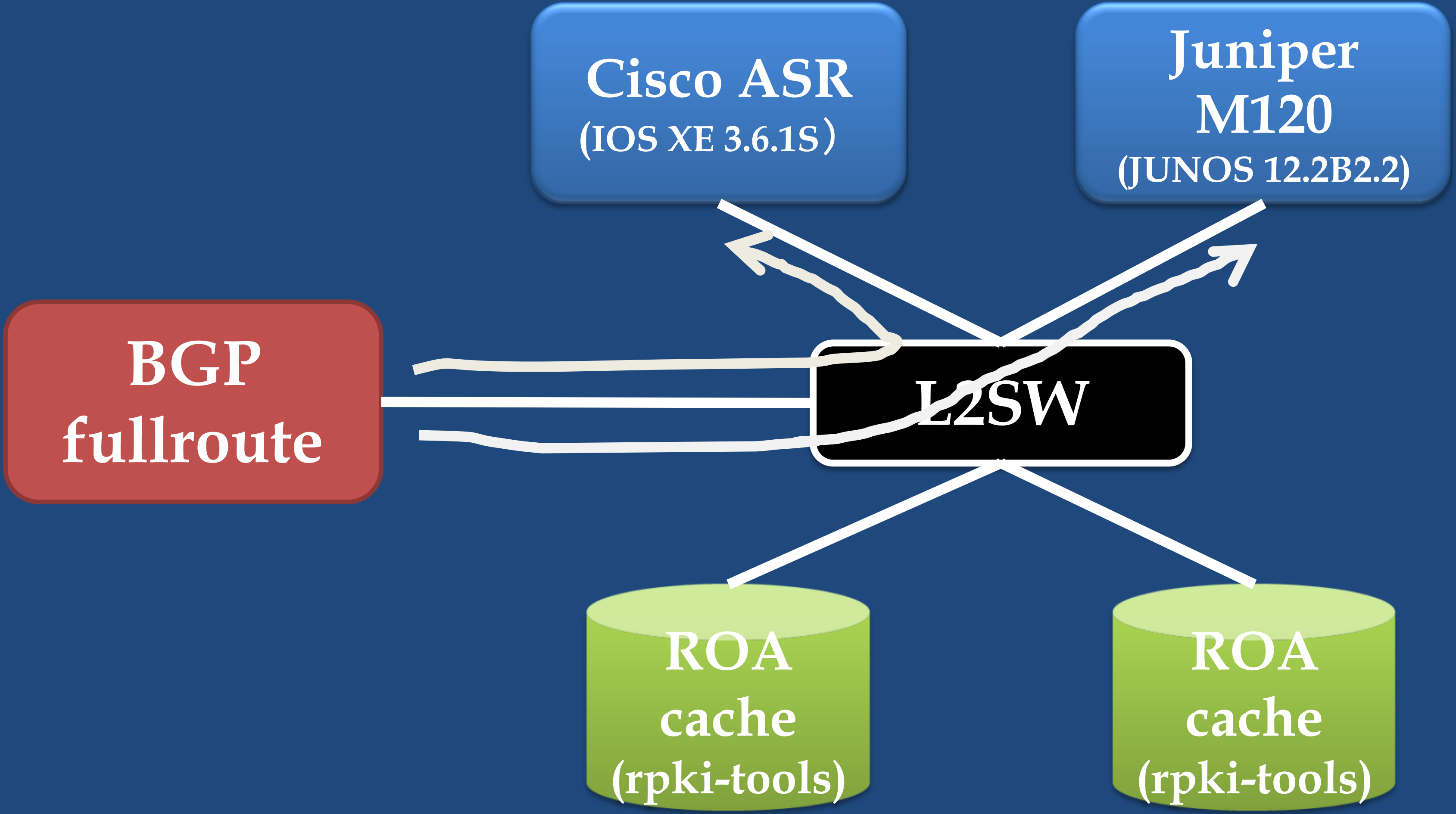
RPKI experience report from Japan

2012/8/29

JPNAP (Internet Multifeed)
Tomoya Yoshida

Test environment



Cisco ASR
(IOS XE 3.6.1S）

Juniper
M120
(JUNOS 12.2B2.2)

BGP
fullroute

L2SW

ROA
cache
(rpki-tools)

ROA
cache
(rpki-tools)

copyright (c) tomo

Brief Result

- Basic implementation was OK
  - validation result : OK
  - could eat 450K(v4)+10K(v6) ROAs
- Improvement will be needed step by step
  - CLI output, need more operational command
- iBGP propagation
  - couldn’t share the rpki status btw C and J
- Should check your OS when you use the rpki function

copyright (c) tomo

What Router dose

1. Storing ROA cache using RTR protocol

2. Validate with ROA cache data inside the router

3. Propagate RPKI validation result with iBGP for other router

copyright (c) tomo

What Router dose

1. Storing ROA cache using RTR protocol

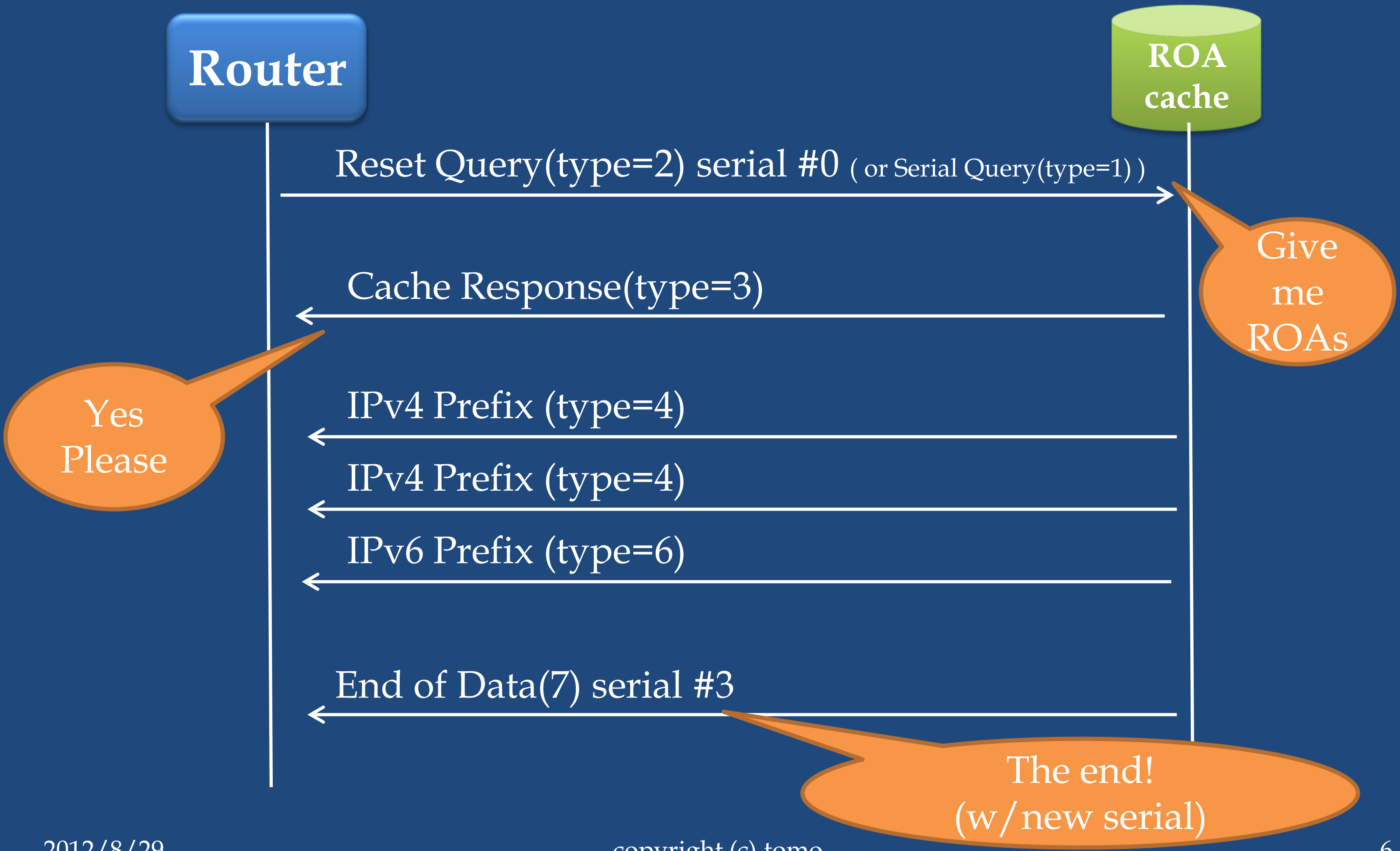
2. Validate with ROA cache data inside the router

