了解BGP RPKI With XR7 Cisco8000白皮书

目录

简介

本文档介绍Cisco IOS® XR平台上的边界网关协议(BGP)资源公钥基础设施(RPKI)功能。

背景信息

前言

本文档讨论BGP RPKI功能及其如何保护带有路由器的BGP免受假/恶意BGP前缀更新的影响。

范围

本文档使用带XR 7.3.1版的Cisco 8000进行演示。但是，BGP RPKI是一种与平台无关的功能，本文档中讨论的概念适用于具有相应等效CLI转换的其他Cisco平台（使用Cisco IOS、Cisco IOS-XE .）。本文档不涉及在区域互联网注册管理机构上添加路由来源授权(ROA)的过程。

先决条件

读者需要了解BGP协议。

免责声明

本文档中使用的任何 Internet 协议 (IP) 地址并不表示实际地址。文档中包含的任何示例、命令显示输出以及图示仅用于演示目的。演示内容中用到的任何实际 IP 地址纯属无意和巧合。

由于错误的前缀通告导致的BGP问题

BGP用作互联网流量的主干。尽管它是互联网核心最重要的组件，但它无法验证入口BGP通告是否源于授权的自治系统。

BGP的这种局限性使其很容易成为各种攻击的候选者。一种常见攻击称为“路由劫持”。利用此攻击可以：

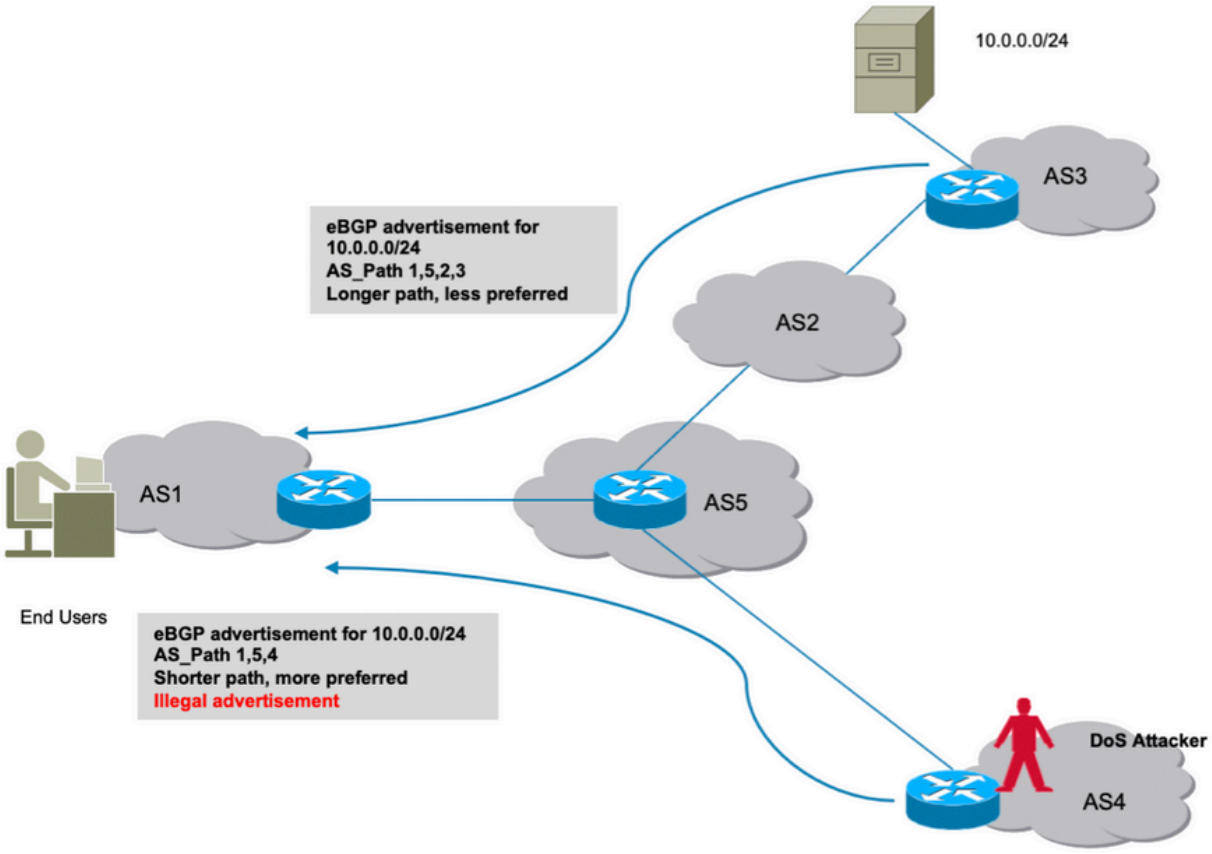
- 盗取IP以发送垃圾邮件，导致IP被拒绝，从而造成拒绝服务。
- 监视流量以获取密码等敏感信息。
- 管理员配置不正确导致的中断。
- 通过安装假服务器来防止流量传输，从而确保拒绝服务。

