瞭解BGP RPKI With XR7 Cisco8000白皮書

目錄

簡介

本檔案介紹Cisco IOS® XR平台上的邊界閘道通訊協定(BGP)資源公開金鑰基礎架構(RPKI)功能。

背景資訊

前言

本檔案將討論BGP RPKI功能，以及它如何保護路由器的BGP免受假/惡意BGP首碼更新的影響。

範圍

本檔案使用Cisco 8000及XR 7.3.1版本作示範。但是，BGP RPKI是一種與平台無關的功能，本文討論的概念適用於具有相應等效CLI轉換的其他Cisco平台（使用Cisco IOS、Cisco IOS-XE．）。本文檔不涉及在區域網際網路登錄檔中新增路由來源授權(ROA)的過程。

必要條件

讀者需要瞭解BGP協定。

簡介

本文檔中使用的任何Internet協定(IP)地址均不是實際地址。本文檔中包含的任何示例、命令顯示輸出和圖示僅作說明之用。在說明性內容中使用實際的IP地址是無意的，而且純屬巧合。

由於錯誤的字首通告導致的BGP問題

BGP充當Internet流量的中樞。雖然它是網際網路核心最重要的元件，但它缺乏驗證輸入BGP通告是否源自授權自治系統的能力。

BGP的這種侷限性使它很容易成為各種攻擊的一個候選者。一種常見攻擊稱為「路由劫持」。利用此攻擊可以：

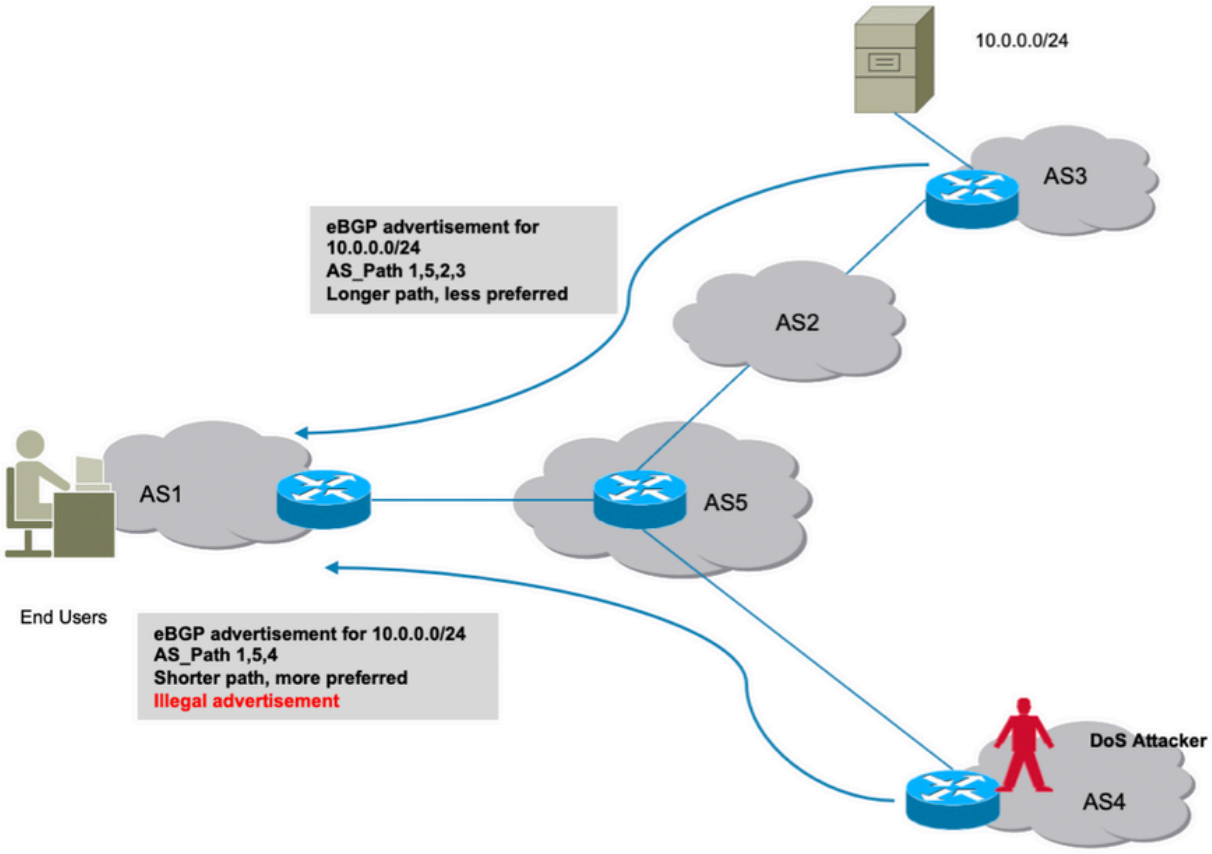
- 竊取IP以傳送垃圾郵件，導致IP被拒絕，從而造成拒絕服務。
- 監視流量以獲取密碼等敏感資訊。
- 管理員配置不正確導致的中斷。
- 通過安裝假伺服器來防止流量傳輸，從而確保拒絕服務。