3. Propagate RPKI validation result with iBGP for other router

copyright (c) tomo

1. RTR(RPKI/Router) Protocol

Start or Restart



Router

ROA cache

Reset Query(type=2) serial #0 (or Serial Query(type=1))

Give me ROAs

Cache Response(type=3)

Yes Please

IPv4 Prefix (type=4)

IPv4 Prefix (type=4)

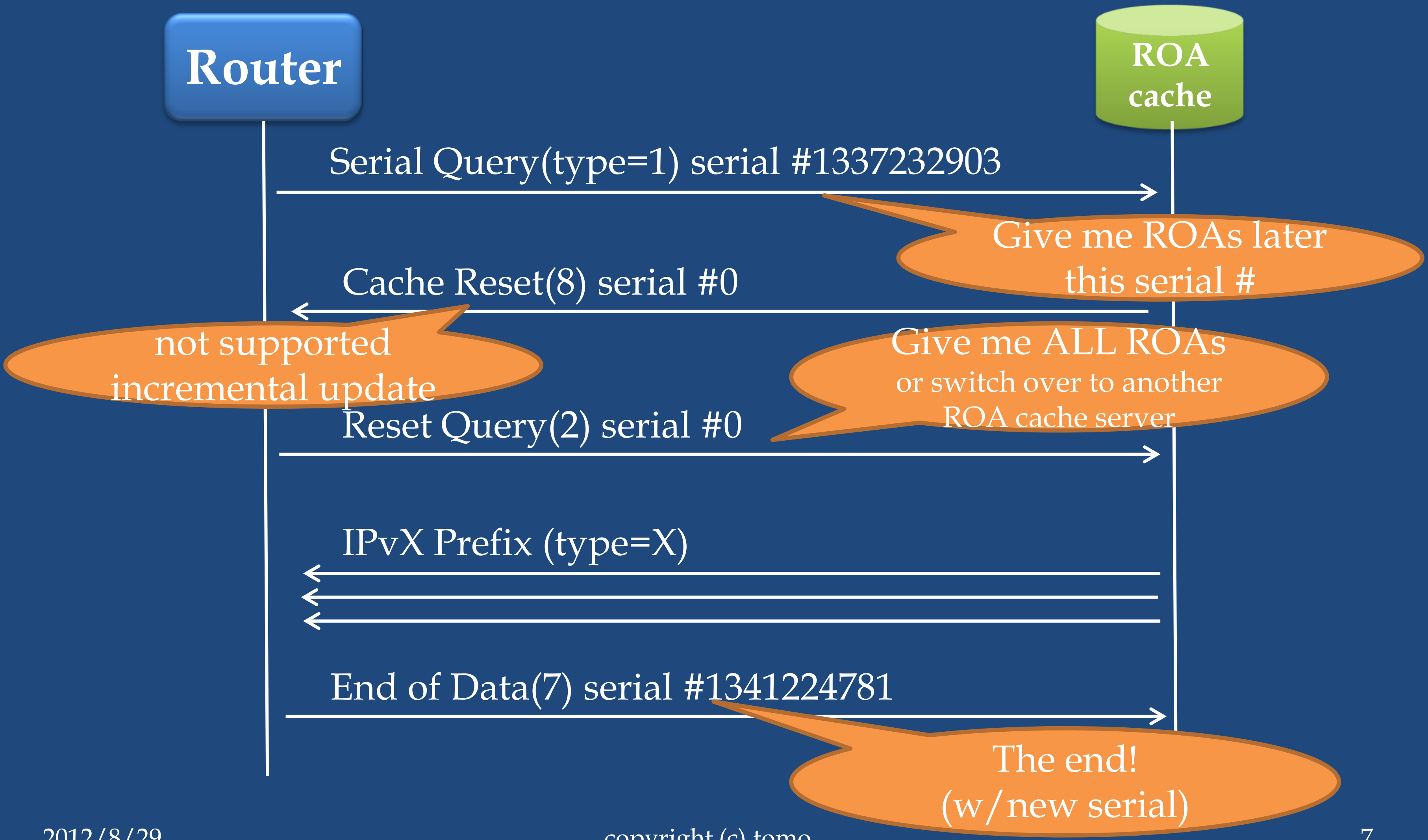
IPv6 Prefix (type=6)

End of Data(7) serial #3

The end!
(w/new serial)

1. RTR(RPKI/Router) Protocol

update (no incremental update)



