Understanding BGP RPKI With XR7 Cisco8000 Whitepaper

内容

概要

このドキュメントでは、Cisco IOS® XRプラットフォームのボーダーゲートウェイプロトコル(BGP)リソース公開キーインフラストラクチャ(RPKI)機能について説明します。

背景説明

はじめに

このドキュメントでは、BGP RPKI機能について説明し、ルータとのBGPを誤ったBGPプレフィックスアップデートや悪意のあるBGPプレフィックスアップデートから保護する方法について説明します。

範囲

このドキュメントでは、XR 7.3.1リリースを搭載したCisco 8000を使用してデモンストレーションを行います。ただし、BGP RPKIはプラットフォームに依存しない機能であるため、このドキュメントで説明する概念は、同等の適切なCLI変換を使用する他のシスコプラットフォーム(Cisco IOS、Cisco IOS-XE(IOS)を使用)にも適用されます。このドキュメントでは、地域インターネットレジストリにRoute Origin Authorizations(ROA)を追加する手順については説明しません。

前提条件

この読者には、BGPプロトコルに関する知識が必要です。

免責事項

このドキュメントで使用されているインターネットプロトコル(IP)アドレスは、実際のアドレスではありません。このドキュメントに含まれる例、コマンドの出力、および図は、説明のみを目的として示されています。説明の内容に実際のIPアドレスを使用することは、意図せず偶然に起こるものです。

不正なプレフィックスアドバタイズメントによるBGPの問題

BGPは、インターネットトラフィックのバックボーンとして機能します。これはインターネットコアの最も重要なコンポーネントですが、入力BGPアナウンスが認可された自律システムから発信されたものかどうかを確認する機能がありません。

このBGPの制限により、さまざまな種類の攻撃の候補として容易に使用できます。一般的な攻撃の1つは「ルートハイジャック」と呼ばれます。この攻撃を利用して、次のことを行うことができます。

- IPを盗んでスパムを送信すると、IPが拒否され、サービス拒否が発生します。
- トラフィックをスパイして、パスワードなどの機密情報を取得します。
- 管理者による設定の誤りによる中断。
- 偽のサーバを使用してトラフィックの配信を防止し、Denial of Service（DoS；サービス拒否）を実現します。

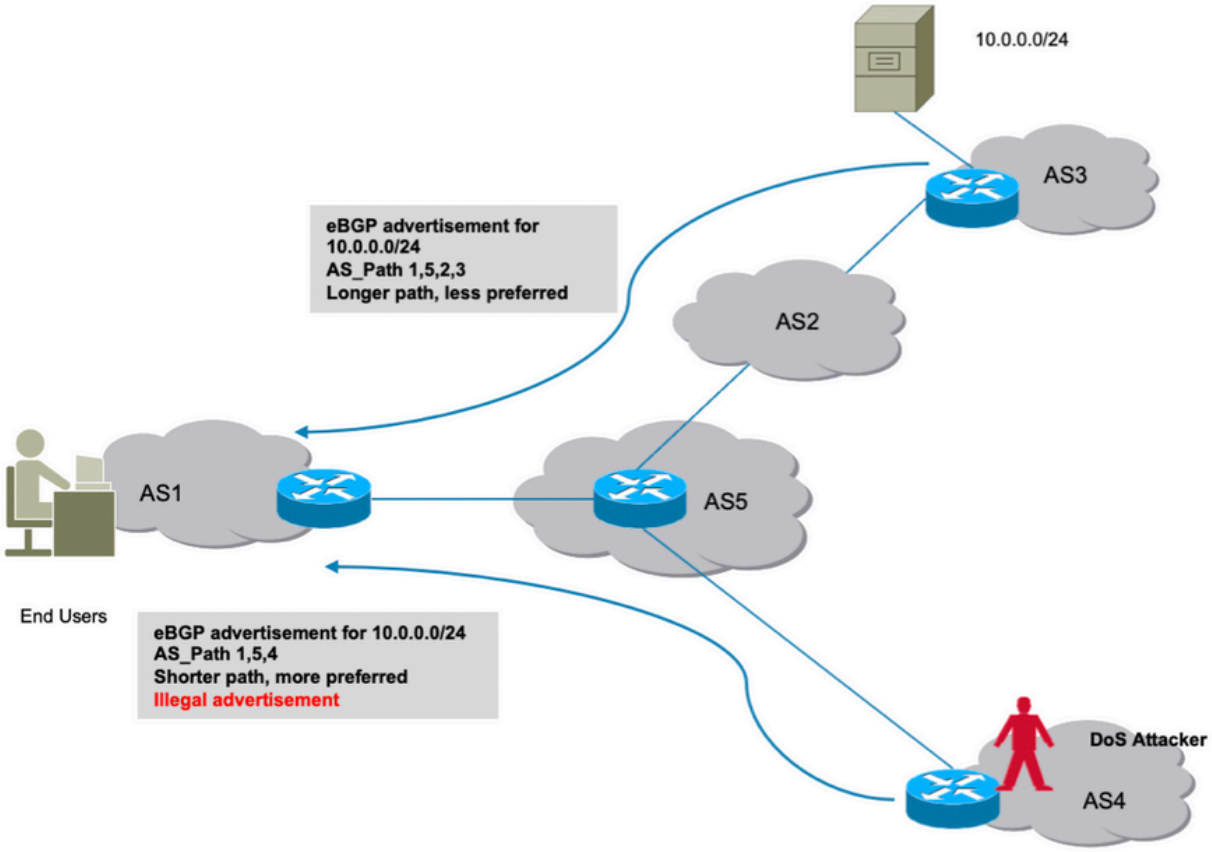
サービス拒否攻撃（一般にDoSと呼ばれる）は、ルータ、スイッチ、サーバなどへの通常のトラ