拒絕服務攻擊（通常稱為DoS）是一種惡意嘗試，會中斷路由器、交換機、伺服器等的正常流量。DoS攻擊種類繁多，此處討論的很少。

路由劫持

請考慮此處顯示的情境。自治系統3(AS3)傳送其字首10.0.0.0/24的合法BGP通告。根據BGP的設計，BGP中沒有阻止攻擊者向網際網路通告相同字首的內容。

如圖所示，AS4中的攻擊者通告相同的字首10.0.0.0/24。BGP最佳路徑演演算法優先使用具有更短AS_Path的路徑。AS_Path 1,5,4通過AS 1,5,2,3贏得較長路徑。因此，來自客戶端的流量現在將被重定向到攻擊者的環境，並且可能會被黑洞，從而造成對終端客戶端的拒絕服務。

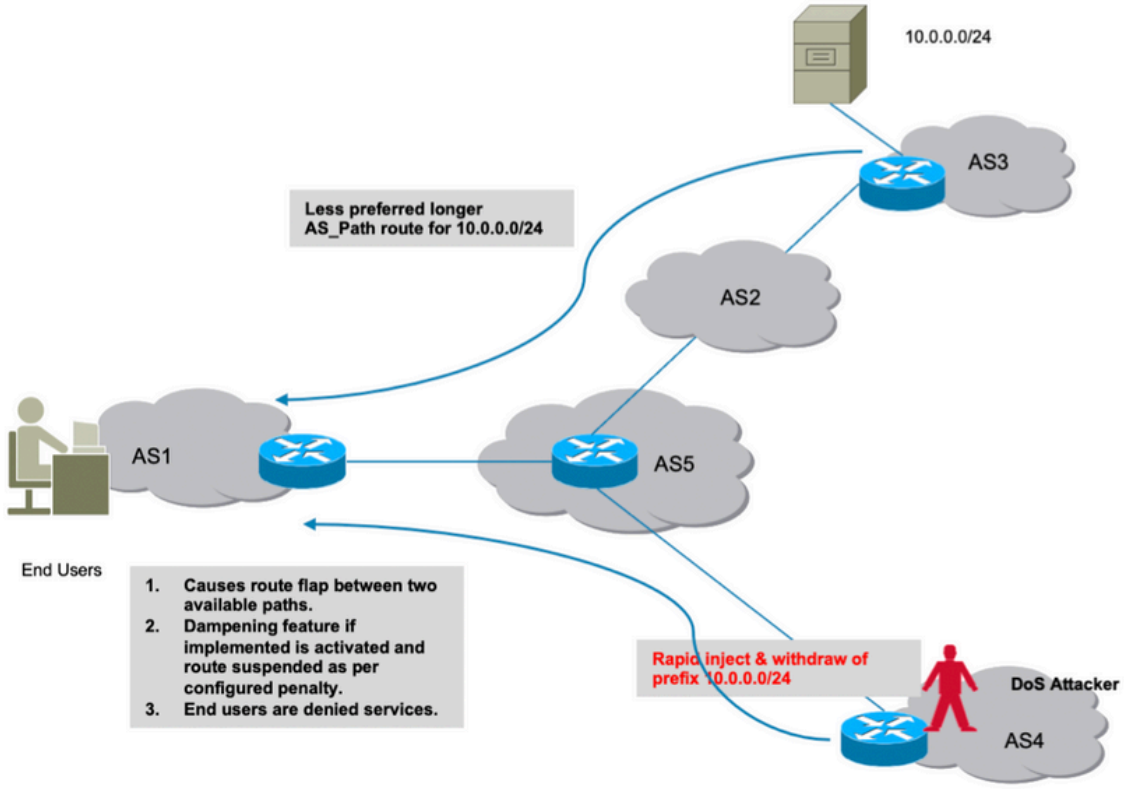


路由劫持

降低系統效能

本節討論另一種拒絕服務的方法。如果配置了Cisco的BGP路由抑制功能，則攻擊者可以在網路中引入快速路由擺動導致持續抖動時，利用此功能。

抑制功能將對合法路由實施處罰，使其無法用於實際流量。此外，這種不道德引起的抖動也會對路由器的CPU、記憶體等資源造成壓力。

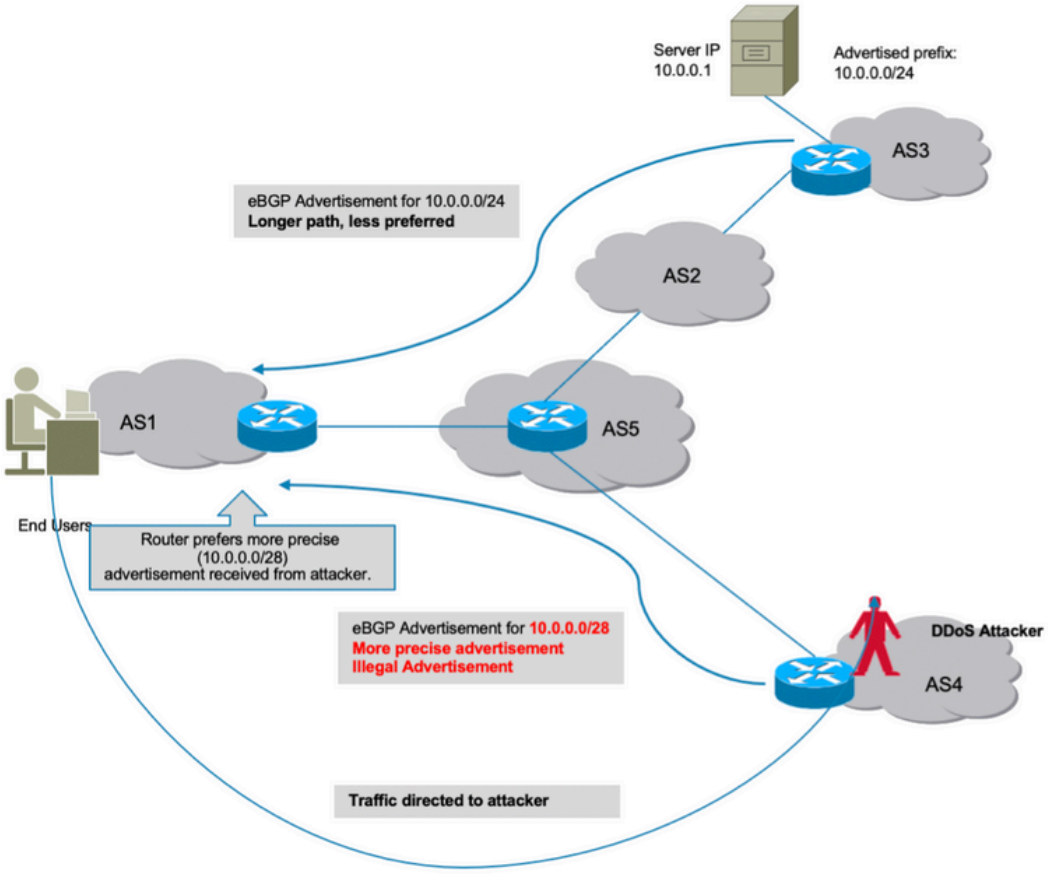


路由抑制

子首碼劫持

如上一節所述，攻擊者如何非法生成字首並造成流量中斷。不幸的是，顛覆並非唯一令人擔憂的原因。在此類攻擊中，實際資料可能會受到危害，其中攻擊者可以掃描接收到的資料以將其用於不道德用途。