Router

ROA cache

Serial Query(type=1) serial #1337232903

Give me ROAs later this serial #

Cache Reset(8) serial #0

not supported incremental update

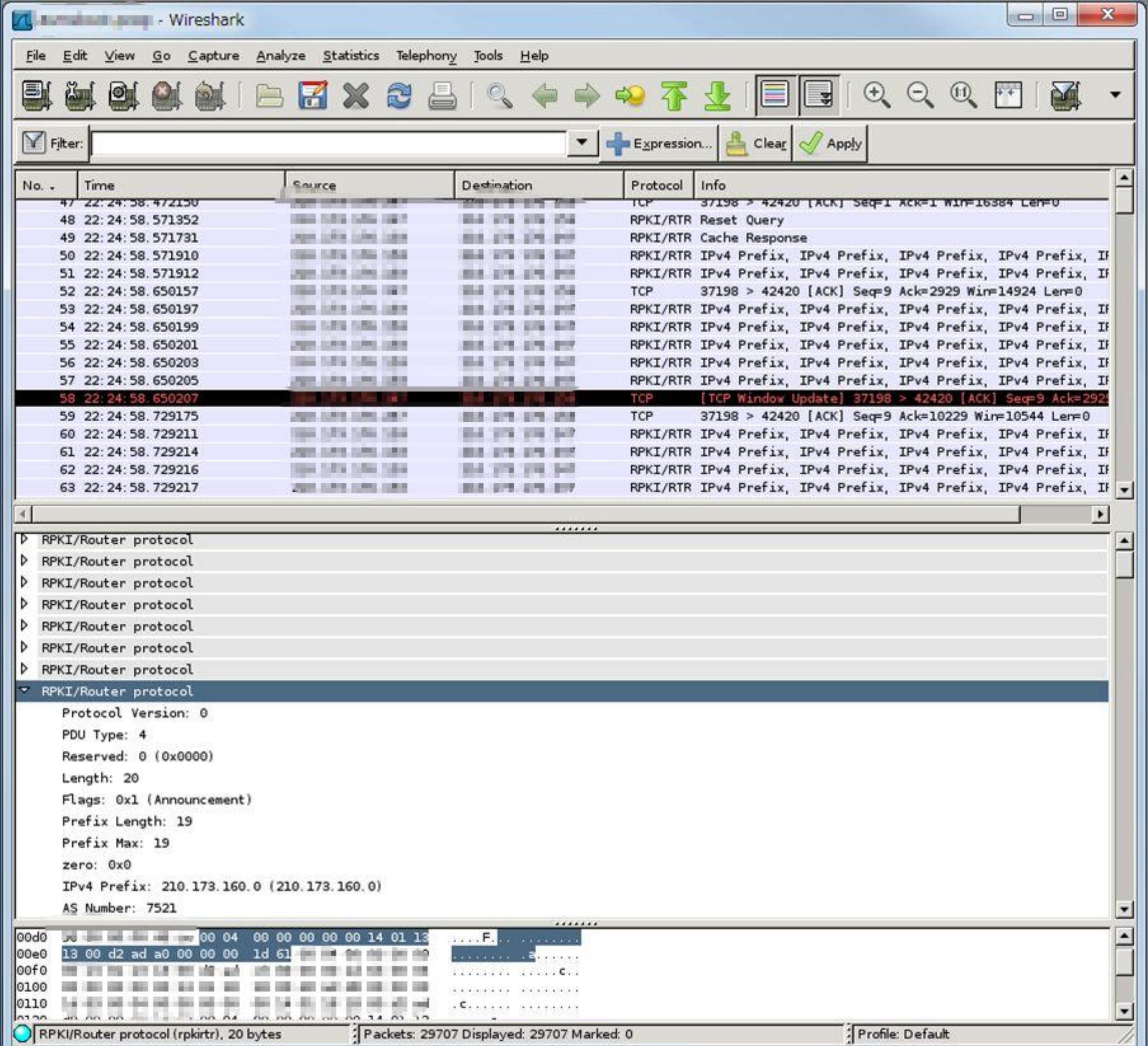
Give me ALL ROAs
or switch over to another ROA cache server

Reset Query(2) serial #0

IPvX Prefix (type=X)

End of Data(7) serial #1341224781

The end!
(w/new serial)