フィックを妨害する悪意のある試みです。さまざまなDoS攻撃がありますが、ここでは説明しません。

ルートのハイジャック

次に示すシナリオを考えてみます。自律システム3(AS3)は、プレフィックス10.0.0.0/24に対して正規のBGPアドバタイズメントを送信します。BGPの設計では、攻撃者が同じプレフィクスをインターネットにアドバタイズすることを妨げるものはBGPにはありません。

示されているように、AS4の攻撃者は同じプレフィクス10.0.0.0/24をアドバタイズします。BGPのベストパスアルゴリズムは、より短いAS_Pathを持つパスを優先します。AS_Path 1,5,4は、AS 1,5,2,3を経由する長いパスよりも優先されます。したがって、クライアントからのトラフィックは攻撃者の環境にリダイレクトされ、ブラックホール化され、エンドクライアントに対するサービス拒否が発生する可能性があります。

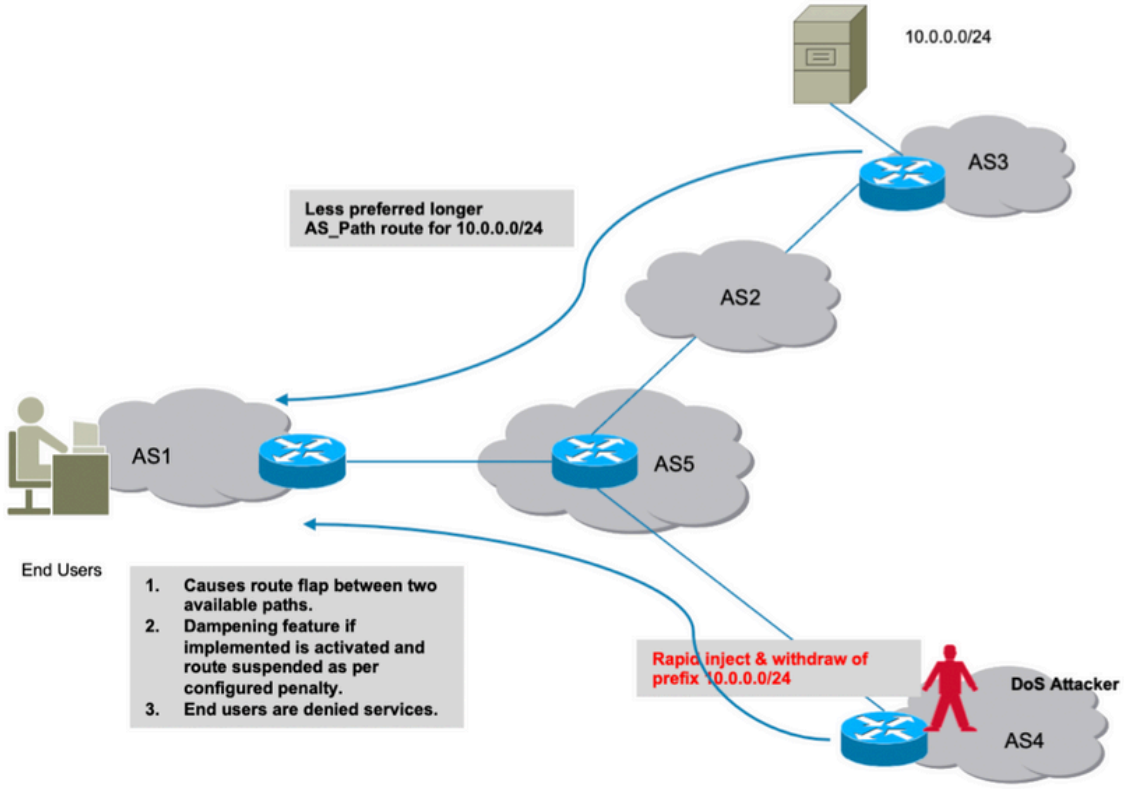


ルートハイジャック

システムパフォーマンスの低下

この項では、サービスを拒否する別の方法について説明します。CiscoのBGPルートダンプニング機能が設定されている場合、攻撃者がネットワーク内で高速ルートフラップを発生させ、絶え間ないチャーンを引き起こすと、この機能が悪用される可能性があります。

