(Ab)Using Route Servers

Elisa Jasinska <elisa.jasinska@ams-ix.net>
Chris Malayter <cmalayter@switchanddata.com>
Agenda

• Why Route Servers?
• What do Route Servers do?
• Current implementations and Route Server Working Group
• Functionality and scalability testing
Agenda

- Why Route Servers?
- What do Route Servers do?
- Current implementations and route Server Working Group
- Functionality and scalability testing
Why Route Servers?

- Internet Exchange (e.g. AMS-IX, DE-CIX, LINX)
- Peering Platform for many Parties
- Route Servers for the Participants
Why Route Servers?

- Peer with as many parties as possible
  ➡ Maintaining lots of BGP Sessions
Why Route Servers?

- Reach a lot of Parties with just one BGP session
Why Route Servers?

• Redundancy ... in case your sessions die ...
Why Route Servers?

- Redundancy ... in case the Route Server dies ...
Why Route Servers?

- Easy entry point for new Members to the Exchange - immediate traffic
Agenda

• Why Route Servers?

• What do Route Servers do?

• Current implementations and route Server Working Group

• Functionality and scalability testing
What do route servers do?

- Receive UPDATEs from every participant

```
19:58:33.721679 IP (tos 0x0, ttl 64, id 44576, offset 0, flags [DF], proto TCP (6), length 117) 
10.23.0.5.58880 > 10.23.0.1.179: P, checksum 0xb892 (correct), 48:113(65) ack 61 win 1460
<nop,nop,timestamp 1976783474 3177206804>: BGP, length: 65
  Update Message (2), length: 65
  ... 
  AS Path (2), length: 10, Flags [T]: 65499 11 12 13
  ... 
  Next Hop (3), length: 4, Flags [T]: 10.23.0.5
  ...
  Updated routes:
    2.0.5.0/24
```

```
19:58:33.723897 IP (tos 0x0, ttl 64, id 42762, offset 0, flags [DF], proto TCP (6), length 117) 
10.23.0.4.33349 > 10.23.0.1.179: P, checksum 0xb033 (correct), 48:113(65) ack 61 win 1460
<nop,nop,timestamp 1976783474 1916183085>: BGP, length: 65
  Update Message (2), length: 65
  ... 
  AS Path (2), length: 10, Flags [T]: 65500 11 12 13
  ... 
  Next Hop (3), length: 4, Flags [T]: 10.23.0.4
  ...
  Updated routes:
    2.0.4.0/24
```
What do route servers do?

- Apply filters for the receiving peers

  from AS65500 accept ANY
to AS65500 announce AS65499

  from AS65499 accept ANY
to AS65499 announce AS65500
What do route servers do?

- Perform “best path” selection for every peer
- Store Routing Information Base (RIB) for every peer

<table>
<thead>
<tr>
<th>flags</th>
<th>destination</th>
<th>gateway</th>
<th>lpref</th>
<th>med</th>
<th>aspath</th>
<th>origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>* &gt;</td>
<td>2.0.4.0/24</td>
<td>10.23.0.4</td>
<td>100</td>
<td>200</td>
<td>65500</td>
<td>11 12 13</td>
</tr>
<tr>
<td>* &gt;</td>
<td>2.0.5.0/24</td>
<td>10.23.0.5</td>
<td>100</td>
<td>200</td>
<td>65499</td>
<td>11 12 13</td>
</tr>
</tbody>
</table>
What do route servers do?

• Forward the RIB contents to the desired peer

19:58:33.901718 IP (tos 0xc0, ttl 1, id 15745, offset 0, flags [DF], proto TCP (6), length 103) 10.23.0.1.179 > 10.23.0.4.33349: P, cksum 0x2b21 (correct), 61:112(51) ack 114 win 17376 <nop,nop,timestamp 1916183105 1976783474>: BGP, length: 51
  Update Message (2), length: 51
    ...
    AS Path (2), length: 10, Flags [T]: 65499 11 12 13
    ...
    Next Hop (3), length: 4, Flags [T]: 10.23.0.5
    ...
    Updated routes:
      2.0.4.0/24