拒绝服务攻击（通常称为DoS）是一种恶意尝试，旨在中断到路由器、交换机、服务器等的正常流量。DoS攻击种类繁多，此处讨论的很少。

路由劫持

请考虑此处显示的场景。自治系统3(AS3)发送其前缀10.0.0.0/24的合法BGP通告。根据BGP的设计，BGP中没有任何内容可以阻止攻击者将相同前缀通告到Internet。

如图所示，AS4中的攻击者通告相同的前缀10.0.0.0/24。BGP最佳路径算法首选具有较短AS_Path的路径。AS_Path 1、5、4通过AS 1、5、2、3在较长路径上胜出。因此，来自客户端的流量现在将被重定向到攻击者的环境，并且可能会被黑洞，从而造成对终端客户端的拒绝服务。

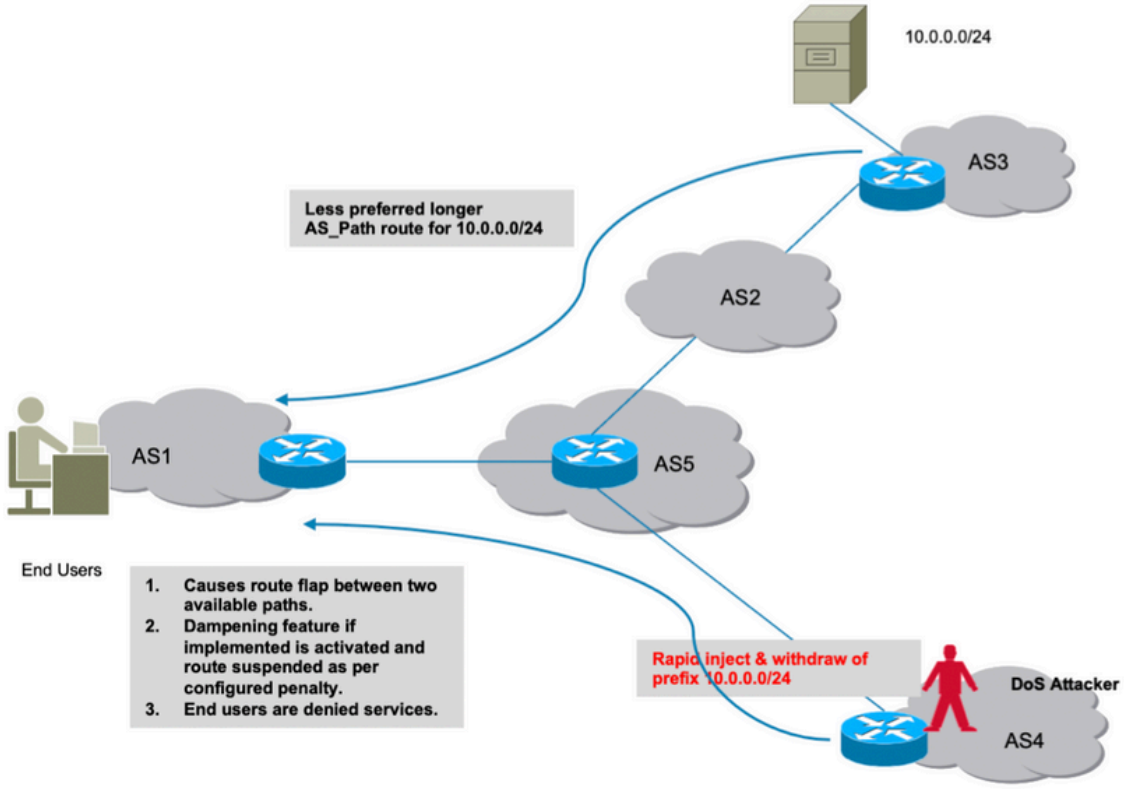


路由劫持

降低系统性能

本节讨论拒绝服务的另一种方式。如果配置了思科的BGP路由抑制功能，则攻击者可以在网络中引入快速路由抖动导致持续抖动时，利用此功能。

阻尼功能将对合法路由施加惩罚并使其无法用于实际流量。此外，这种不道德引起的抖动会对路由器的CPU、内存等资源造成压力。

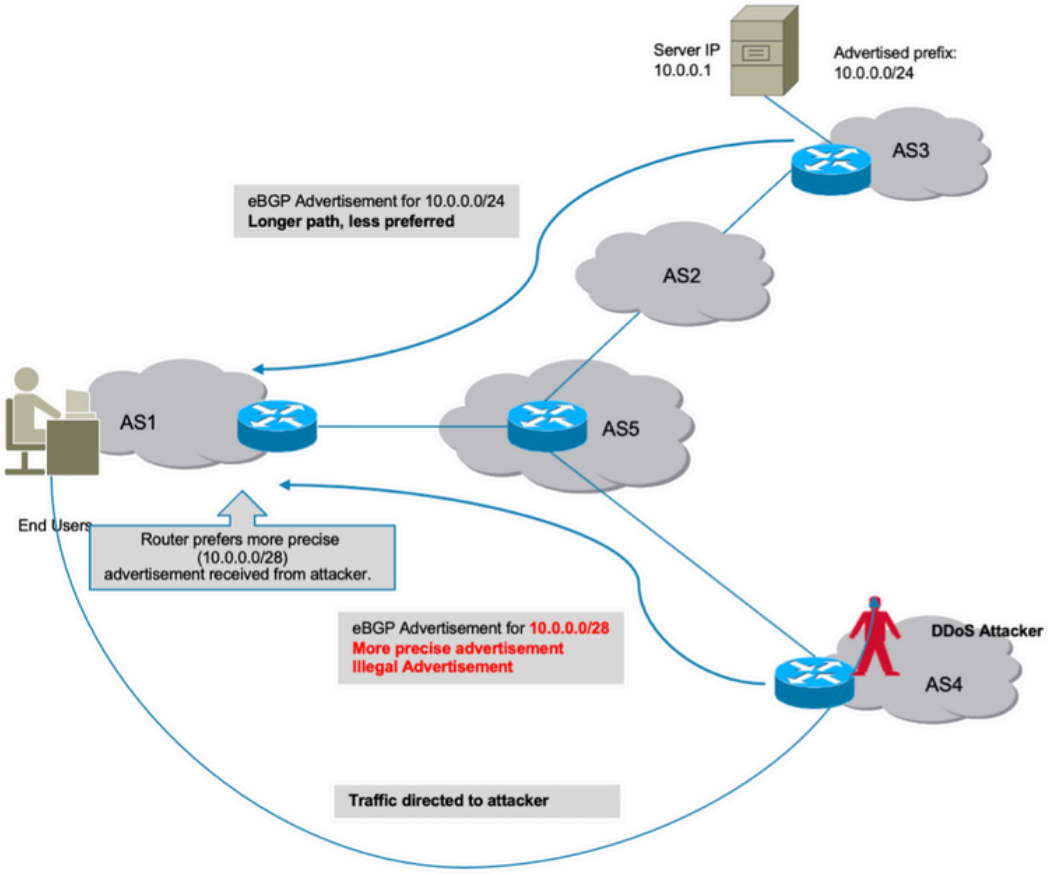


路由惩罚

子前缀劫持

如上一节所述，攻击者如何非法生成前缀并导致流量中断。不幸的是，颠覆并不是唯一令人担忧的原因。在此类攻击中，实际数据可能受到危害，其中攻击者可以扫描接收到的数据以将其用于不道德用途。

