

Route Server Bake-off

Andy Davidson, LONAP

UKNOF15, Rochdale, January 21st 2010

Acknowledgements and particular thanks to also to:

Elisa Jasinska, AMS-IX

Chris Malayter, Switch and Data

Mo Shivji, LINX

Wolfgang Hennerbichler, VIX

Robert Wozny, PLIX

Sebastian Spies, DECIX

Agenda

- What Route Servers Do
- Situation in November 2009
- Functional Tests
- Performance/Scalability Tests

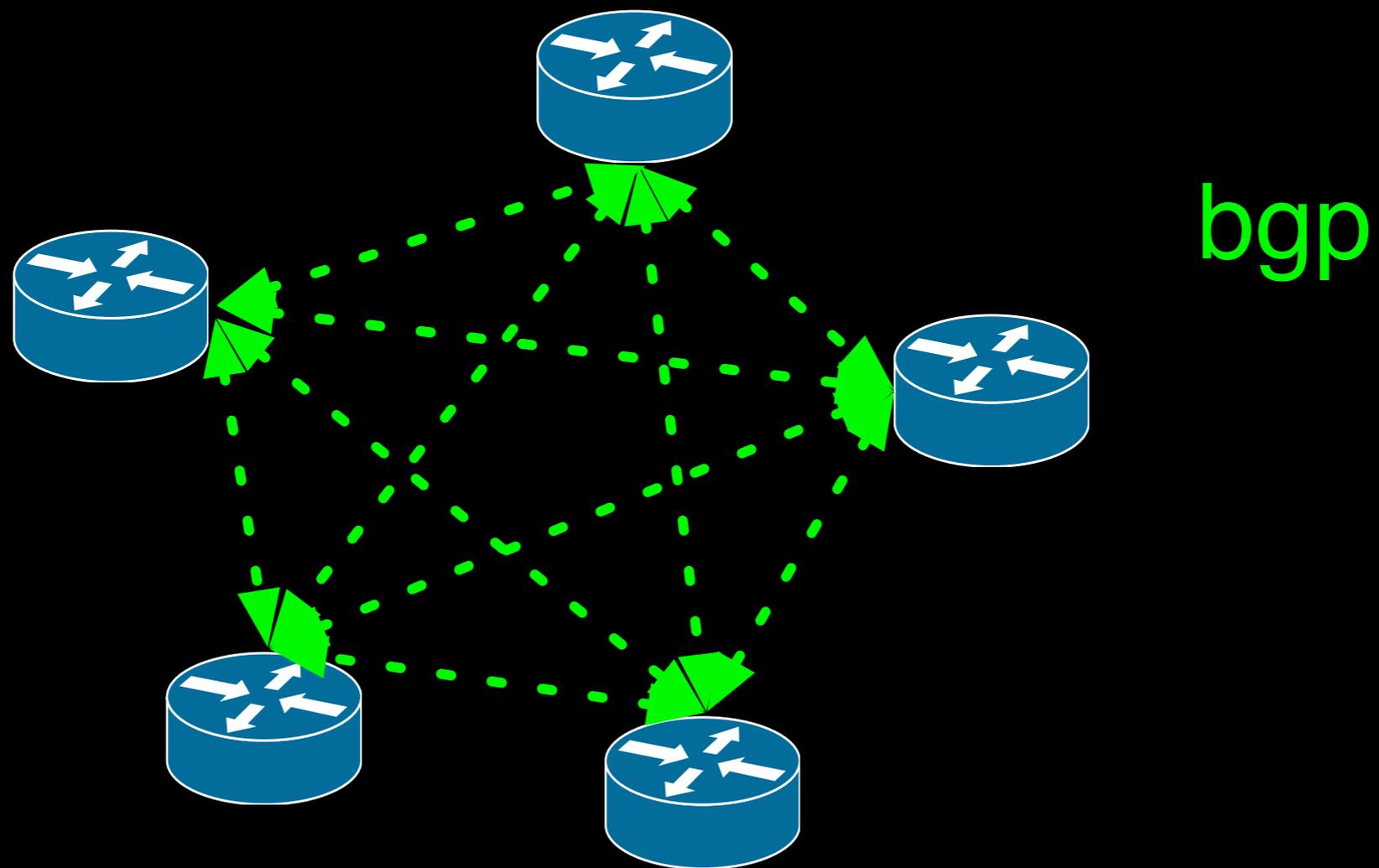
Agenda

- **What Route Servers Do**
- Situation in November 2009
- Functional Tests
- Performance/Scalability Tests