同樣，劫持一條路線也可能通過非法宣傳一條更精確的路線來完成。BGP偏好匹配時間較長的首碼，且此行為可能會被錯誤利用，如下圖所示。



子字首劫持

所討論的所有攻擊都源於這樣一個事實：BGP無法識別這些惡意通告的字首的來源AS是否有效。要解決此問題，需要路由器能夠在其資料庫中保留的「true」和「trusted」資料來源。然後，每次收到新通告後，路由器現在能夠交叉驗證從BGP對等體接收的字首的AS源資訊和從驗證器接收的本地資料庫資訊。

因此，路由器能夠區分好通告和壞（非法）通告，並且路由器本身增加了避免前面討論的所有攻擊的能力。 BGP RPKI提供所需的可信資訊源。

RPKI

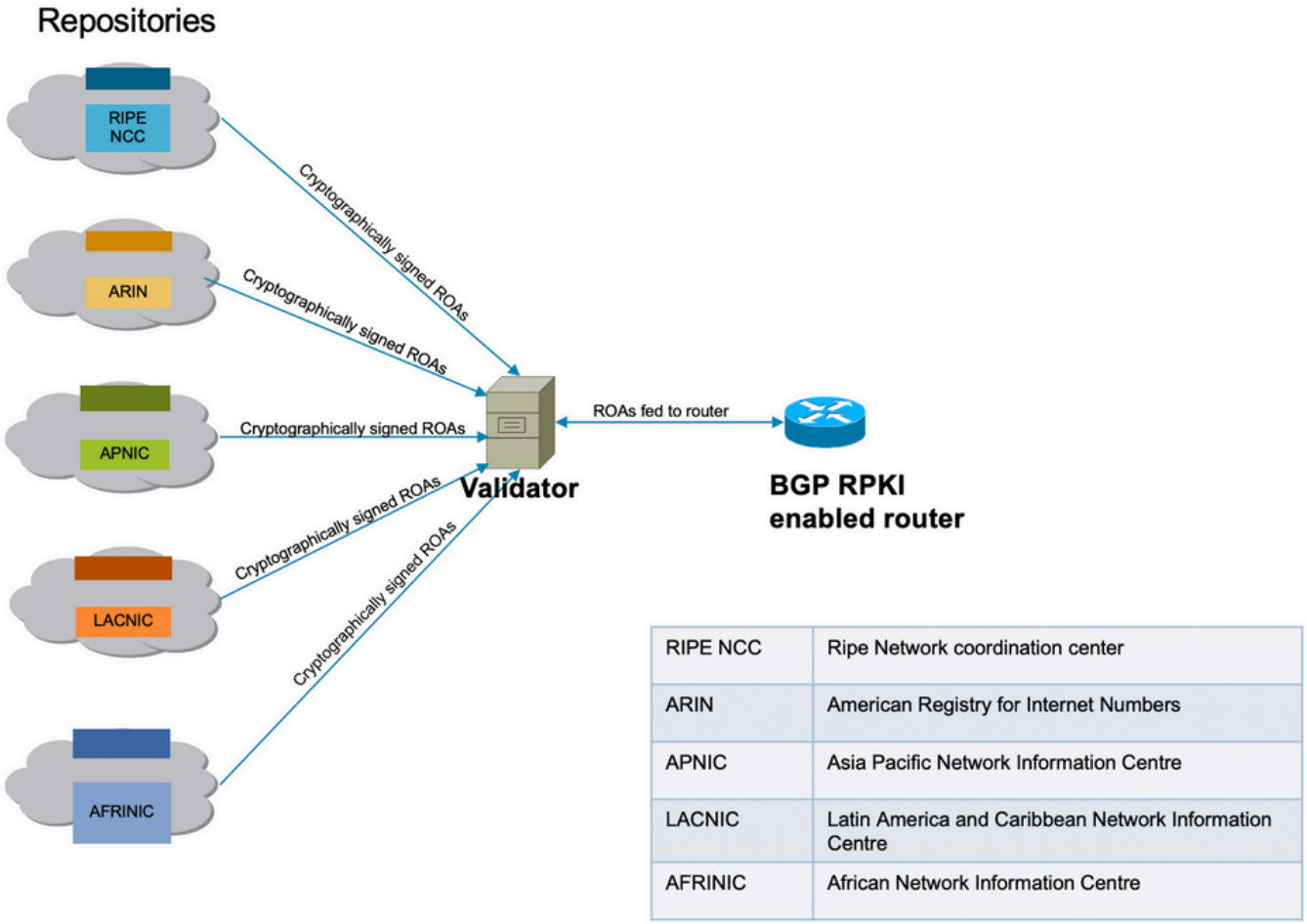
RPKI使用包含ROA的儲存庫。ROA包含有關首碼及其關聯的BGP AS編號的資訊。路由來源授權是一個加密簽名的語句。