ダンプニング機能は、正当なルートにペナルティを課し、実際のトラフィックでは使用できなくなります。また、このような非倫理的に誘発されるフラップは、CPUやメモリなどのルータのリソースに負荷を与えます。

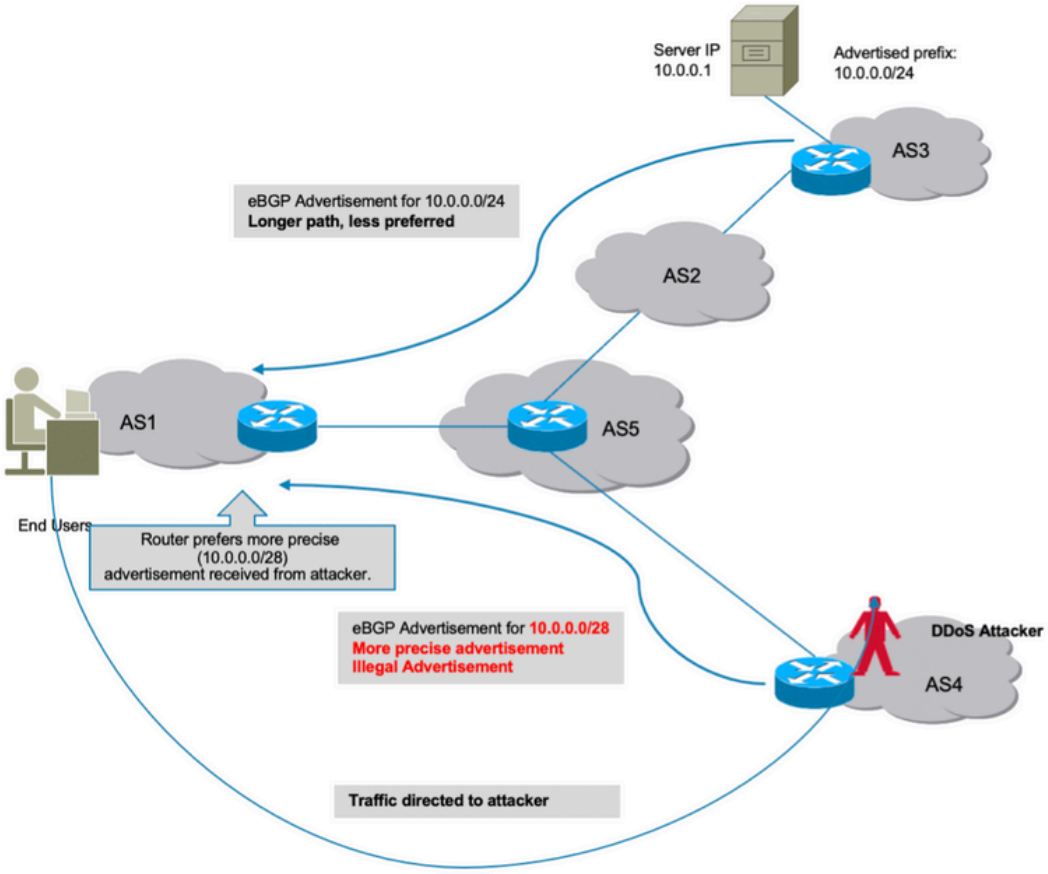


ルート ダンプニング

サブプレフィックスハイジャック

前のセクションで説明したように、攻撃者がプレフィックスを不正に発信し、トラフィックの中断を引き起こす方法です。残念ながら、混乱が懸念の唯一の原因ではありません。このような攻撃では、実際のデータが侵害される可能性があり、攻撃者は受信したデータをスキャンして非倫理的な使用を行うことができます。

同様に、ルートのハイジャックは、より正確なルートを不正にアドバタイズすることによって行われる可能性があります。図に示すように、BGPは一致するプレフィックスを優先するため、この動作が誤って悪用される可能性があります。



サブプレフィックスハイジャック

ここで説明する攻撃はすべて、悪意をもってアドバタイズされたプレフィックスの発信元ASが有効かどうかをBGPが特定できなかったことに起因します。これを修正するには、ルータがデータベースに保持できる「真の」信頼できるデータソースが必要です。その後、新しいアドバタイズメントを受信するたびに、ルータはBGPピアから受信したプレフィクスのASオリジン情報を、バリデータからのローカルデータベース情報と照合できるようになります。