19:58:33.903463 IP (tos 0xc0, ttl 1, id 12268, offset 0, flags [DF], proto TCP (6), length 103) 10.23.0.1.179 > 10.23.0.5.58880: P, cksum 0x377e (correct), 61:112(51) ack 114 win 17376 <nop,nop,timestamp 3177206824 1976783474>: BGP, length: 51
  Update Message (2), length: 51
    ...
    AS Path (2), length: 10, Flags [T]: 65500 11 12 13
    ...
    Next Hop (3), length: 4, Flags [T]: 10.23.0.4
    ...
    Updated routes:
      2.0.5.0/24
What do route servers do?
Agenda

- Why Route Servers?
- What do Route Servers do?
- Current implementations and Route Server Working Group
- Functionality and scalability testing
Current Implementations

- Quagga
- OpenBGPD
- BIRD
Route Server Working Group

- Andy Davidson - LONAP
- Chris Malayter - Switch & Data
- Elisa Jasinska - AMS-IX
- Mo Shivji - LINX
- Robert Wozny - PL-IX
- Sebastian Spies - DE-CIX
- Wolfgang Hennerbichler - VIX
Agenda

• Why Route Servers?
• What do Route Servers do?
• Current implementations and route Server Working Group
• Functionality and scalability testing
Functional Testing
AS4 / 32 Bit ASN

• All three implementations support AS4
• All three versions tested as of 12/4/2009 to properly implement AS4
IPv6

• All three implementations support IPv6
• We highly recommend running a current version of any of three implementations
• MANY bugs fixed between 10/1/2009-1/1/2010
• Running a port of a route server is ill-advised and can leave a bad taste in your mouth!
Scalability
Testing

• 100 sessions, set up from IXIA
• 500 or 1000 prefixes per session
• Additional random flapping
Quagga

- Single threaded implementation
- Issues with performing its tasks on time
- CPU thrashing during periods of instability
- Bug causing crash during flapping
Quagga CPU

Quagga cpu usage
multiple-rib; 100 sessions
IXIA vs. Quagga
500 prefixes per session with random flapping
lab4
2 x Intel(R) Xeon(R) CPU 3050 @ 2.13GHz
OpenBGPD CPU

OpenBGPD cpu usage
multiple-rib; 100 sessions
IXIA vs. OpenBGPD
1000 prefixes per session
lab2.paix.net

4 x Intel(R) Xeon(TM) CPU 3.60GHz (GenuineIntel 686-class) 3.61 GHz
BIRD CPU

BIRD cpu usage
multiple-rib; 100 sessions
IXIA vs. BIRD
500 prefixes per session with random flapping
lab6.paix.net
4 x Intel(R) Xeon(TM) CPU 3.80GHz

The diagram shows the CPU usage over time for the BIRD system. The x-axis represents time in seconds, ranging from 0 to 60000, and the y-axis represents CPU usage in percent, ranging from 0 to 90. The graph illustrates the dynamic nature of CPU usage, with fluctuations throughout the duration of the test session.
OpenBGPD

- Multi-threaded implementation
- Session thread keeps sessions active while instability is occurring
- 1GB memory limitation per process on i386 and a 4 GB memory limitation on amd64
OpenBGPD Mem

OpenBGPD mem usage
multiple-rib; 100 sessions
IXIA vs. OpenBGPD
1000 prefixes per session
lab2.paix.net

4 x Intel(R) Xeon(TM) CPU 3.60GHz (GenuineIntel 686-class) 3.61 GHz
Quagga Mem

Quagga mem usage
multiple-rib; 100 sessions
IXIA vs. Quagga
500 prefixes per session with random flapping
lab4
2 x Intel(R) Xeon(R) CPU 3050 @ 2.13GHz
BIRD

• Single threaded implementation
• Amazing scheduling system
• The most stable route server we tested
• Discovered odd memory freeing issues in Linux glibc
BIRD Mem

BIRD mem usage
multiple-rib; 100 sessions
IXIA vs. BIRD
500 prefixes per session with random flapping
lab6.paix.net
4 x Intel(R) Xeon(TM) CPU 3.80GHz

mem usage (Mbytes)

memory
free memory

time (seconds)
Thank you!

Questions?

<elisa.jasinska@ams-ix.net>
<cmalayter@switchanddata.com>