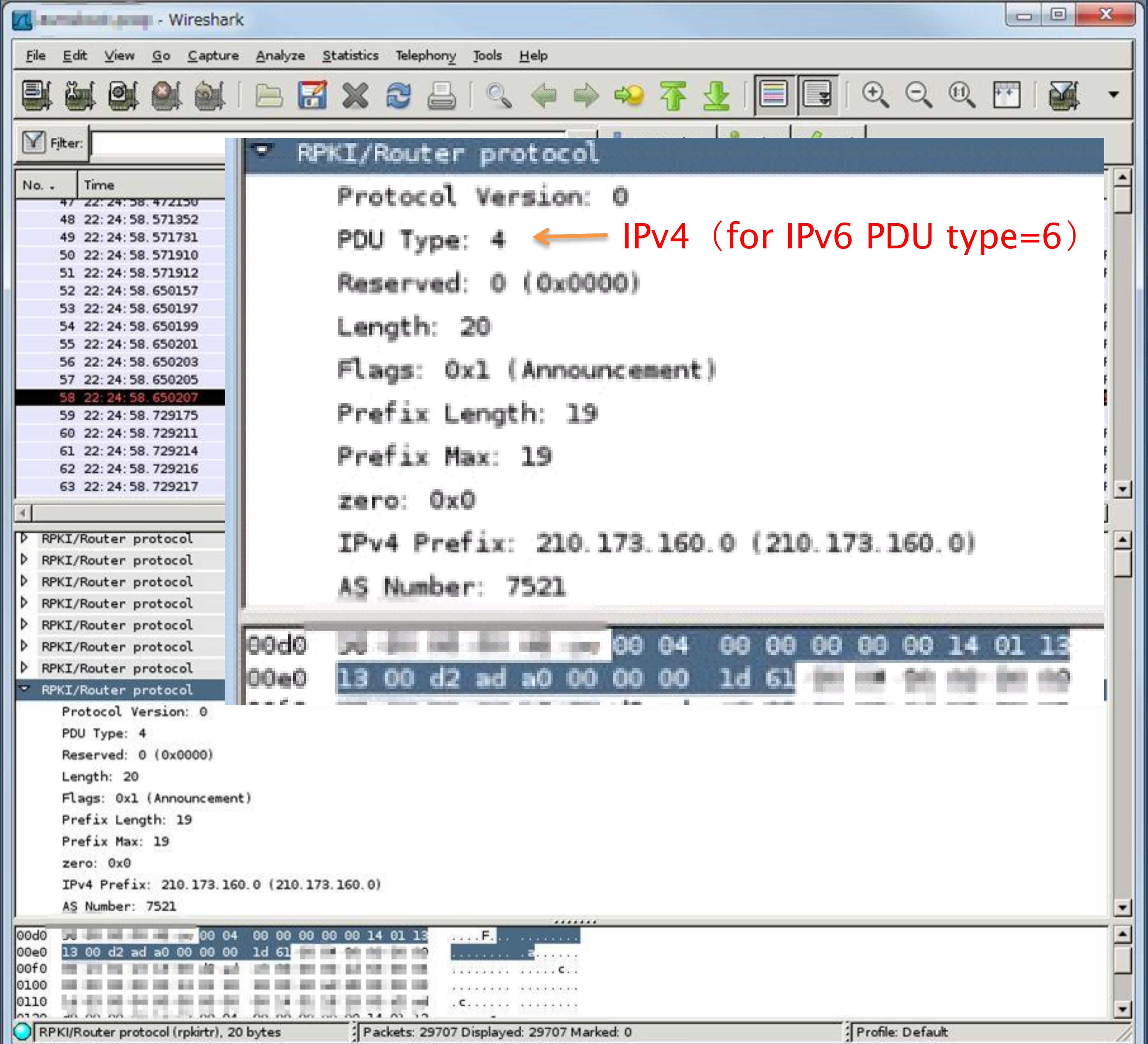
[redacted] - Wireshark
File Edit View Go Capture Analyze Statistics Telephony Tools Help
Filter:
Expression...
Clear
Apply

```
No. .  Time             Source      Destination  Protocol  Info
  47   22:24:58.472150  [redacted]  [redacted]   TCP       37198 > 42420 [ACK] Seq=1 Ack=1 Win=16384 Len=0
  48   22:24:58.571352  [redacted]  [redacted]   RPKI/RTR  Reset Query
  49   22:24:58.571731  [redacted]  [redacted]   RPKI/RTR  Cache Response
  50   22:24:58.571910  [redacted]  [redacted]   RPKI/RTR  IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IP
  51   22:24:58.571912  [redacted]  [redacted]   RPKI/RTR  IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IP
  52   22:24:58.650157  [redacted]  [redacted]   TCP       37198 > 42420 [ACK] Seq=9 Ack=2929 Win=14924 Len=0
  53   22:24:58.650197  [redacted]  [redacted]   RPKI/RTR  IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IP
  54   22:24:58.650199  [redacted]  [redacted]   RPKI/RTR  IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IP
  55   22:24:58.650201  [redacted]  [redacted]   RPKI/RTR  IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IP
  56   22:24:58.650203  [redacted]  [redacted]   RPKI/RTR  IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IP
  57   22:24:58.650205  [redacted]  [redacted]   RPKI/RTR  IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IP
  58   22:24:58.650207  [redacted]  [redacted]   TCP       [TCP Window Update] 37198 > 42420 [ACK] Seq=9 Ack=292
  59   22:24:58.729175  [redacted]  [redacted]   TCP       37198 > 42420 [ACK] Seq=9 Ack=10229 Win=10544 Len=0
  60   22:24:58.729211  [redacted]  [redacted]   RPKI/RTR  IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IP
  61   22:24:58.729214  [redacted]  [redacted]   RPKI/RTR  IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IP
  62   22:24:58.729216  [redacted]  [redacted]   RPKI/RTR  IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IP
  63   22:24:58.729217  [redacted]  [redacted]   RPKI/RTR  IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IPv4 Prefix, IP
```

```
▷ RPKI/Router protocol
▷ RPKI/Router protocol
▷ RPKI/Router protocol
▷ RPKI/Router protocol
▷ RPKI/Router protocol
▷ RPKI/Router protocol
▷ RPKI/Router protocol
▽ RPKI/Router protocol
    Protocol Version: 0
    PDU Type: 4
    Reserved: 0 (0x0000)
    Length: 20
    Flags: 0x1 (Announcement)
    Prefix Length: 19
    Prefix Max: 19
    zero: 0x0
    IPv4 Prefix: 210.173.160.0 (210.173.160.0)
    AS Number: 7521
```

```
00d0  [redacted] 00 04  00 00 00 00 00 14 01 13   ....F... ........
00e0  13 00 d2 ad a0 00 00 00  1d 61 [redacted]   ........ .a......
00f0  [redacted]                                  ........ ......c..
0100  [redacted]                                  ........ ........
0110  [redacted]                                  .c...... ........
```

RPKI/Router protocol (rpkirtr), 20 bytes | Packets: 29707 Displayed: 29707 Marked: 0 | Profile: Default


[redacted] - Wireshark
File Edit View Go Capture Analyze Statistics Telephony Tools Help
Filter:
No.
Time

```
47  22:24:58.472150
48  22:24:58.571352
49  22:24:58.571731
50  22:24:58.571910
51  22:24:58.571912
52  22:24:58.650157
53  22:24:58.650197
54  22:24:58.650199
55  22:24:58.650201
56  22:24:58.650203
57  22:24:58.650205
58  22:24:58.650207
59  22:24:58.729175
60  22:24:58.729211
61  22:24:58.729214
62  22:24:58.729216
63  22:24:58.729217
```

```
RPKI/Router protocol
    Protocol Version: 0
    PDU Type: 4
    Reserved: 0 (0x0000)
    Length: 20
    Flags: 0x1 (Announcement)
    Prefix Length: 19
    Prefix Max: 19
    zero: 0x0
    IPv4 Prefix: 210.173.160.0 (210.173.160.0)
    AS Number: 7521

00d0  [redacted] 00 04  00 00 00 00 00 14 01 13
00e0  13 00 d2 ad a0 00 00 00  1d 61 [redacted]
```