5個地區網際網路註冊機構(RIR)是RPKI的信任錨點。Internet Assigned Numbers Authority(IANA)是派發IP字首的樹頂端。RIR是層次結構中的下一項。它們將子字首分配給本地網際網路註冊機構(LIR)和大型網際網路服務提供商(ISP)。他們為這些字首簽署證書。下一層分配這些證書的子字首，並使用上面的證書簽署自己的證書來認證他們自己的分配。它們通常使用自己的發佈點來託管證書和ROA。每個證書列出其簽名的子證書的發佈點。因此，RPKI形成了一個證書樹，該樹反映了IP地址分配樹。信賴方擁有的RPKI驗證器會輪詢所有發佈點，以查詢更新的證書和ROA（以及CRL和清單）。它們從信任錨點開始，並遵循到子證書發佈點的連結。

ROA通過RIR輸入到儲存庫，但也可通過其他登記處（國家或地方）進行相同操作。這項責任也可以委託給ISP，由RIR進行適當的監督和驗證。

目前，有五個由RIPE NCE、ARIN、APNIC、LACNIC和AFRINIC維護的ROA庫。

網路中存在的驗證器與這些儲存庫通訊，並下載受信任的ROA資料庫以構建其快取。這是RPKI的合併副本，定期從全域性RPKI直接或間接獲取/刷新。然後，驗證器將此資訊提供給路由器，使路由器能夠將傳入的BGP通告與RPKI表進行比較，以便做出安全決策。



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基礎設施連線

驗證器

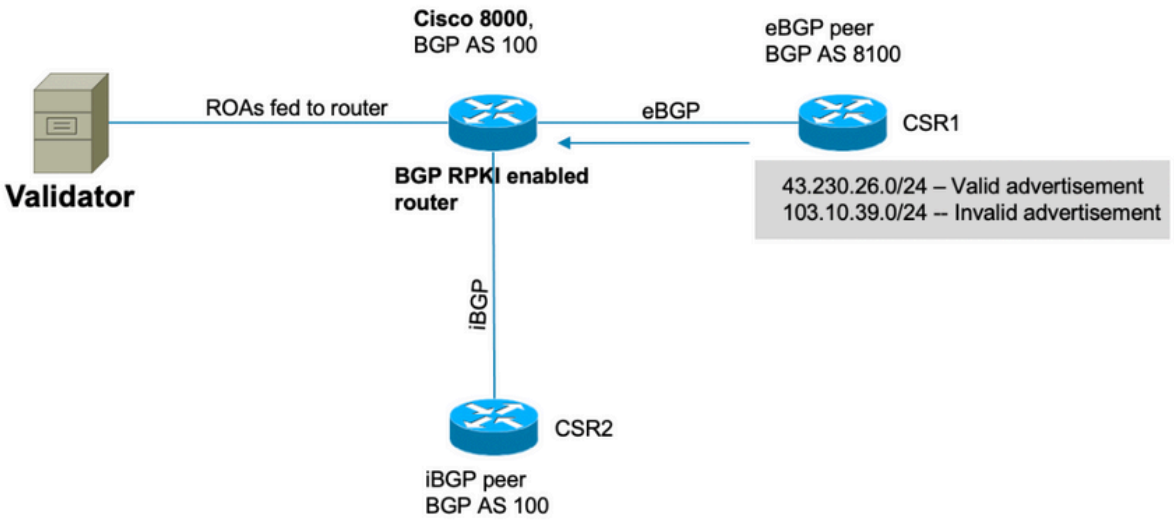
此演示使用RIPE驗證器。驗證器將通過建立TCP作業階段與路由器通訊。在本演示中，驗證器偵聽其IP 192.168.122.120和埠3323。