したがって、ルータは良好なアドバタイズメントと不良（不正）なアドバタイズメントを区別でき、前述したすべての攻撃を回避する機能がルータに本質的に追加されます。 BGP RPKIは、必要な信頼できる情報源を提供します。

RPKI

RPKIは、ROAを含むリポジトリを使用します。ROAには、プレフィクスとそれに関連付けられたBGP AS番号に関する情報が含まれています。ルートの起点(origin)認可は、暗号署名付きの文です。

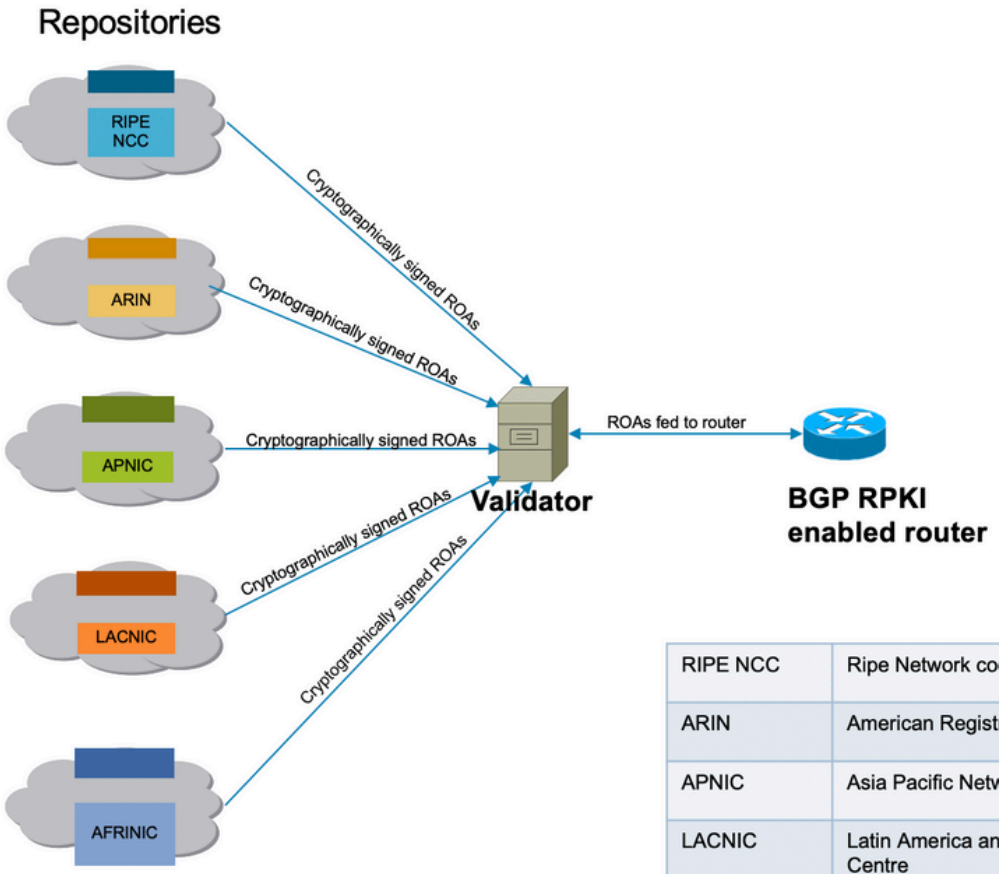
5つの地域インターネットレジストリ(RIR)は、RPKIのトラストアンカーです。Internet Assigned Numbers Authority(IANA)は、IPプレフィクスを配布するツリーの最上部です。RIRは階層の次にあります。ローカルインターネットレジストリ(LIR)と大規模なインターネットサービスプロバイダー(ISP)にサブプレフィックスを割り当てます。これらのプレフィックスの証明書に署名します。次のレベルでは、これらのサブプレフィックスを割り当て、上の証明書を使用して独自の証明書に署名し、独自の割り当てを証明します。通常は、独自の公開ポイントを使用して証明書とROAをホストします。各証明書には、署名した子証明書の公開ポイントが一覧表示されます。したがって、RPKIはIPアドレス割り当てのツリーをミラーリングする証明書のツリーを形成します

。証明書利用者が所有するRPKIバリデータは、すべての公開ポイントをポーリングして、更新された証明書とROA（およびCRLとマニフェスト）を検索します。信頼アンカーから開始し、子証明書の公開ポイントへのリンクに従います。