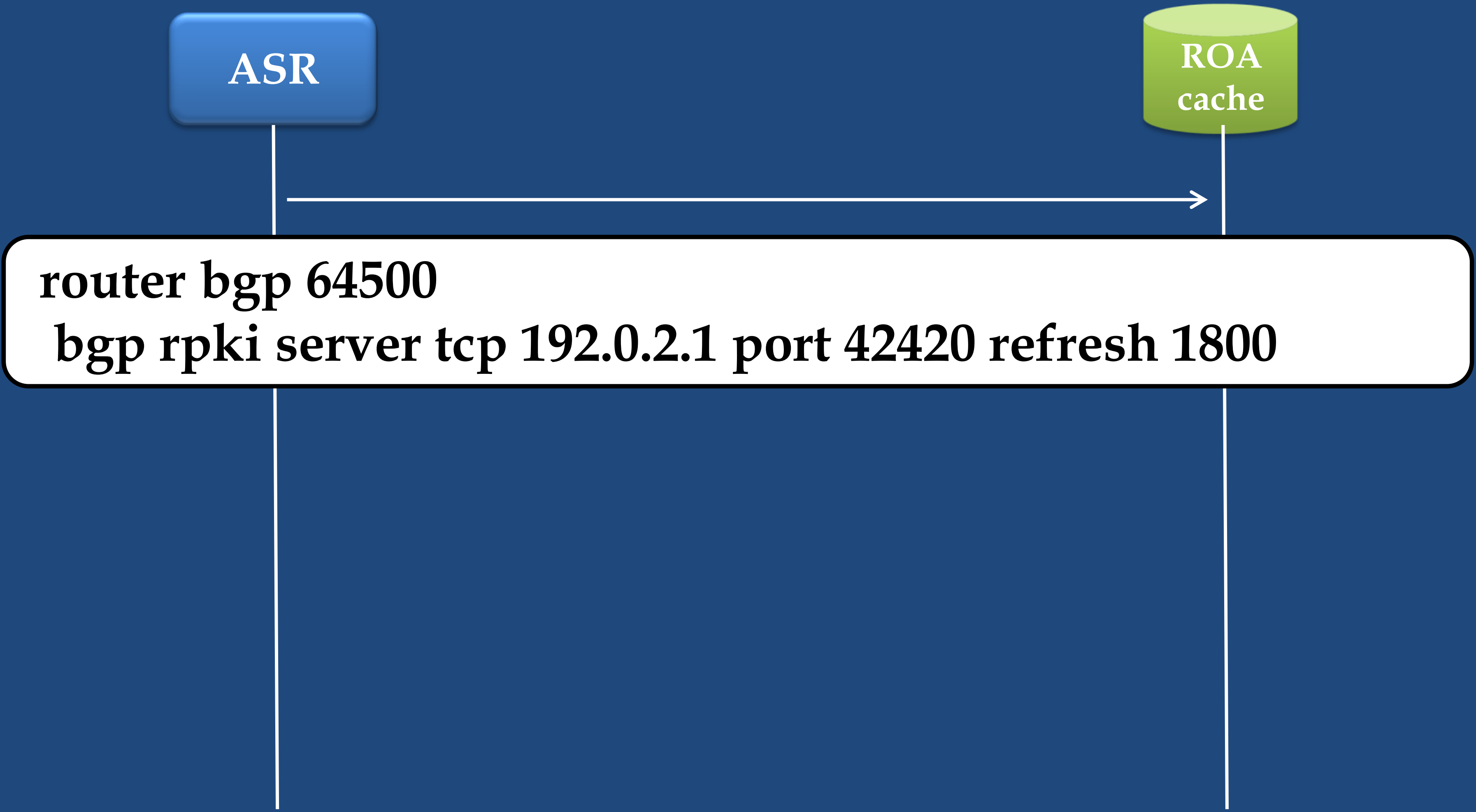
IPv4 （for IPv6 PDU type=6）

```
▷ RPKI/Router protocol
▷ RPKI/Router protocol
▷ RPKI/Router protocol
▷ RPKI/Router protocol
▷ RPKI/Router protocol
▷ RPKI/Router protocol
▷ RPKI/Router protocol
▽ RPKI/Router protocol
    Protocol Version: 0
    PDU Type: 4
    Reserved: 0 (0x0000)
    Length: 20
    Flags: 0x1 (Announcement)
    Prefix Length: 19
    Prefix Max: 19
    zero: 0x0
    IPv4 Prefix: 210.173.160.0 (210.173.160.0)
    AS Number: 7521
```

```
00d0  [redacted] 00 04  00 00 00 00 00 14 01 13   ....F.. ........
00e0  13 00 d2 ad a0 00 00 00  1d 61 [redacted]   ........ .a......
00f0  [redacted]                                  ........ .....c..
0100  [redacted]                                  ........ ........
0110  [redacted]                                  .c...... ........
```

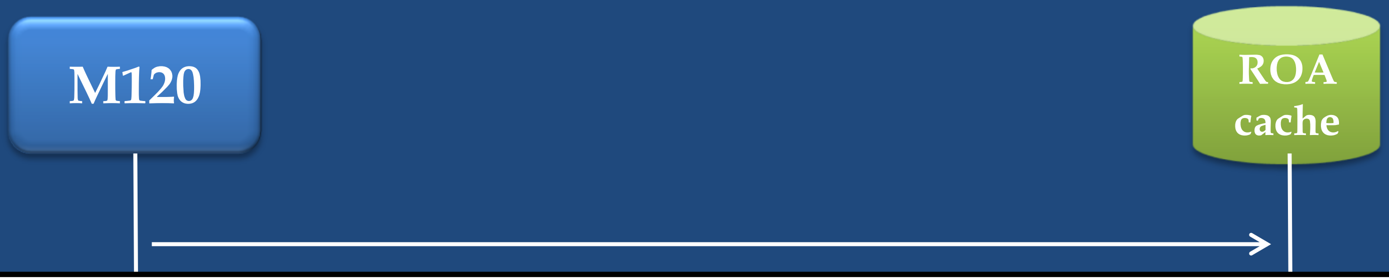
RPKI/Router protocol (rpkirtr), 20 bytes | Packets: 29707 Displayed: 29707 Marked: 0 | Profile: Default

1. RTR(RPKI/Router) Protocol



```
router bgp 64500
 bgp rpki server tcp 192.0.2.1 port 42420 refresh 1800
```

1. RTR(RPKI/Router) Protocol



```
routing-options {
 validation {
 group ROA {
          session 192.0.2.1 {
              refresh-time 1800;
              port 42420;
              local-address 192.0.2.13;
          }
      }
}
```

copyright (c) tomo

ASR

```
asr>show bgp ipv4 unicast rpki table
Load for five secs: 0%/0%; one minute: 1%; five minutes: 1%
Time source is NTP, 11:26:41.011 JST Fri Jul 6 2012
452768 BGP sovc network entries using 72442880 bytes of memory
455551 BGP sovc record entries using 14577632 bytes of memory

Network          Maxlen  Origin-AS  Source  Neighbor
1.0.0.0/24       24      15169      0       192.0.2.1/42420
1.0.4.0/22       22      56203      0       192.0.2.1/42420
1.0.16.0/23      23      2519       0       192.0.2.1/42420
1.0.18.0/23      23      2519       0       192.0.2.1/42420
1.0.20.0/23      23      2519       0       192.0.2.1/42420
1.0.22.0/23      23      2519       0       192.0.2.1/42420
1.0.24.0/24      24      2519       0       192.0.2.1/42420
1.0.24.0/23      23      2519       0       192.0.2.1/42420
1.0.25.0/24      24      2519       0       192.0.2.1/42420
1.0.26.0/24      24      2519       0       192.0.2.1/42420
```

ASR

```
asr>show bgp ipv6 unicast rpki table
Load for five secs: 2%/0%; one minute: 6%; five minutes: 2%
Time source is NTP, 10:56:34.272 JST Fri Jul 6 2012
9851 BGP sovc network entries using 1812584 bytes of memory
9932 BGP sovc record entries using 317824 bytes of memory

Network              Maxlen  Origin-AS  Source  Neighbor
2001::/32            32      1101       0       192.0.2.1/42420
2001::/32            32      6939       0       192.0.2.1/42420
2001::/32            32      12859      0       192.0.2.1/42420
2001:200::/32        32      2500       0       192.0.2.1/42420
2001:200:900::/40    40      7660       0       192.0.2.1/42420
2001:200:905::/48    48      56218      0       192.0.2.1/42420
2001:200:C00::/40    40      7530       0       192.0.2.1/42420
2001:200:C000::/35   35      23634      0       192.0.2.1/42420
2001:200:E000::/35   35      7660       0       192.0.2.1/42420
```

ASR has only..

```
asr>show ip bgp rpki ?
  servers  Display RPKI cache server information
  table    Display RPKI table entries
```

M120

```
m120> show validation session
Session                    State  Flaps   Uptime #IPv4/IPv6 records
192.0.2.1                   Up      0    00:45:05 455550/9931
192.0.2.2                   Up      0    17:25:52 1/1
```