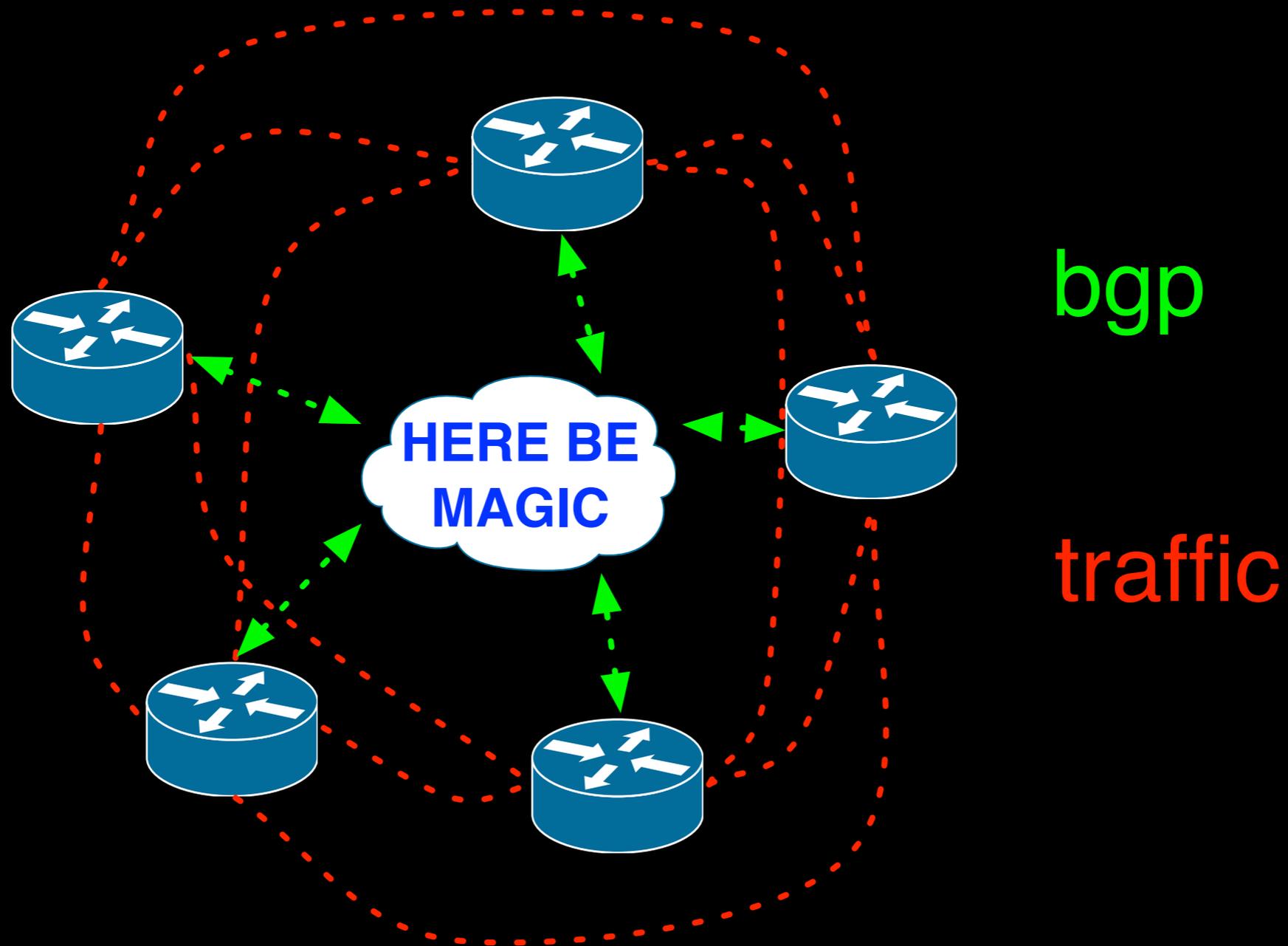
Why Operators Care

- At IXP but not peering - **and want to**
- At IXP and peering with route-server, **concerned about stability**
- Operator, looking at internal r-s **for scaling**