同样，劫持一条路线也可以通过非法宣传一条更精确的路线来完成。BGP首选较长匹配的前缀，并且此行为可能会被错误利用，如图所示。



子前缀劫持

讨论的所有攻击都源于这样一个事实：BGP无法识别这些恶意通告的前缀的源AS是否有效。要解决此问题，需要路由器可在其数据库中保留的“true”和“trusted”数据源。然后，每次收到新通告后，路由器现在能够使用来自验证器的本地数据库信息交叉验证从BGP对等体接收的前缀的AS源信息。

因此，路由器能够区分好通告和坏（非法）通告，并且路由器本身增加了避免前面讨论的所有攻击的能力。 BGP RPKI提供所需的可信信息源。

RPKI

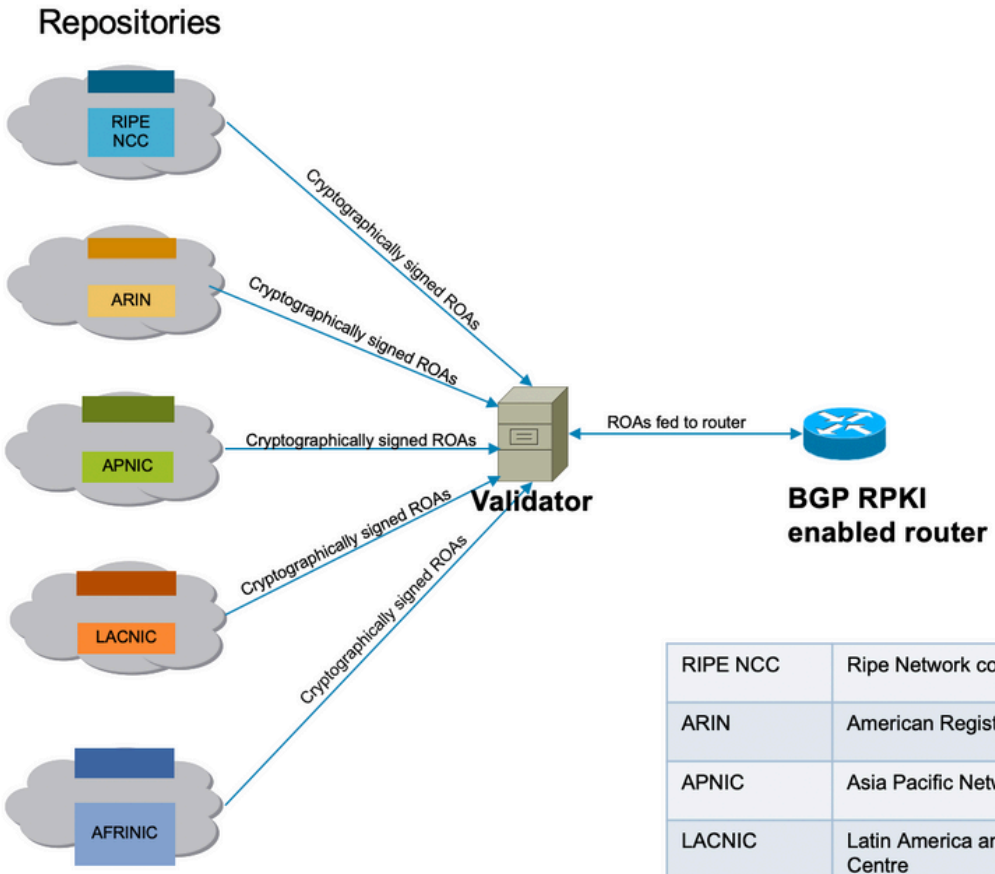
RPKI使用包含ROA的存储库。ROA包含有关前缀及其关联的BGP AS编号的信息。路由来源授权是一个加密签名的语句。