ROAはRIRを通じてリポジトリに入力されますが、他のレジストリ（国内またはローカル）を通じて同じことができます。この責任は、RIRによる適切な監視と検証によってISPに委任することもできます。

現時点では、RIPE NCE、ARIN、APNIC、LACNIC、AFRINICの5つのROAリポジトリが維持されています。

ネットワークに存在するバリデータは、これらのリポジトリと通信し、信頼できるROAデータベースをダウンロードしてキャッシュを構築します。これはRPKIの結合コピーであり、グローバルRPKIから直接または間接的に定期的にフェッチ/リフレッシュされます。次にバリデータはこの情報をルータに送信し、安全な判断を下すために着信BGPアナウンスとRPKIテーブルを比較できるようにします。



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

RPKIインフラストラクチャ接続

バリデータ

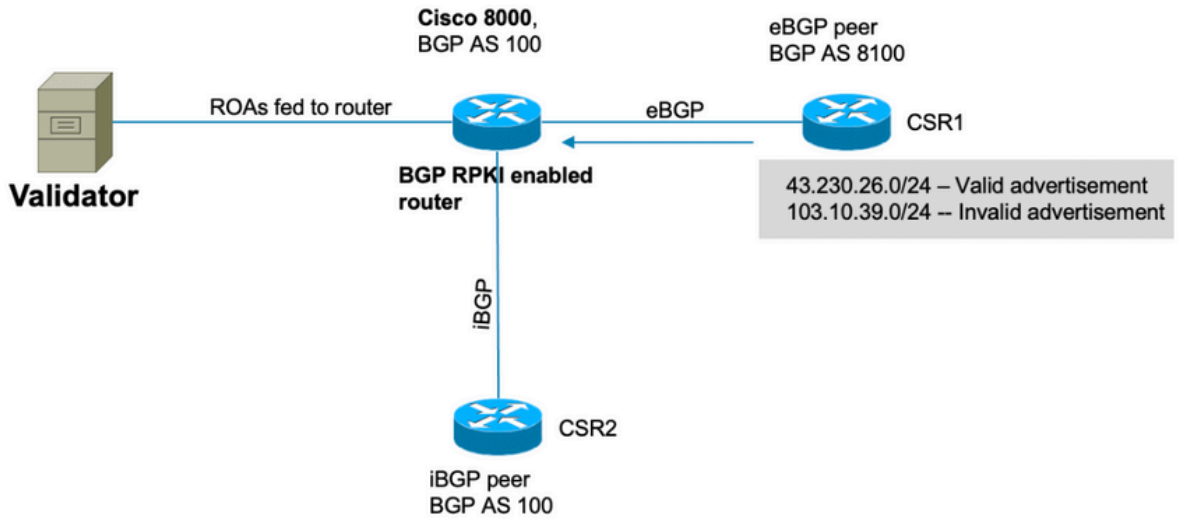
このデモンストレーションでは、RIPEバリデータを使用します。バリデータは、TCPセッションを確立することによってルータと通信します。このデモンストレーションでは、バリデータはIP 192.168.122.120およびポート3323をリッスンします。

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

IANAはこの通信にポート3323を指定しています。更新タイマーは、ローカルリポジトリが同期され、更新を維持するために更新されるまでの時間間隔を定義します。

BGP RPKIのデモンストレーション

トポロジ



トポロジ

> **注**：このデモンストレーションでは、単にBGP RPKIメカニズムを説明するために、ランダムなパブリックAS番号とプレフィクスを使用します。パブリックIPはRPKIによって主にパブリックプレフィックス保護のために使用され、RIRで作成されるすべてのROAはパブリックプレフィックスです。最後に、このドキュメントで説明されているアクションや設定などは、これらのパブリックIPやASに何ら影響を与えません。

設定

```
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.
```

BGP RPKIセッション

ルータは、ROAキャッシュをルータのメモリにダウンロードするために、バリデータ（IP:192.168.122.120、ポート3323）とのTCPセッションを確立します。

```
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
```

ルータでのROAのダウンロード

バリデータはROA情報をルータにフィードします。このキャッシュは、ルータが古い情報を保持する可能性を最小限に抑えるために、定期的に更新されます。このデモンストレーションでは、900秒のリフレッシュ時間が設定されています。 次に示すように、Cisco 8000ルータはIPv4とIPv6 ROAを172632バリデータから28350ウンロードしています。

```
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
```

確認

このセクションでは、BGP RPKIの動作の仕組み、およびルータの誤った/不正なアドバタイズメントを防止する方法について説明します。

Origin-Asの有効性の有効化

デフォルトでは、ルータはバリデータからROAをフェッチしますが、そのように設定されるまで使用を開始しません。その結果、これらのプレフィックスは「D」または無効としてマークされます。

```
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 ?
```