Very similar to BGP CLI

copyright (c) tomo

M120

```
m120> show validation database session 192.0.2.2
RV database for instance master

Prefix              Origin-AS Session                    State   Mismatch
210.173.160.0/19-24      7521 192.0.2.2                  valid
2001:3a0::/32-64         7521 192.0.2.2                  valid

  IPv4 records: 1
  IPv6 records: 1
```

M120

```
m120> show validation database session 192.0.2.1
RV database for instance master

Prefix           Origin-AS Session               State   Mismatch
1.0.0.0/24-24        15169 192.0.2.1                valid
1.0.4.0/22-22        56203 192.0.2.1                valid
1.0.16.0/23-23        2519 192.0.2.1                valid
1.0.18.0/23-23        2519 192.0.2.1                valid
1.0.20.0/23-23        2519 192.0.2.1                valid
1.0.22.0/23-23        2519 192.0.2.1                valid
1.0.24.0/23-23        2519 192.0.2.1                valid
 ・ ・ ・
2001::/32-32          1101 192.0.2.1                valid
2001::/32-32          6939 192.0.2.1                valid
```

IPv6 ROAs comes after the IPv4… better to be shown per address-family

M120

```
m120> show validation database origin-autonomous-system 7521
RV database for instance master

Prefix              Origin-AS Session                State  Mismatch
210.173.160.0/19-19       7521 192.0.2.1              valid
210.173.160.0/19-24       7521 192.0.2.2              valid
2001:3a0::/32-32          7521 192.0.2.1              valid
2001:3a0::/32-64          7521 192.0.2.2              valid

  IPv4 records: 2
  IPv6 records: 2
```

can search by origin ASN, GOOD!

What Router dose

1. Storing ROA cache using RTR protocol

2. Validate with ROA cache data inside the router

3. Propagate RPKI validation result with iBGP for other router

Validation status

Valid

Invalid

Not Found

2. Validation (ASR)

```
router bgp 64500
 address-family ipv4
  bgp bestpath prefix-validate allow-invalid

route-map rpki permit 10
 match rpki invalid
 set community 65400:2 additive
!
route-map rpki permit 20
 match rpki not-found
 set community 65400:1 additive
!
route-map rpki permit 30
 match rpki valid
 set community 65400:0 additive
```

Reflect the invalid route
to the bestpath selection

2. Validation (M120)

```
protocols {
  bgp {
    group RPKI-fullroute {
      neighbor 192.0.2.254 {
        import validation;
        peer-as 131079;
      }
      neighbor 2001:7fa:7:1:0:13:1079:1 {
        import validation;
        peer-as 131079;
      }
    }
```

2. Validation (M120)

```
policy-statement validation {
      term valid {
         from {
            protocol bgp;
            validation-database valid;
         }
         then {
            validation-state valid;
            community set rpki-valid;
            community add origin-validation-state-valid;
            accept;
         }
      }
      term invalid {
         from {
            protocol bgp;
            validation-database invalid;
         }
         then {
            validation-state invalid;
            community set rpki-invalid;
            community add origin-validation-state-invalid;
            accept;
         }
      }
```

```
term unknown {
         from {
            protocol bgp;
            validation-database unknown;
         }
         then {
            validation-state unknown;
            community set rpki-unknown;
            community add origin-validation-state-unknown;
            accept;
         }
      }
   }

 community origin-validation-state-invalid members
0x43:65400:2;
   community origin-validation-state-unknown members
0x43:65400:1;
   community origin-validation-state-valid members
0x43:65400:0;
   community rpki-invalid members 65400:2;
   community rpki-unknown members 65400:1;
   community rpki-valid members 65400:3;
```

ASR(IPv4)

```
asr>show ip bgp 210.173.160.0/19
Load for five secs: 0%/0%; one minute: 1%; five minutes: 1%
Time source is NTP, 10:54:10.058 JST Fri Jul 6 2012
BGP routing table entry for 210.173.160.0/19, version 938277
Paths: (2 available, best #1, table default)
  Not advertised to any peer
  Refresh Epoch 1
  131079 7521
    192.0.2.254 from 192.0.2.254 (210.173.172.118)
      Origin IGP, localpref 100, valid, external, best
      Community: 65400:1
      path 7FC93CD9B9C0 RPKI State valid
  Refresh Epoch 1
  131079 7521, (received-only)
    192.0.2.254 from 192.0.2.254 (210.173.172.118)
      Origin IGP, localpref 100, valid, external
      path 7FC93CD9B958 RPKI State valid
```