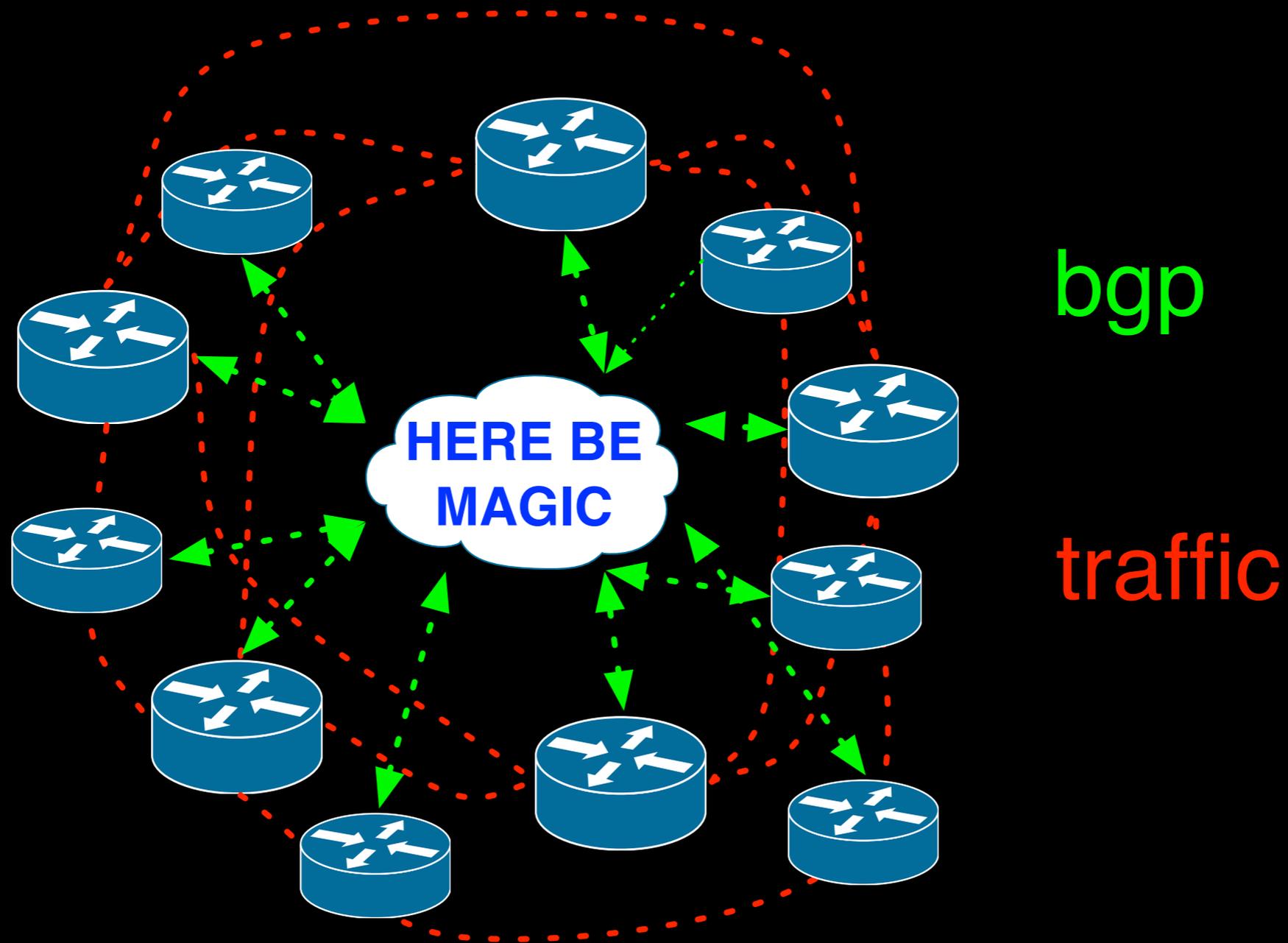
What route-servers do