ルータでas-originの有効性チェックを有効にするには、該当するアドレスファミリに対してこのコマンドをアクティブにします。

```
router bgp 100

 address-family ipv4 unicast

 bgp origin-as validation enable

 !
```

このコマンドをアクティブにすると、ルータはBGPテーブルに存在するプレフィックスをバリデータから受信したROA情報と照合してスキャンし、3つの状態の1つがプレフィックス (BGPまたはBGP) に割り当てられます。

```
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 ?
```

最適パスの計算を行う際に、ルータがプレフィクス検証状態情報を使用できるようにするには、このコマンドが必要です。これはデフォルトでは有効になっていません。これは、ベストパスの計算に有効性情報を使用せずに、このドキュメントで後述するルートポリシーで使用することを可能にするオプションがあるためです。

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as use validity

 !
```

プレフィクスの有効状態

プレフィックスが見つかる可能性のある状態は3つあります。

```
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：プレフィクスが次の2つの条件のいずれかを満たしていることを示します。1.1つ以上のRoute Origin Authorizations(ROA)に一致しますが、起点ASがAS-PATH上の起点ASに一致するROAの一致はありません。2.ROAで指定された最小長で1つ以上のROAに一致しますが、最小長に一致するすべてのROAに対して、指定された最大長よりも長くなります。基準ASは条件#2には関係ありません。
- Valid:RPKIキャッシュテーブルでプレフィックスとASペアが見つかることを示します。
- Not Found – プレフィクスが有効または無効なプレフィクスの中に含まれていないことを示します。

この項では、各プレフィクスとその状態について詳しく説明します。

1. 203.0.113.0/24：有効

AS 8100のeBGPピアはこのルートを発信し、Cisco8000ノードにアドバタイズしました。送信元AS(8100)がROAの送信元AS（バリデータから受信）と一致するため、このプレフィックスは有効とマークされ、ルータのルーティングテーブルにインストールされます。

```
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
```

ルートはBGPテーブルにインストールされます。

```
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
```

これは最適なBGPプレフィックスであり、RPKIごとに有効であるため、ルーティングテーブルに正常にインストールされます。

```
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 – 無効

ROAに含まれる発信元AS情報と、eBGPピアからBGPメッセージを介して受信した発信元AS情報が競合しているため、このプレフィックスは無効です。203.0.113.1/24は、発信元AS 8100のBGPを介して受信されます。

```
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 ?
```

ただし、バリデータから受信したROAは、このプレフィクスがAS 10021に属していることを示します。

```
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
```

受信したBGPアナウンス(AS 8100)内のASオリジン情報が、ROA(AS 10021)で受信した実際のASオリジンと一致しないため、プレフィックスは無効としてマークされ、ルーティングテーブルにインストールされません。

```
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 Not Found

これはプライベートプレフィックスであり、ROAキャッシュには存在しません。BGPはこのプレフィクスを「Not found」と宣言しました。

```
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
```

RPKIは引き続き採用されるため、「見つからない」プレフィクスがルーティングテーブルにインストールされます。そうでない場合、BGPはRPKIデータベースに登録されていないこれらの正規のプレフィックスを無視します。

無効なプレフィックスを許可

推奨されませんが、ソフトウェアには、無効なプレフィックスがベストパス計算アルゴリズムに参加できるようにするノブが用意されています。

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as allow invalid

 !
```

この設定では、ルータはベストパス計算で無効なプレフィックスを考慮しますが、これは「無効」としてマークされます。この出力は、ベストパスとしてマークされた「203.0.113.1/24」を示しています。

```
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 ?
```

次の出力に示すように、無効に保たれているにもかかわらず、プレフィクスはbestとマークされます。

```
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
```

ルータは依然として無効なプレフィクスを最後のオプションとして扱い、有効なプレフィクスが使用可能な場合は無効なプレフィクスよりも常に有効なプレフィクスを優先することに注意してください。

ルータでの手動ROA設定

何らかの理由で、特定のプレフィックスのROAがまだ作成、受信、または遅延されていない場合は、ルータで手動ROAを設定できます。たとえば、次に示すように、プレフィクス「192.168.122.1/32」は「Not Found」としてマークされます。

```
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 ?
```

次に示すように、手動ROAを設定できます。このコマンドは、プレフィクス「192.168.122.1/32」をAS 8100に関連付けます。

```
router bgp 100

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

この設定では、プレフィクスの状態が「N」から「V」に変わります。

```
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 ?
```

ルートポリシーとプレフィクスの検証状態

プレフィクス状態の結果は、ルートポリシーの作成に使用できます。これらの状態はmatch文で使用でき、管理者が必要とするアクションを実行できます。 この例では、無効な状態のすべてのプレフィックスを照合し、それらのプレフィックスに対して重み値12345を設定します。

```
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
 !
 !
!
```