copyright (c) tomo

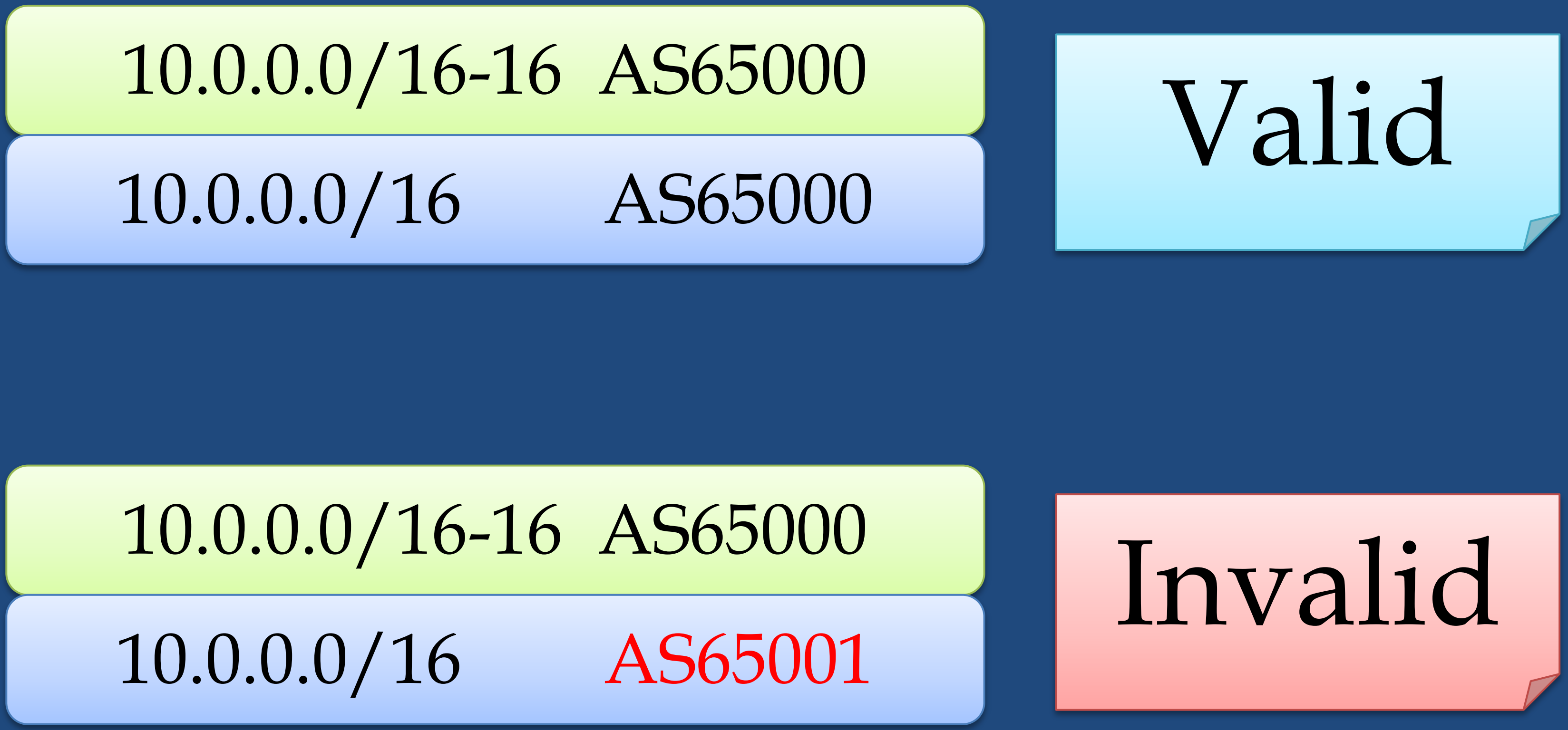
ASR(IPv6)

```
asr>show bgp ipv6 uni 2001:3a0::/32
Load for five secs: 0%/0%; one minute: 1%; five minutes: 1%
Time source is NTP, 10:52:19.542 JST Fri Jul 6 2012
BGP routing table entry for 2001:3A0::/32, version 16909
Paths: (2 available, best #2, table default)
  Advertised to update-groups:
     2
  Refresh Epoch 1
  131079 7521
    2001:7FA:7:1:0:13:1079:1 from 2001:7FA:7:1::250:1 (210.173.161.247)
      Origin IGP, localpref 100, valid, internal
      Community: 65400:3
      Extended Community: 0x43:65400:0
      path 7FC93120BC28 RPKI State valid
  Refresh Epoch 1
  131079 7521, (received & used)
    2001:7FA:7:1:0:13:1079:1 (FE80::205:85FF:FE16:C001) from 2001:7FA:7:1:0:13:1079:1
       (210.173.172.118)
      Origin IGP, localpref 100, valid, external, best
      path 7FC931062E08 RPKI State valid
```