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

IANA為此通訊指定了埠3323。刷新計時器定義同步和更新本地儲存庫以保持更新的時間間隔。

BGP RPKI演示

拓撲



拓撲

> **注意**：此演示使用隨機公共AS編號和字首只是為了說明BGP RPKI機制。公共IP是使用，因為RPKI主要用於公共字首保護，而在RIR上建立的所有ROA都是公共字首。最後，本文檔中描述的任何操作、配置等都不會以任何方式影響這些公共IP和AS。

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會話

路由器與驗證器（IP: 192.168.122.120，埠3323）建立TCP會話，以便將ROA快取下載到路由器的記憶體中。

```
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下載

驗證器將ROA資訊提供給路由器。此快取記憶體會定期刷新，以便最小化路由器儲存過時資訊的可能性。在此演示中，已配置刷新時間900秒。 如圖所示，Cisco 8000路由器已從驗172632器下載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
```

驗證

本節展示BGP RPKI如何發揮作用，以及如何防止路由器發生錯誤/非法通告。

啟用Origin-As有效性

預設情況下，路由器從驗證器讀取ROA，但直到配置為使用ROA。因此，這些字首被標籤為「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 ?
```

為了啟用路由器進行來源有效性檢查，請為相關地址系列啟用此命令。

```
router bgp 100

 address-family ipv4 unicast

 bgp origin-as validation enable

 !
```

啟用此命令時，它會使路由器根據從驗證器接收的ROA資訊掃描其BGP表中存在的字首，並且三種狀態之一被分配給字首。

```
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 ?
```

要使路由器在進行最佳路徑計算時使用字首驗證狀態資訊，需要使用此命令。預設情況下，此功能不會啟用，因為它允許您不將有效性資訊用於最佳路徑計算，但仍允許您在本文檔稍後將討論的路由策略中使用它。

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as use validity

 !
```

字首有效性狀態

在三種狀態下可以找到字首。

```
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 ?
```

- 無效 — 表示字首滿足以下兩個條件之一：1。它匹配一個或多個**路由來源授權**(ROA)，但是如果來源AS與AS-PATH上的來源AS匹配，則沒有ROA匹配。2.它以在ROA中指定的最小長度匹配一個或多個ROA，但對於與最小長度匹配的所有ROA，其長度大於指定的最大長度。Origin AS與條件#2無關。
- 有效 — 表示在RPKI快取表中找到字首和AS對。
- Not Found — 表示字首不在有效或無效字首之列。

本節詳細討論每個字首及其狀態。

1. 203.0.113.0/24 — 有效

AS 8100中的eBGP對等體生成此路由並通告給Cisco8000節點。由於來源AS(8100)與ROA中的來源AS（從驗證器接收）相符，因此此首碼標示為有效，並安裝到路由器的路由表中。

```
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
```

該路由安裝在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
```

由於這是最佳的BGP首碼，且每個RPKI也有效，因此它已成功安裝到路由表中。

```
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 — 無效

此字首無效，因為ROA中包含的原始AS資訊和通過BGP消息從eBGP對等體接收的原始AS資訊之間存在衝突。203.0.113.1/24通過BGP接收，源地址為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 ?
```

但是，從驗證器收到的ROA顯示此字首屬於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
```

由於收到的BGP通告(AS 8100)中的AS源資訊與ROA(AS 10021)中收到的實際AS源資訊不匹配，因此字首被標籤為「無效」(Invalid)，並且未安裝在路由表中。

```
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

這是私人首碼，且不存在於ROA快取中。BGP將此字首宣告為「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
```