次の出力は、無効なプレフィクスが適用された重み付け12345を示しています。

```
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
```

拡張コミュニティによるプレフィクス検証情報の共有

BGPルータは、BGP拡張コミュニティを介して、プレフィクス検証状態を他のルータと共有することもできます（バリデータからのローカルキャッシュがない）。これにより、バリデータとの

セッションですべてのROAをダウンロードすることで、ネットワーク内の各ルータのオーバーヘッドを節約できます。

これは、BGP拡張コミュニティによって可能になります。

このコマンドにより、ルータは「プレフィクス検証」情報をiBGPピアと共有できます。

```
router bgp 100

 address-family ipv4 unicast

 bgp origin-as validation signal ibgp
```

Cisco 8000ルータが次のように設定されると、ピアへのBGPアップデートにはプレフィクス検証情報が含まれます。この場合、ネイバーiBGPルータはIOS-XEルータです。

```
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
```

この拡張コミュニティマッピングは、0x4300 0x0000（状態を示す4バイト）を使用して理解できます。

状態を示す4バイトは、次のいずれかの値を持つ32ビットの符号なし整数として扱われます。

- 0 – 有効
- 1 – 見つかりません
- 2 – 無効

プレフィクス203.0.113.1/24のコミュニティは0x4300:0:2で、「Invalid」プレフィクスにマッピングされます。このように、csr2ルータ自体のローカルキャッシュがないにもかかわらず、ルータはプレフィクス検証状態に基づいて決定を下すことができます。

プレフィクス検証状態を使用して、ルートマップまたはBGPベストパスアルゴリズムで照合できるようになりました。

BGP RPKIの実装に関する推奨事項

ROA作成のベストプラクティス

これらは、RPKI-Observatoryで観測された到達不能ネットワークに基づくいくつかの推奨事項です。RPKI天文台は、配備されたRPKIの風景の複数の側面を分析します。

- 任意のプレフィックスに対してROAが作成された場合は、そのプレフィックスをBGPでアナウンスすることを推奨します。ROAがない場合は、そのROAに含まれるASNのふりをしてプレフィックスを使用するだけで、他の誰かがそれをアナウンスできます。
- ROAがプレフィックス長より大きいmaxlenで作成される場合、ROAは、元のプレフィックスの下にある可能性のあるすべてのプレフィックスについて、maxlenまでのROAを作成することと同じです。これらのプレフィックスすべてをBGPでアナウンスすることを強く推奨します。
- プレフィックスに対してROAが作成され、プレフィックス所有者が元のプレフィックスのサブプレフィックスをアナウンスした場合、ROAはそのサブプレフィックスを無効にします。サブプレフィクスのROAも、元のROAのmaxlenも、サブプレフィクスをカバーするように拡張する必要があります。
- 組織がプレフィックスを所有しているが、それをBGPでアナウンスしない予定の場合は、AS0のプレフィックス用のROAを作成する必要があります。これにより、AS0はどのASパスにも現れないため、プレフィクスのアナウンスはすべて無効になります。
- 同じプレフィクスを発信元とするASNが複数ある場合は、そのプレフィクスのROAをASNごとに作成する必要があります。その結果、ルータに同じプレフィクスに対する複数のROAがある場合は、それらのいずれかに一致するBGPアドバタイズメントが有効になります。同じプレフィクスに対する複数のROAは互いに競合しません。
- 「A」がその顧客「B」のプレフィックスを発信し、「B」の代わりにそのプレフィックスのROAを作成する場合、「A」はアナウンスに「B」のASNを付加するか、「B」がプレフィックス自体を発信する必要があります。

XR BGPルータでのRPKIのパフォーマンスへの影響

ルートポリシーを使用したCPUでのROAアップデートの影響

ROAが更新される際、ルータに「validation-state is」が含まれるネイバーの口ーカル入力ルートポリシーがある場合は、更新された新しいROAに基づいてプレフィクスのステータスを再検証することが重要になります。これは、ルータがピアにBGP REFRESH要求を送信することで実現されます。

BGPネイバーがこのメッセージを次のように受信すると、ネイバーはプレフィックスを再度送信し、着信ルートポリシーは着信プレフィックスを再検証できます（BGPネイバーのプレフィクスは100 %以上です）。

```
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
```

ROAが更新されるたびに多くのネイバーが同時に更新されると、問題が増大します。ネイバーの着信ルートポリシーが複雑で、多くの処理が必要な場合、ROAの更新後に数分にわたってCPUの使用率が高くなります。ネイバーの着信ルートポリシーに「validation-state is」コマンドが含まれていない場合、これらのREFRESHメッセージは発生しません。