copyright (c) tomo

Validation Result（1/7）

ROA
BGP



10.0.0.0/16-16 AS65000

10.0.0.0/16 AS65000

Valid

10.0.0.0/16-16 AS65000

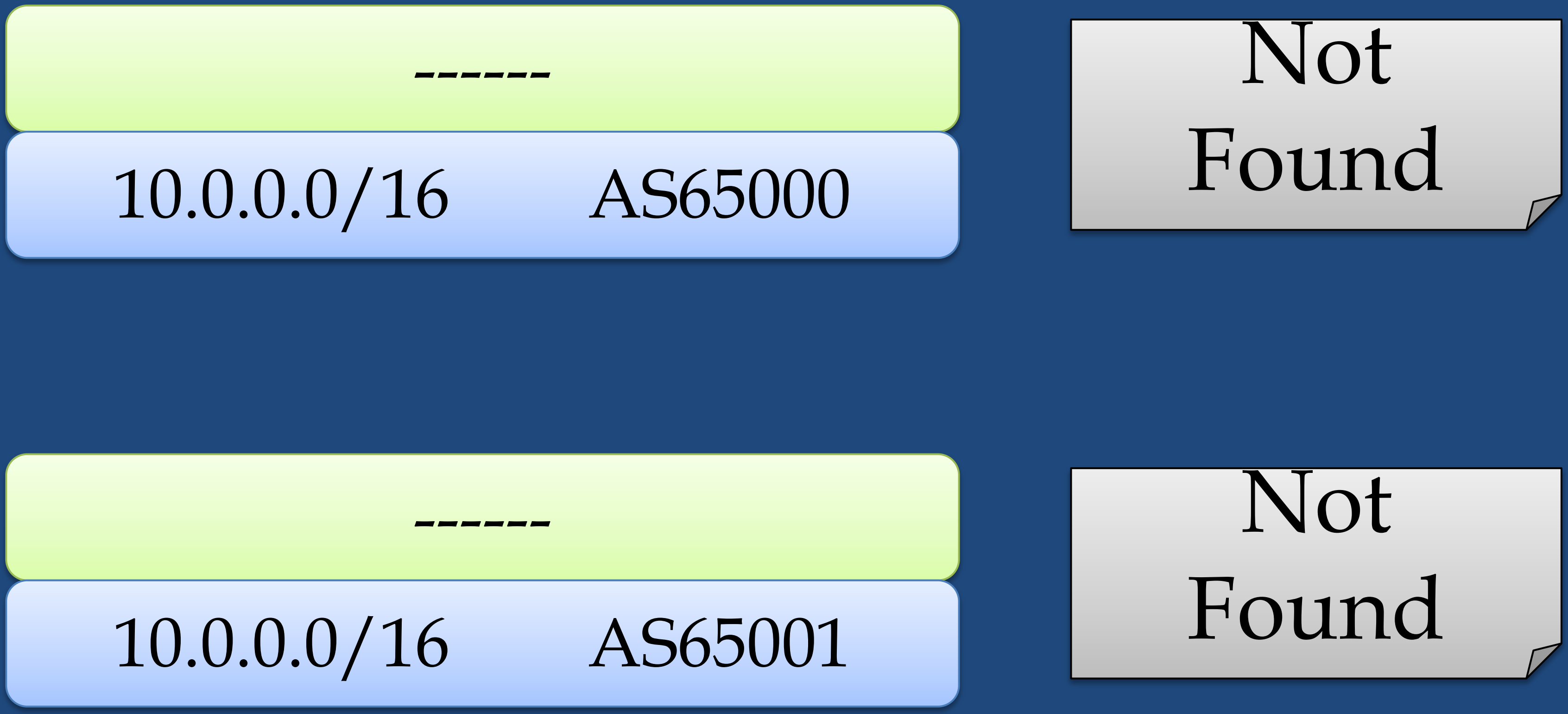
10.0.0.0/16 AS65001

Invalid

copyright (c) tomo

Validation Result（2/7）

ROA
BGP



10.0.0.0/16 AS65000

Not Found

10.0.0.0/16 AS65001

Not Found

Validation Result（3/7）

ROA

BGP

10.0.0.0/16-16 AS65000

10.0.0.0/8 AS65000

Not Found

10.0.0.0/16-16 AS65000

10.0.0.0/17 AS65000

Invalid

10.0.0.0/16-24 AS65000

10.0.0.0/17 AS65000

Valid

Validation Result（4/7）

🟩 ROA
🟦 BGP

🟩 10.0.0.0/16-16 AS65000
🟩 10.0.0.0/16-16 AS65001
🟦 10.0.0.0/16 AS65000

Valid

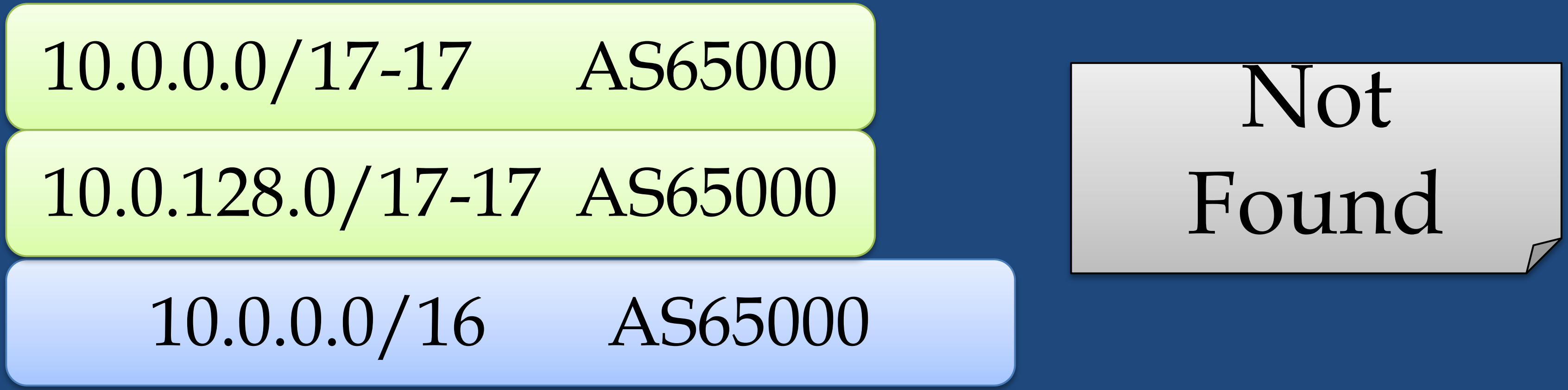
🟩 10.0.0.0/16-16 AS65000
🟩 10.0.0.0/16-16 AS65001
🟦 10.0.0.0/16 AS65001

Valid

Validation Result（5/7）

ROA

BGP



10.0.0.0/17-17 AS65000

10.0.128.0/17-17 AS65000

10.0.0.0/16 AS65000

Not Found

Validation Result（6/7）

ROA

BGP

10.0.0.0/16-24 AS0

10.0.0.0/8 AS65000

Not Found

10.0.0.0/16-24 AS0

10.0.0.0/24 AS65000

Invalid

10.0.0.0/16-24 AS0

10.0.0.0/32 AS65000

Invalid

Validation Result（7/7）

ROA
BGP

10.0.0.0/16-24 AS65000
10.0.0.0/24 {AS65000}

Not Found

10.0.0.0/16-24 AS65000
10.0.0.0/24 {AS65001}

Not Found

10.0.0.0/16-24 AS65000
10.0.0.0/24 {AS65000, AS65001}

Not Found

Example of some prefix

```
asr#show ip bgp 109.5.117.0/24
BGP routing table entry for 109.5.117.0/24, version 231295
131079 7521 2497 15557 41334
  192.0.2.254 from 192.0.2.254 (210.173.172.118)
    Origin IGP, localpref 100, valid, external, best
    Community: 65400:1
    path 7F9B26F111D0 RPKI State invalid
 Refresh Epoch 1
 131079 7521 2497 15557 41334, (received-only)
  192.0.2.254 from 192.0.2.254 (210.173.172.118)
    Origin IGP, localpref 100, valid, external
    path 7F9B26F11168 RPKI State valid
```

Why Invalid?

Example of some prefix

- It’s only way to see ROAs cache data from the top on the router …

copyright (c) tomo

Finally..

```
#ROA
・・・
109.0.0.0/11      11     15557     0      210.173.176.117/42420
```

109.0.0.0/11-11 AS15557

109.5.117.0/24 AS41334

Invalid

Operational Issue

- It's difficult to recognize why it's invalid
  - we should find out the ROA which cover that prefix

- It's better to be able to search that ROA using some command w/prefix or adding the reason(ROA) for BGP information

One idea

- >show ip bpg roa 109.5.117.0/24

Prefix	ASN
109.0.0.0/11	15557

Another idea

```
asr#show ip bgp 109.5.117.0/24
BGP routing table entry for 109.5.117.0/24, version 231295
131079 7521 2497 15557 41334
    192.0.2.254 from 192.0.2.254 (210.173.172.118)
      Origin IGP, localpref 100, valid, external, best
      Community: 65400:1
      path 7F9B26F111D0 RPKI State invalid (ROA: 109.0.0.0/11)
  Refresh Epoch 1
  131079 7521 2497 15557 41334, (received-only)
    192.0.2.254 from 192.0.2.254 (210.173.172.118)
      Origin IGP, localpref 100, valid, external
      path 7F9B26F11168 RPKI State valid
```

What Router dose

1. Storing ROA cache using RTR protocol

2. Validate with ROA cache data inside the router

3. Propagate RPKI validation result with iBGP for other router

Extended Community

	Valid	Not found	Invalid
Cisco ASR	0x43:0:0	0x43:0:1	0x43:0:2
Juniper M120	0x43:X:0	0x43:X:1	0x43:X:2

For Juniper, X: ASN

Couldn’t communicate each other…

Challenge 450K (v4)+10K (v6) ROA

- Built test ROAs from RIS
- Import to the router from ROA cache server

copyright (c) tomo

Challenge 450K (v4)+10K (v6) ROA

※M120 JUNOS 12.2B2.2 is a beta code

	Time to receive ALL ROAs
Cisco ASR	20s
Juniper M120	4m25s

w/IPv4 410K fullroute, w/IPv6 9K fullroute

Challenge 450K (v4)+10K (v6) ROA

	Time to receive IPv4 fullroute	Time to receive IPv6 fullroute
Cisco ASR	1m2s	3s
Juniper M120	2m35s	9s

w/IPv4 450K ROAs, IPv6 10K ROAs

copyright (c) tomo

Challenge 450K (v4)+10K (v6) ROA

	Time to receive IPv4 fullroute	Time to receive IPv6 fullroute
Cisco ASR	1m4s	1s
Juniper M120	2m34s	12s

w/ no ROA

copyright (c) tomo

Other Issues

- When RTR session goes down, the RPKI status will be not found for all the bgp route after a while
  - Invalid => not found
  - we need several RTR sessions or care your filtering policy
- In case of the router reload, which one is faster, receiving ROAs or receiving BGP routes?
  - If receiving BGP is match faster than ROA, the router propagate the invalid route to others

My Impression

- Basic function has been almost completed
- There is still need to improve variously especially from the operational point of view

copyright (c) tomo