Não encontrado".

```
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
```

Como o RPKI ainda é adotado, os prefixos 'não encontrados' são instalados na tabela de roteamento. Caso contrário, fará com que o BGP ignore esses prefixos legítimos que não estão registrados no banco de dados RPKI.

Permitir prefixo inválido

Embora não seja recomendado, o software fornece um botão para permitir que prefixos inválidos participem do algoritmo de cálculo do melhor caminho.

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as allow invalid

 !
```

Com essa configuração, o roteador considera prefixos inválidos para o cálculo do melhor caminho, enquanto isso é marcado como "inválido". Esta saída mostra ‘203.0.113.1/24’ marcado como o melhor caminho.

```
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 ?
```

Como mostrado nesta saída, o prefixo é marcado como melhor apesar de ser mantido como inválido.

```
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
```

Éimportante observar que um roteador ainda trata o prefixo inválido como a última opção e sempre prefere um prefixo válido a um prefixo inválido, se disponível.

Configuração ROA manual no roteador

Se, por algum motivo, um ROA para um determinado prefixo ainda não tiver sido criado, recebido ou estiver atrasado, um ROA manual poderá ser configurado no roteador. Por exemplo, o prefixo ‘192.168.122.1/32’ é marcado como ‘Não encontrado’, como mostrado aqui.

```
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 ?
```

Um ROA manual pode ser configurado conforme mostrado aqui. Este comando associa o prefixo '192.168.122.1/32' com AS 8100.

```
router bgp 100

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

Com essa configuração, o estado do prefixo muda de 'N' para '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 ?
```

Política de rota e estado de validação de prefixo

O resultado do estado do prefixo pode ser usado para criar políticas de rota. Esses estados podem ser usados em uma instrução match e ações desejadas pelo administrador podem ser tomadas. Este exemplo corresponde a todos os prefixos com um estado inválido e define um valor de peso de 12345 para eles.

```
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
 !
 !
!
```

Esta saída mostra um peso de 12345 aplicado de prefixo inválido.

```
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
```

Compartilhar informações de validação de prefixo por meio da comunidade estendida

Como roteador BGP também pode compartilhar o estado de validação de prefixo com outros roteadores (sem cache local do validador) através da comunidade estendida BGP. Isso salva a sobrecarga de cada roteador na rede com uma sessão com o validador e o download de todos os ROAs.

Isso é possibilitado pela comunidade estendida de BGP.

Esse comando permite que o roteador compartilhe informações de ‘validação de prefixo’ com peers iBGP.

```
router bgp 100

 address-family ipv4 unicast

 bgp origin-as validation signal ibgp
```

Uma vez que o roteador Cisco 8000 é configurado como mostrado, as atualizações de BGP para pares incluem informações de validação de prefixo. Nesse caso, o roteador iBGP vizinho é um roteador 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
```

Esse mapeamento de comunidade estendida pode ser entendido com o uso de 0x4300 0x0000 (4 bytes indicando o estado).

Os quatro bytes que indicam o estado são tratados como um inteiro não assinado de 32 bits com um dos seguintes valores:

- 0 - Válido
- 1 - Não encontrado
- 2 - Inválido

A comunidade do prefixo 203.0.113.1/24 é 0x4300:0:2 que mapeia para o prefixo ‘Inválido’. Dessa forma, o roteador csr2, apesar de não ter cache local próprio, ainda é capaz de tomar decisões

com base no estado de validação de prefixo.

O estado de validação do prefixo agora pode ser usado para corresponder em um mapa de rota ou no algoritmo do melhor caminho BGP.

Recomendações para implementação de BGP RPKI

Boas práticas para a criação de ROA

Estas são algumas recomendações baseadas em redes inalcançáveis observadas no Observatório RPKI. O Observatório RPKI analisa vários aspectos do cenário RPKI implantado.

- Se um ROA é criado para qualquer prefixo, então é recomendável anunciar esse prefixo no BGP. Na ausência dele, outra pessoa pode anunciá-lo simplesmente fingindo ser um ASN contido nesse ROA e usar o prefixo.
- Se o ROA é criado com um maxlen maior que o comprimento do prefixo, então é equivalente a criar ROAs para todos os prefixos possíveis sob o prefixo original até o maxlen. É altamente recomendável anunciar todos esses prefixos no BGP.
- Se um ROA for criado para um prefixo e o proprietário do prefixo anunciar um subprefixo do prefixo original, o ROA invalidará esse subprefixo. Um ROA para o subprefixo também ou o maxlen do ROA original deve ser estendido para cobrir o subprefixo.
- Se uma organização possui um prefixo, mas planeja não anunciá-lo no BGP, então um ROA para o prefixo para AS0 deve ser criado. Isso invalidará qualquer anúncio de prefixo porque AS0 não pode aparecer em nenhum caminho AS.
- Se houver vários ASNs originando o mesmo prefixo, os ROAs desse prefixo deverão ser criados para cada um dos ASNs. Consequentemente, se um roteador tiver vários ROAs para o mesmo prefixo, um anúncio BGP que corresponda a qualquer um deles será válido. ROAs múltiplos para o mesmo prefixo não entram em conflito entre si.
- Se ‘A’ estiver originando um prefixo para seu cliente ‘B’ e criar um ROA para esse prefixo em nome de ‘B’, então ‘A’ deve anexar o ASN ‘B’ ao anúncio ou fazer com que o ‘B’ origine o próprio prefixo.

Impacto no desempenho de RPKI em roteadores BGP XR

Efeito da atualização de ROA na CPU com política de rota

Quando os ROAs são atualizados e se o roteador tem uma política de rota de ingresso local para um vizinho que contém um "estado de validação é", torna-se importante revalidar o status dos prefixos com base nos novos ROAs atualizados. Isso é obtido pelo roteador que envia uma solicitação BGP REFRESH ao seu peer.