5个地区Internet注册管理机构(RIR)是RPKI的信任锚点。互联网编号指派机构(IANA)是派发IP前缀的树的顶部。RIR位于层次结构中的下一层。它们将子前缀分配给本地Internet注册管理机构(LIR)和大型互联网服务提供商(ISP)。他们为这些前缀签署证书。下一层分配这些前缀的子前缀，并使用上面的证书签署自己的证书来验证自己的分配。它们通常使用自己的发布点来托管证书和ROA。每个证书列出其签名的子证书的发布点。因此，RPKI形成了镜像IP地址分配树的证书树。信赖方拥有的RPKI验证器会轮询所有发布点，以查找更新的证书和ROA（以及CRL和清单）。它们从信任锚点开始，然后点击子证书发布点的链接。

ROA通过RIR输入存储库，但也可通过其他注册机构（国家或地方）进行相同操作。RIR还可以将这项责任委托给ISP，并对其进行适当的监督和验证。

目前，有五个由RIPE NCE、ARIN、APNIC、LACNIC和AFRINIC维护的ROA存储库。

网络中的验证程序与这些存储库通信，并下载可信ROA数据库以构建其缓存。这是RPKI的合并副本，定期从全局RPKI直接或间接获取/刷新。然后，验证器将此信息提供给路由器，使路由器能够将传入BGP通告与RPKI表进行比较，以便做出安全决策。



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基础设施连接

验证程序

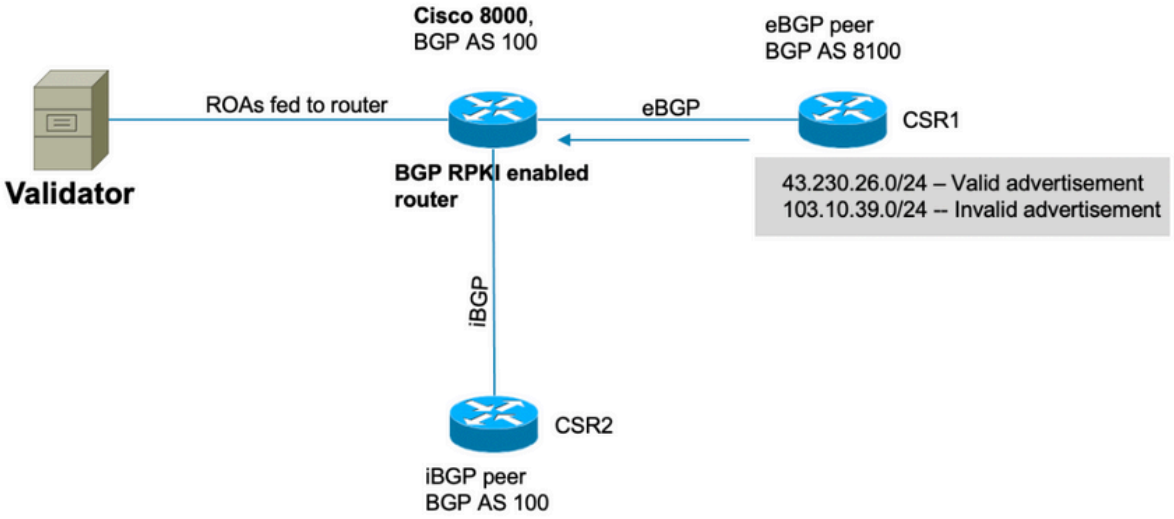
此演示使用RIPE验证器。验证器将通过建立TCP会话与路由器通信。在本演示中，验证程序侦听其IP 192.168.122.120和端口3323。

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

IANA为此通信指定了端口3323。刷新计时器定义本地存储库进行同步和更新以保持更新的时间间隔。

BGP RPKI演示

拓扑



拓扑

> **注意**：此演示使用随机公共AS编号和前缀，只是为了说明BGP RPKI机制。公共IP是因为RPKI主要用于公共前缀保护，并且在RIR上创建的所有ROA都是公共前缀。最后，本文档中描述的任何操作、配置等都不会以任何方式影响这些公共IP和AS。

配置

```
 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会话

路由器使用验证器（IP: 192.168.122.120，端口3323）建立TCP会话，以将ROA缓存下载到路由器的内存。

```
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
```

路由器上的ROA下载

验证程序将ROA信息提供给路由器。此缓存会定期刷新，以尽量减少路由器保存过时信息的可能性。在本演示中，已配置刷新时间900秒。 如图所示，Cisco 8000路由器已从验172632器下载IPv28350和IPv6 ROA。