由於仍然採用RPKI，因此路由表中會安裝「not-found」字首。否則，BGP將忽略這些未在RPKI資料庫中註冊的合法字首。

允許無效字首

雖然不建議這樣做，但軟體確實提供了一個旋鈕，以允許無效字首參與最佳路徑計算演算法。

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as allow invalid

 !
```

透過此組態，路由器在標籤為「無效」時，確實會為最佳路徑計算考慮無效字首。此輸出顯示「203.0.113.1/24」被標籤為最佳路徑。

```
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 ?
```

如以下輸出所示，儘管字首保持無效，但字首仍被標籤為最佳。

```
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
```

請注意，路由器仍會將「無效」首碼視為最後一個選項，且如果可用的首碼，則總是會偏好有效首碼。

路由器上的手動ROA配置

如果由於某種原因，尚未建立、接收或延遲特定字首的ROA，則可以在路由器上配置手動ROA。例如，字首「192.168.122.1/32」標籤為「Not Found」，如此處所示。

```
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 ?
```

可以配置手動ROA，如下圖所示。此命令將「192.168.122.1/32」字首「與AS 8100關聯」。

```
router bgp 100

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

透過此組態，首碼的狀態會從「N」變更為「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 ?
```

路由策略和字首驗證狀態

字首狀態結果可用於建立路由策略。可以在match語句中使用這些狀態，並可執行管理員所需的操作。 此示例匹配具有無效狀態的所有字首，並為它們設定權重值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

 !

 !

!
```

此輸出顯示應用了無效的字首權重為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
```

通過擴展社群共用字首驗證資訊

因為BGP路由器還可以通過BGP擴展社群與其他路由器（驗證器沒有本地快取）共用字首驗證狀態。這節省了網路中每台路由器與驗證器會話以及下載所有ROA的開銷。

BGP延伸社群可以做到這一點。

此命令使路由器能夠與iBGP對等體共用「prefix-validation」資訊。

```
router bgp 100

 address-family ipv4 unicast

 bgp origin-as validation signal ibgp
```

如圖所示配置Cisco 8000路由器後，BGP更新對等體會包含字首驗證資訊。在這種情況下，鄰居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
```
可以使用0x4300 0x000（4位元組表示狀態）來瞭解此擴展社群對映。

指示狀態的四個位元組被視為32位無符號整數，它具有以下值之一：

- 0 — 有效
- 1 — 未找到
- 2 — 無效

字首203.0.113.1/24的社群是0x4300:0:2，該社群對映到「無效」字首。如此一來，即使路由器本身沒有本地快取，CSR2路由器仍能根據首碼驗證狀態做出決策。

現在，字首驗證狀態可用於在路由對映中或BGP最佳路徑演算法中進行匹配。

BGP RPKI實施建議

建立ROA的良好做法

這些是基於RPKI — 天文台觀測到的不可達網路的一些建議。RPKI觀測站從多個方面分析部署的RPKI環境。

- 如果為任何字首建立了ROA，則建議在BGP中通告該字首。如果沒有該名稱，則其他人可以通過簡單假裝是ROA中包含的ASN並使用字首來宣佈該名稱。
- 如果建立的ROA的maxlen大於字首長度，則其等效於為原始字首下一直到maxlen的所有可能字首建立ROA。強烈建議在BGP中通告所有這些字首。
- 如果為字首建立ROA，並且字首所有者宣佈原始字首的子字首，則ROA將使該子字首失效。子字首的ROA以及原始ROA的maxlen必須擴展以覆蓋子字首。
- 如果組織擁有字首，但計畫不在BGP中通告它，則必須為AS0的字首建立ROA。這將使任何字首通告失效，因為AS0不能出現在任何AS路徑中。
- 如果有多個ASN源自同一個字首，則必須為每個ASN建立該字首的ROA。因此，如果路由器具有同一字首的多個ROA，則與其中任何一個匹配的BGP通告將有效。同一字首的多個路由協定不會相互衝突。
- 如果「A」為其客戶「B」生成一個字首，並代表「B」為該字首建立一個ROA，則「A」必須在公告前加上「B」的ASN，或具有「B」生成字首本身。

RPKI對XR BGP路由器效能的影響

ROA更新對使用路由策略的CPU的影響

更新ROA時，如果路由器為包含「validation-state is」的鄰居提供本地入口路由策略，則根據新的ROA重新驗證字首的狀態非常重要。這通過路由器向其對等體傳送BGP REFRESH請求來實現。