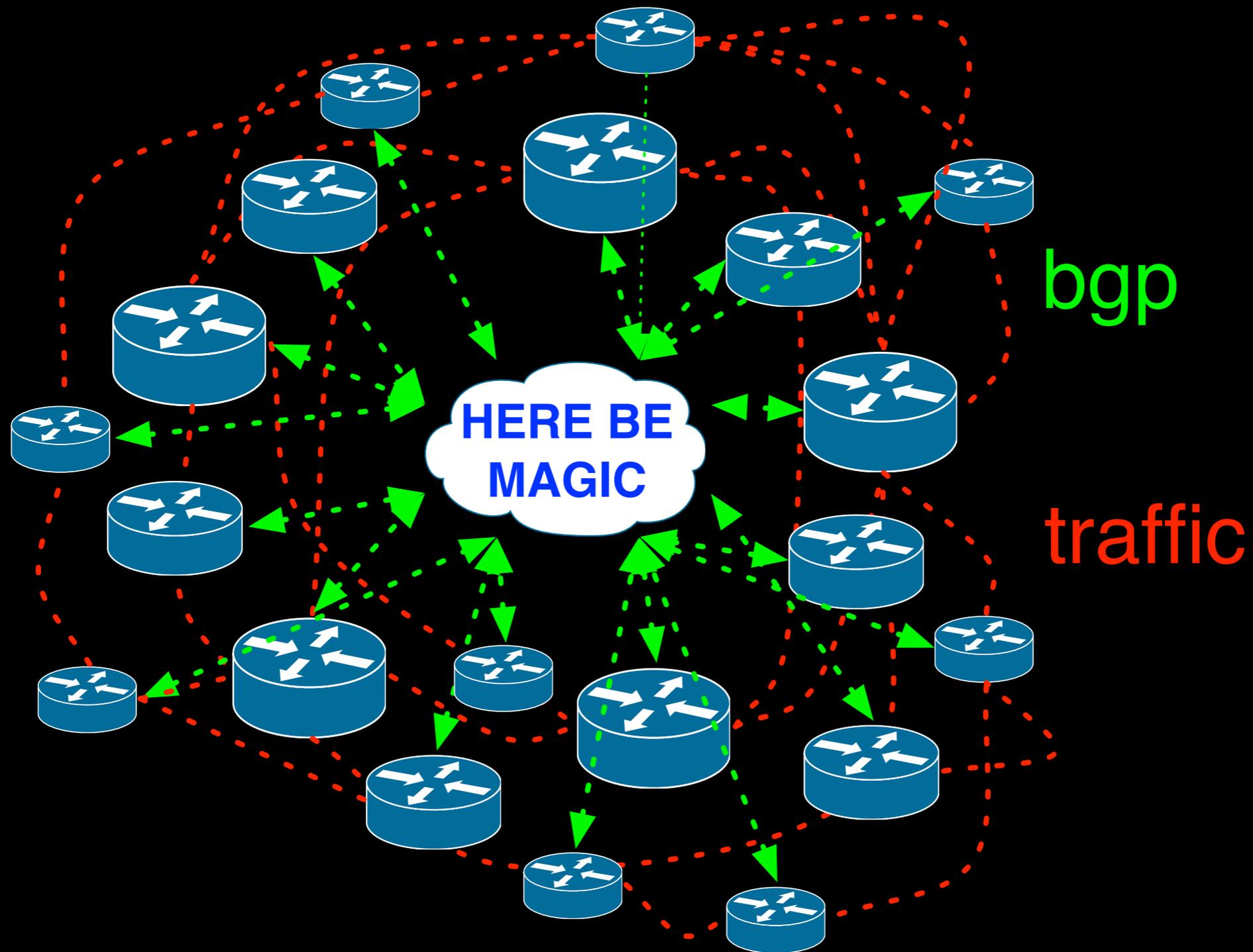
What route-servers do