```
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
```

验证

本部分演示了BGP RPKI如何发挥作用，以及它如何防止路由器发生错误/非法通告。

启用Origin-As有效性

默认情况下，路由器从验证器获取ROA，但只有在配置为使用ROA后才会开始使用。因此，这些前缀被标记为“D”或禁用。

```
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 ?
```

要启用路由器进行原点有效性检查，请激活相关地址系列的此命令。

```
router bgp 100

 address-family ipv4 unicast

 bgp origin-as validation enable

 !
```

当您激活此命令时，它会使路由器根据从验证器收到的ROA信息扫描其BGP表中存在的前缀，并且三个状态之一分配给前缀（例如BGP）。

```
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 ?
```

为了使路由器能够在进行最佳路径计算时使用前缀验证状态信息，需要此命令。默认情况下不启用此功能，因为它允许您不将有效性信息用于最佳路径计算，但仍允许您将其用于本文档稍后将介绍的路由策略。

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as use validity

 !
```

前缀有效状态

可在三种状态中找到前缀。

```
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 ?
```

- 无效 — 表示前缀满足以下两个条件之一：1。它匹配一个或多个**路由来源授权**(ROA)，但是如果源AS与AS-PATH上的源AS匹配，则没有ROA匹配。2.它在ROA中指定的最小长度上匹配一个或多个ROA，但是对于与最小长度匹配的所有ROA，其长度大于指定的最大长度。源AS与条件#2无关。
- 有效 — 表示在RPKI缓存表中找到前缀和AS对。
- Not Found — 表示前缀不在有效或无效的前缀之列。

本节详细讨论每个前缀及其状态。

1. 203.0.113.0/24 — 有效

AS 8100中的eBGP对等体生成此路由并通告给Cisco8000节点。由于源AS(8100)与ROA中的源AS（从验证器接收）匹配，因此该前缀被标记为有效并安装在路由器的路由表中。

```
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
```

该路由安装在BGP表中。

```
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
```

由于这是最佳的BGP前缀，而且每个RPKI也有效，因此它成功安装到路由表中。

```
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 — 无效

此前缀无效，因为ROA中包含的源AS信息和通过BGP消息从eBGP对等体接收的源AS信息之间存在冲突。203.0.113.1/24通过BGP与源AS 8100接收。

```
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 ?
```

但是，从验证器收到的ROA显示该前缀属于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
```

由于收到的BGP通告(AS 8100)中的AS源信息与ROA(AS 10021)中收到的实际AS源信息不匹配，因此前缀被标记为“无效”(Invalid)且未安装在路由表中。

```
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.未找到192.168.122.1/32

这是私有前缀，不在ROA缓存中。BGP将此前缀声明为“Not found”。

```
RP/0/RP0/CPU0:Cisco8000#show bgp 192.168.122.1/32
Thu Jan 21 05:44:39.861 UTC
BGP routing table entry for 192.168.122.1/32
Versions:
 Process           bRIB/RIB  SendTblVer
 Speaker                 33          33
Last Modified: Jan 21 00:03:33.344 for 05:41:06
Paths: (1 available, best #1)
 Not advertised to any peer
 Path #1: Received by speaker 0
 Not advertised to any peer
 8100
```

```
   10.0.12.2 from 10.0.12.2 (192.168.122.105)

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

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

     Origin-AS validity: not-found
```

由于仍采用RPKI，因此路由表中会安装“not-found”前缀。否则，BGP将忽略这些未在RPKI数据库中注册的合法前缀。

允许无效前缀

虽然不建议这样做，但软件确实提供了一个命令，以允许无效前缀参与最佳路径计算算法。

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as allow invalid

 !
```

使用此配置时，路由器会考虑无效前缀以进行最佳路径计算，而此前缀标记为“无效”。此输出显示标记为最佳路径的“203.0.113.1/24”。

```
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 ?
```

如此输出所示，尽管前缀保持无效，但将其标记为最佳。

```
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
```