當BGP鄰居收到此消息（如圖所示）時，鄰居再次傳送其字首，並且入站路由策略可以重新驗證傳入字首。

```
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
```
每當更新ROA時，當許多鄰居同時刷新時，問題就會放大。如果鄰居入站路由策略很複雜並且需要大量處理，則在ROA更新後幾分鐘內會導致CPU使用率較高。如果鄰居入站路由策略不包含「validation-state is」命令，則不會出現這些REFRESH消息。

如果為鄰居配置了「總是進行軟重新配置入站」，則不會傳送BGP REFRESH消息，但仍然會以相同的速率執行相同的路由策略，並且可以預期相同的CPU使用情況。

建議優先使用「bgp bestpath origin-as use validity」方法，而不是配置路由策略，原因如下6.2.2所述。

將ROA更新對CPU的影響降至最低

避免出現此處所說明問題的最佳方法是將**bestpath origin-as use validity** without **validation-state**列入策略。

```
router bgp 100

 address-family ipv4 unicast

 bgp bestpath origin-as use validity

 !
```
此命令保留路由器上接收到的無效路由，但阻止它成為最佳路徑。將不會安裝或進一步通告它。就好比扔下它一樣。如果下次ROA更新生效，則無需刷新，它自動符合最佳路徑的條件，無需執行策略。

如果使用者偏好允許「無效」首碼而不使用它們，則除了**bestpath origin-as use validity**外，使用組態**best path origin-as allow invalid**。

在這種情況下，當ROA更改時，自動更新最佳路徑，而不需要REFRESH消息。為了取消首選，路由意味著在BGP路由選擇期間，RPKI無效路徑被視為比通向相同目的地的任何其他路徑更不優選。這類似於為其分配權重或小於0的本地首選項。

RPKI無效的數量相對較少，並且保留在表中不會對資源產生重大影響。

> **注意**：為了使用「bestpath origin-as use validity」，路由的所有路徑（包括IBGP路徑）都必須具有正確的RPKI有效性。如果不是，則仍可使用route-policy中的validation-state測試。

路由器不會根據ROA資料庫驗證IBGP路由。IBGP路由從RPKI擴展社群獲得RPKI有效性。如果收到IBGP路由時沒有使用此擴展社群，則其validation-state設定為not-found。

BGP RPKI記憶體空間

每個ROA消耗索引和資料記憶體。如果兩個ROA具有相同的IP字首，但具有不同的max_len，或者從不同的RPKI伺服器接收，則它們共用相同的索引但具有不同的資料。記憶體要求可能有所不同，因為記憶體開銷不是固定的。建議超支10%。與32位平台相比，64位平台要求每個記憶體對象擁有更多記憶體。表中列出了索引對象和資料對象的IOS-XR記憶體使用情況（以位元組為單位）。這些數字中包括一些基本恆定的開銷。

	32位平台（位元組）	64位平台（位元組）
IPv4索引	74	111
IPv6索引	86	125
資料	34	53

本節採用兩種方案來解釋ROA如何使用記憶體。

案例 1.路由器上配置的三台RPKI伺服器

考慮使用3台RPKI伺服器的路由器，每台在64位路由處理器上提供200,000個IPv4 ROA和20,000個IPv6 ROA將需要此記憶體：

20000 *(125 + 3*53)+ 200000 *(111 + 3*53)位元組= 59.68百萬位元組

在計算記憶體時，來自三個不同驗證器的相同字首的ROA共用相同的索引值。

案例 2.在路由器上配置單個RPKI伺服器

不帶ROA的BGP進程記憶體：

```
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
```

在未使用任何ROA的情況下，BGP進程佔用25 MB記憶體。

使用ROA的BGP進程記憶體：

```
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
```

在未使用任何ROA的情況下，BGP進程佔用25 MB記憶體。

使用ROA的BGP進程記憶體：

```
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路由器運行64位作業系統。它接172796了IPv4 ROA和28411 ROA。

記憶體（位元組）= 172,796 x [111（索引）+ 53（資料）] + 28411 x [125（索引）+ 53（資料）]。

這些計算得出的大小約為27 MB，大約是上述路由器記憶體中注意到的增量。

關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。