Quando os vizinhos BGP recebem esta mensagem como mostrado, os vizinhos enviam seus prefixos novamente e a política de rota de entrada pode revalidar os prefixos de entrada .

```
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
```

O problema aumenta quando muitos vizinhos são atualizados ao mesmo tempo, sempre que os ROAs são atualizados. Se a política de rota de entrada vizinha for complexa e exigir muito processamento, altos resultados de CPU por alguns minutos após uma atualização de ROA. Essas mensagens REFRESH não ocorrem se a política de rota de entrada vizinha não contiver um comando "validation-state is".

Se "soft-reconfiguration inbound always" estiver configurado para um vizinho, então as mensagens BGP REFRESH não serão enviadas, mas as mesmas políticas de rota ainda serão executadas na mesma taxa e o mesmo uso de CPU pode ser esperado.

Recomenda-se preferir a abordagem "bgp bestpath origem-como validade de uso" em vez de configurar uma política de rota pelos motivos explicados em 6.2.2 abaixo.

Minimize o impacto da CPU causado pela atualização de ROA

A melhor maneira de evitar o problema explicado aqui é usar a **validade de origem de melhor caminho como uso** sem **estado de validação** na política.

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as use validity

 !
```

Esse comando mantém uma rota inválida recebida no roteador, mas impede que ela se torne o melhor caminho. Ele não será instalado ou anunciado novamente. É tão bom quanto deixá-lo cair. Se com a próxima atualização de ROA ela se tornar válida, nenhuma ATUALIZAÇÃO será necessária e ela se tornará elegível automaticamente para o melhor caminho sem que seja necessária a execução da política.

Se o usuário preferir permitir prefixos "inválidos" e não usá-los, então, além da **validade de bestpath origem-as use**, use a configuração **melhor caminho origem-as allow invalid**.

Nesse caso, quando um ROA muda, o melhor caminho é atualizado automaticamente sem exigir uma mensagem REFRESH. A fim de despreferir, uma rota significa que durante a seleção da rota BGP o caminho inválido RPKI é considerado menos preferível do que qualquer outro caminho para o mesmo destino. É semelhante a atribuir-lhe um peso ou preferência local menor que 0.

O número de RPKI inválidos é relativamente pequeno e mantido na tabela não resulta em um impacto significativo sobre os recursos.

> **Observação**: para usar a "origem do melhor caminho como validade do uso", todos os caminhos de uma rota, incluindo os caminhos do IBGP, devem ter a validade correta do RPKI. Caso contrário, o teste de estado de validação na política de rota ainda pode ser usado.

As rotas IBGP não são validadas pelo roteador em relação ao banco de dados ROA. As rotas IBGP ganham uma validade RPKI da comunidade estendida RPKI. Se a rota IBGP for recebida sem essa comunidade estendida, seu estado de validação será definido como não encontrado.

BGP RPKI Memory Footprint (Espaço de memória BGP RPKI)

Cada ROA consome memória para o índice e os dados. Se dois ROAs forem para o mesmo prefixo IP, mas tiverem max_len diferente ou forem recebidos de servidores RPKI diferentes, eles compartilharão o mesmo índice, mas terão dados separados. Os requisitos de memória podem variar porque a sobrecarga de memória não é constante. Recomenda-se um orçamento excessivo de 10%. As plataformas de 64 bits exigem mais memória para cada objeto de memória do que as plataformas de 32 bits. O uso de memória do IOS-XR em bytes para um objeto de índice e um objeto de dados está na tabela. Alguns custos indiretos constantes são incluídos nos números.

	Plataforma de 32 bits (bytes)	Plataforma de 64 bits(bytes)
índice de IPv4	74	111
índice de IPv6	86	125
dados	34	53

Esta seção utiliza dois cenários para explicar como os ROAs consomem memória.

Cenário 1. Três servidores RPKI configurados no roteador

Considere um roteador usando 3 servidores RPKI, cada um fornecendo 200.000 ROAs IPv4 e 20.000 ROAs IPv6 em um processador de rota de 64 bits exigirá esta memória:

20000 * (125 + 3*53) + 200000 * (11 + 3*53) bytes = 59,68 milhões de bytes

Ao calcular a memória, o ROA para o mesmo prefixo de três validadores diferentes compartilhou o mesmo valor de índice.

Cenário 2. Servidores RPKI únicos configurados no roteador

Memória de processo BGP sem ROAs:

```
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
```

O processo BGP é visto consumindo 25 MB de memória sem ROAs.

Memória de processo BGP com 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
```

O processo BGP é visto consumindo 25 MB de memória sem ROAs.

Memória de processo BGP com 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
```

O roteador Cisco 8000 executa um sistema operacional de 64 bits. Ele recebeu 172796 ROA IPv4 e 28411 ROA.

Memória (Bytes) = 172.796 x [111 (índice) + 53 (dados)] + 28411 x [125 (índice) + 53 (dados)].

Esses cálculos fornecem ~27 MB, que é aproximadamente o incremento observado na memória do roteador acima.

Sobre esta tradução
A Cisco traduziu este documento com a ajuda de tecnologias de tradução automática e humana para oferecer conteúdo de suporte aos seus usuários no seu próprio idioma, independentemente da localização.

Observe que mesmo a melhor tradução automática não será tão precisa quanto as realizadas por um tradutor profissional.

A Cisco Systems, Inc. não se responsabiliza pela precisão destas traduções e recomenda que o documento original em inglês (link fornecido) seja sempre consultado.