请注意，路由器仍然将无效前缀视为最后一个选项，并且始终首选有效前缀，而不是无效前缀（如果可用）。

路由器上的手动ROA配置

如果由于某种原因，尚未创建、接收或延迟特定前缀的ROA，则可以在路由器上配置手动ROA。例如，前缀“192.168.122.1/32”标记为“未找到”，如此处所示。

```
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 ?
```

手动ROA可配置如图所示。此命令将前缀“192.168.122.1/32”与AS 8100关联。

```
router bgp 100

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

通过此配置，前缀的状态从“N”更改为“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 ?
```

路由策略和前缀验证状态

前缀状态结果可用于创建路由策略。可以在match语句中使用这些状态，并可执行管理员所需的操作。 此示例匹配具有无效状态的所有前缀，并为它们设置权重值12345。

```
route-policy Invalid

 if validation-state is invalid then

   set weight 12345

 endif

end-policy

!



router bgp 100

 remote-as 8100

 address-family ipv4 unicast

  route-policy Invalid in

 !

 !

!
```

此输出显示应用了无效的前缀权重为12345。

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

Thu Jan 21 06:57:33.816 UTC

BGP routing table entry for 203.0.113.1/24

Versions:

 Process           bRIB/RIB  SendTblVer

 Speaker                38          38
```

```
Last Modified: Jan 21 06:54:04.344 for 00:03:29

Paths: (1 available, best #1)

 Not advertised to any peer

 Path #1: Received by speaker 0

 Not advertised to any peer

 8100

   10.0.12.2 from 10.0.12.2 (192.168.122.105)

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

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

     Origin-AS validity: invalid
```

通过扩展社区共享前缀验证信息

BGP路由器还可以通过BGP扩展社区与其他路由器（没有来自验证器的本地缓存）共享前缀验证状态。这节省了网络中每台路由器与验证器进行会话并下载所有ROA的开销。

这可以由BGP扩展社区实现。

此命令使路由器能够与iBGP对等体共享“前缀验证”信息。

```
router bgp 100

 address-family ipv4 unicast

 bgp origin-as validation signal ibgp
```

如图所示配置Cisco 8000路由器后，对等体的BGP更新会包含前缀验证信息。在这种情况下，邻居iBGP路由器是IOS-XE路由器。

```
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
```
使用0x4300 0x000（4字节表示状态）可以理解此扩展社区映射。

表示状态的四个字节被视为32位无符号整数，它具有以下值之一：

- 0 — 有效
- 1 — 未找到
- 2 — 无效

前缀203.0.113.1/24的社区是0x4300:0:2，该社区映射到“无效”前缀。这样，csr2路由器即使没有自己的本地缓存，仍能够基于前缀验证状态做出决策。

现在，前缀验证状态可用于在路由映射或BGP最佳路径算法中进行匹配。

BGP RPKI实施建议

创建ROA的良好做法

这些建议基于在RPKI-Observatory观察到的不可达网络。RPKI Observatory分析已部署的RPKI环境的多个方面。

- 如果为任何前缀创建ROA，则建议在BGP中通告该前缀。如果没有该前缀，则其他人可以通过简单伪装成ROA中包含的ASN并使用前缀来通告它。
- 如果创建的ROA的maxlen大于前缀长度，则它相当于为maxlen之前的原始前缀下所有可能的前缀创建ROA。强烈建议在BGP中通告所有这些前缀。
- 如果为前缀创建ROA，并且前缀所有者宣布原始前缀的子前缀，则ROA将使该子前缀失效。子前缀的ROA以及原始ROA的maxlen必须扩展以覆盖子前缀。
- 如果组织拥有前缀，但计划不在BGP中通告它，则必须为AS0的前缀创建ROA。这将使任何前缀通告失效，因为AS0不能出现在任何AS路径中。
- 如果有多个ASN源自同一个前缀，则必须为每个ASN创建该前缀的ROA。因此，如果路由器具有用于相同前缀的多个ROA，则与其中任何一个ROA匹配的BGP通告将有效。同一前缀的多个ROA不会相互冲突。
- 如果“A”为其客户“B”生成一个前缀，并代表“B”为该前缀创建ROA，则“A”必须在通告中预置“B”的ASN，或使“B”生成前缀本身。

RPKI对XR BGP路由器性能的影响

ROA更新对具有路由策略的CPU的影响

当ROA更新时，如果路由器具有包含“validation-state is”的邻居的本地入口路由策略，则根据新更新的ROA重新验证前缀的状态非常重要。这是通过路由器向其对等体发送BGP REFRESH请求来实现的。

当BGP邻居收到此消息时（如图所示），邻居会再次发送其前缀，入站路由策略可以重新验证传入前缀。