ネイバーに対して「soft-reconfiguration inbound always」が設定されている場合、BGP

REFRESHメッセージは送信されませんが、同じルートポリシーは同じレートで実行され、同じCPU使用率が予想されます。

次の6.2.2で説明されている理由から、ルートポリシーの設定よりも「bgp bestpath origin-as use validity」アプローチを優先することを推奨します。

ROAアップデートによるCPUへの影響の最小化

ここで説明する問題を回避する最善の方法は、ポリシーで**validation-state is**を指定せずに**bestpath origin-as use validity**を使用することです。

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as use validity

 !
```

このコマンドは、ルータで受信した無効なルートを保持しますが、ベストパスになるのを防ぎます。インストールも、さらにアドバタイズもされません。落とすも同然だ。次回のROA更新で有効になった場合は、REFRESHは必要なく、ポリシーの実行が必要なく、自動的にベストパスの対象となります。

ユーザが「無効な」プレフィクスを許可して使用しない場合は、**bestpath origin-as use validity**に加えて、**best path origin-as allow invalid**の設定を使用します。

この場合、ROAが変更されると、REFRESHメッセージを必要とせずにベストパスが自動的に更新されます。優先を解除するために、ルートは、BGPルート選択中にRPKI無効パスが同じ宛先への他のどのパスよりも優先されないと見なされることを意味します。これは、ウェイトの割り当てや、ローカルプリファレンスを0未満に設定する場合と似ています。

RPKI無効の数は比較的少なく、テーブルに保持されているため、リソースに大きな影響を与えることはありません。

> **注**:「bestpath origin-as use validity」を使用するには、IBGPパスを含むルートのすべてのパスが正しいRPKIの有効性を持つ必要があります。そうでない場合は、ルートポリシーのvalidation-stateのテストを引き続き使用できます。

IBGPルートは、ルータによってROAデータベースに対して検証されません。IBGPルートは、RPKI拡張コミュニティからRPKIの有効性を取得します。IBGPルートがこの拡張コミュニティなしで受信された場合、その検証状態はnot-foundに設定されます。

BGP RPKIメモリフットプリント

各ROAは、インデックスとデータのメモリを消費します。2つのROAが同じIPプレフィックスに対するものであっても、max_lenが異なっていたり、異なるRPKIサーバから受信されたりする場合、これらのROAは同じインデックスを共有しますが、データは異なります。メモリオーバーヘッドは一定ではないため、必要メモリはさまざまです。10 %の予算超過が推奨されます。64ビットプラットフォームでは、32ビットプラットフォームよりも多くのメモリが各メモリオブジェク

トに必要です。インデックスオブジェクトとデータオブジェクトのIOS-XRメモリ使用量（バイト単位）がテーブルに表示されます。ほとんど一定のオーバーヘッドが数値に含まれています。

	32ビットプラットフォーム（バイト）	64ビットプラットフォーム（バ）
IPv4インデックス	74	111
IPv6インデックス	86	125
data	34	53

このセクションでは、ROAがメモリを消費する方法を説明するために2つのシナリオを使用します。

シナリオ 1.ルータに設定された3台のRPKIサーバ

3台のRPKIサーバを使用するルータで、それぞれが64ビットルートプロセッサ上で200,000のIPv4 ROAと20,000のIPv6 ROAを提供する場合は、次のメモリが必要になります。

20000 * (125 + 3*53) + 200000 * (111 + 3*53)バイト= 5968万バイト

メモリの計算中、3つの異なるバリデータからの同じプレフィクスに対するROAは同じインデックス値を共有しました。

シナリオ 2.ルータに設定された単一のRPKIサーバ

ROAのないBGPプロセスメモリ：

```
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
```

BGPプロセスがROAなしで25 MBのメモリを消費していることがわかります。

ROAによるBGPプロセスメモリ：

```
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
```

BGPプロセスがROAなしで25 MBのメモリを消費していることがわかります。

ROAによるBGPプロセスメモリ：

```
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
```

Cisco 8000ルータは64ビットOSを実行します。IPv4 ROA172796IPv4 ROAを受信28411ました。

メモリ（バイト）= 172,796 x [111（インデックス）+ 53（データ）] + 28411 x [125（インデックス）+ 53（データ）]。

これらの計算では、上記のルータのメモリで認識される増分とほぼ同じ27 MBが得られます。

翻訳について
シスコは世界中のユーザにそれぞれの言語でサポート コンテンツを提供するために、機械と人による翻訳を組み合わせて、本ドキュメントを翻訳しています。ただし、最高度の機械翻訳であっても、専門家による翻訳のような正確性は確保されません。シスコは、これら翻訳の正確性について法的責任を負いません。原典である英語版（リンクからアクセス可能）もあわせて参照することを推奨します。