```
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
```
每当更新ROA时，当许多邻居同时刷新时，问题就会放大。如果邻居入站路由策略很复杂且需要大量处理，则在ROA更新后几分钟内会导致高CPU使用率。如果邻居入站路由策略不包含“validation-state is”命令，则不会出现这些REFRESH消息。

如果为邻居配置了“始终进行软重新配置入站”，则不会发送BGP REFRESH消息，但仍然会以相同的速率执行相同的路由策略，而且可以预期相同的CPU使用率。

出于下面6.2.2中说明的原因，建议使用“bgp bestpath origin-as use validity”方法而不是配置路由策略。

将ROA更新对CPU的影响降至最低

避免此处所说明问题的最佳方法是使用**bestpath origin-as use validity**而无**validation-state**在策略中。

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as use validity

 !
```
此命令将收到的无效路由保留在路由器上，但阻止它成为最佳路径。不会安装或进一步通告它。就好比扔掉一样。如果下次ROA更新生效，则无需刷新，它将自动符合最佳路径的条件，无需执行策略。

如果用户喜欢允许“无效”前缀而不使用它们，则除了**bestpath origin-as use validity**外，请使用配置最佳路径**origin-as allow invalid**。

在这种情况下，当ROA更改时，最佳路径会自动更新，无需REFRESH消息。为了取消首选，路由意味着在BGP路由选择期间，RPKI无效路径被视为比通往同一目的地的任何其他路径更不优选。这类似于为其指定权重或小于0的本地优先级。

RPKI无效的数量相对较少，并且保存在该表中不会导致对资源的显着影响。

> **注**：要使用“最佳路径原点为使用有效性”，路由的所有路径（包括IBGP路径）都必须具有正确的RPKI有效性。如果不是，则仍可使用路由策略中的验证状态测试。

路由器不会根据ROA数据库验证IBGP路由。IBGP路由从RPKI扩展社区获得RPKI有效性。如果接收到没有此扩展社区的IBGP路由，则其validation-state设置为not-found。

BGP RPKI内存空间

每个ROA消耗索引和数据内存。如果两个ROA用于同一IP前缀，但具有不同的max_len，或者从不同的RPKI服务器接收，则它们共享相同的索引，但具有不同的数据。内存要求可能不同，因为内存开销不是固定的。建议超支10%。与32位平台相比，64位平台需要为每个内存对象提供更多内存。表中列出了索引对象和数据对象的IOS-XR内存使用情况（以字节为单位）。这些数字中包括一些基

本为常数的开销。

	32位平台（字节）	64位平台（字节）
IPv4索引	74	111
IPv6索引	86	125
数据	34	53

本节采用两种方案来解释ROA如何消耗内存。

场景 1.路由器上配置了三台RPKI服务器

假设一台路由器使用3台RPKI服务器，每台在64位路由处理器上提供200,000个IPv4 ROA和20,000个IPv6 ROA将需要此内存：

20000 *(125 + 3*53)+ 200000 *(111 + 3*53)字节= 59.68百万字节

在计算内存时，来自三个不同验证器的相同前缀的ROA共享相同的索引值。

场景 2：路由器上配置的单个RPKI服务器

没有ROA的BGP进程内存：

```
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进程占用25 MB内存，无任何ROA。

使用ROA的BGP进程内存：

```
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进程占用25 MB内存，无任何ROA。

使用ROA的BGP进程内存：

```
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路由器运行64位操作系统。它接收172796IPv4 ROA和28411 ROA。

内存（字节）= 172,796 x [111（索引）+ 53（数据）] + 28411 x [125（索引）+ 53（数据）]。

这些计算得出约27 MB，大约是上述路由器内存中发现的增量。

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。