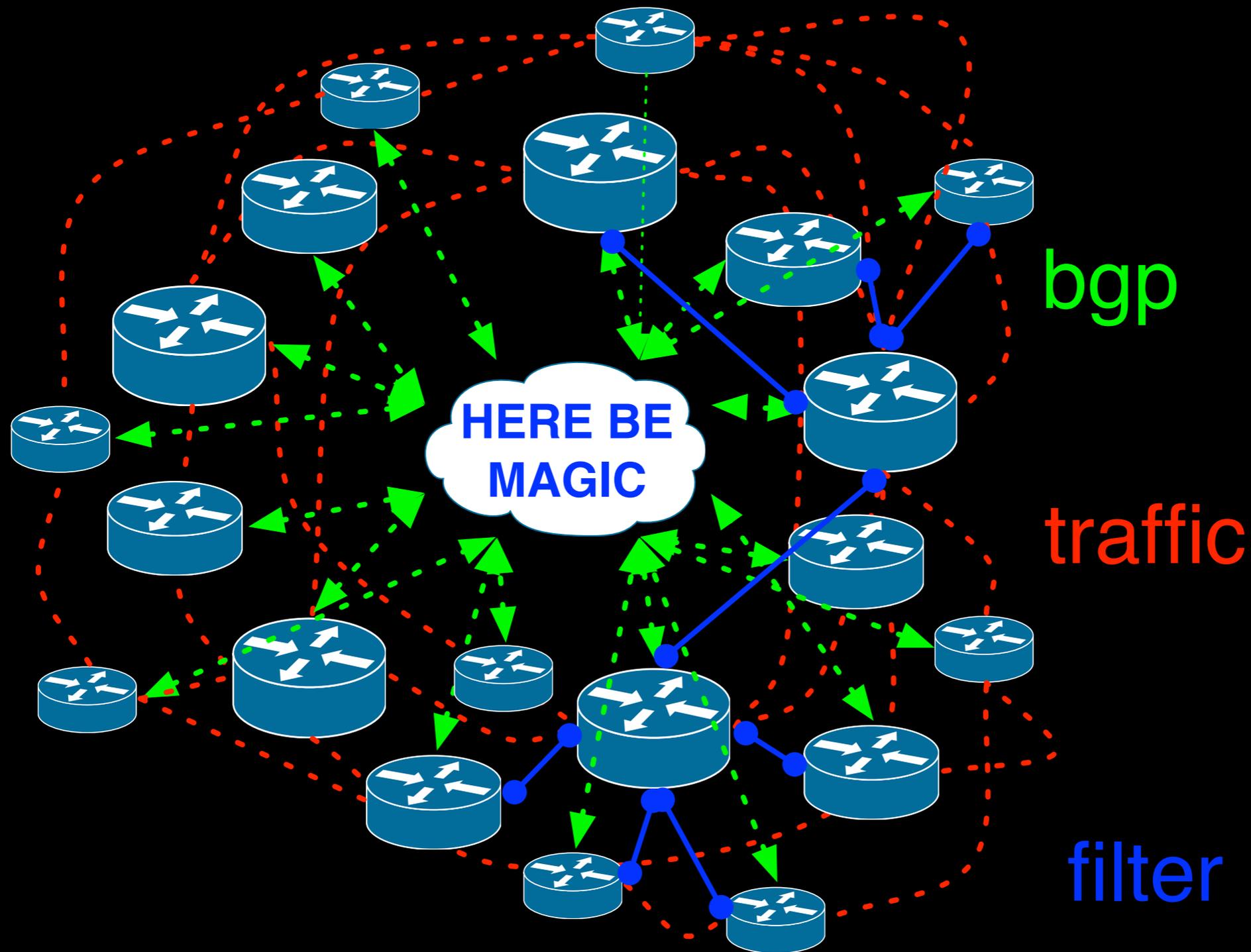
What route-servers do



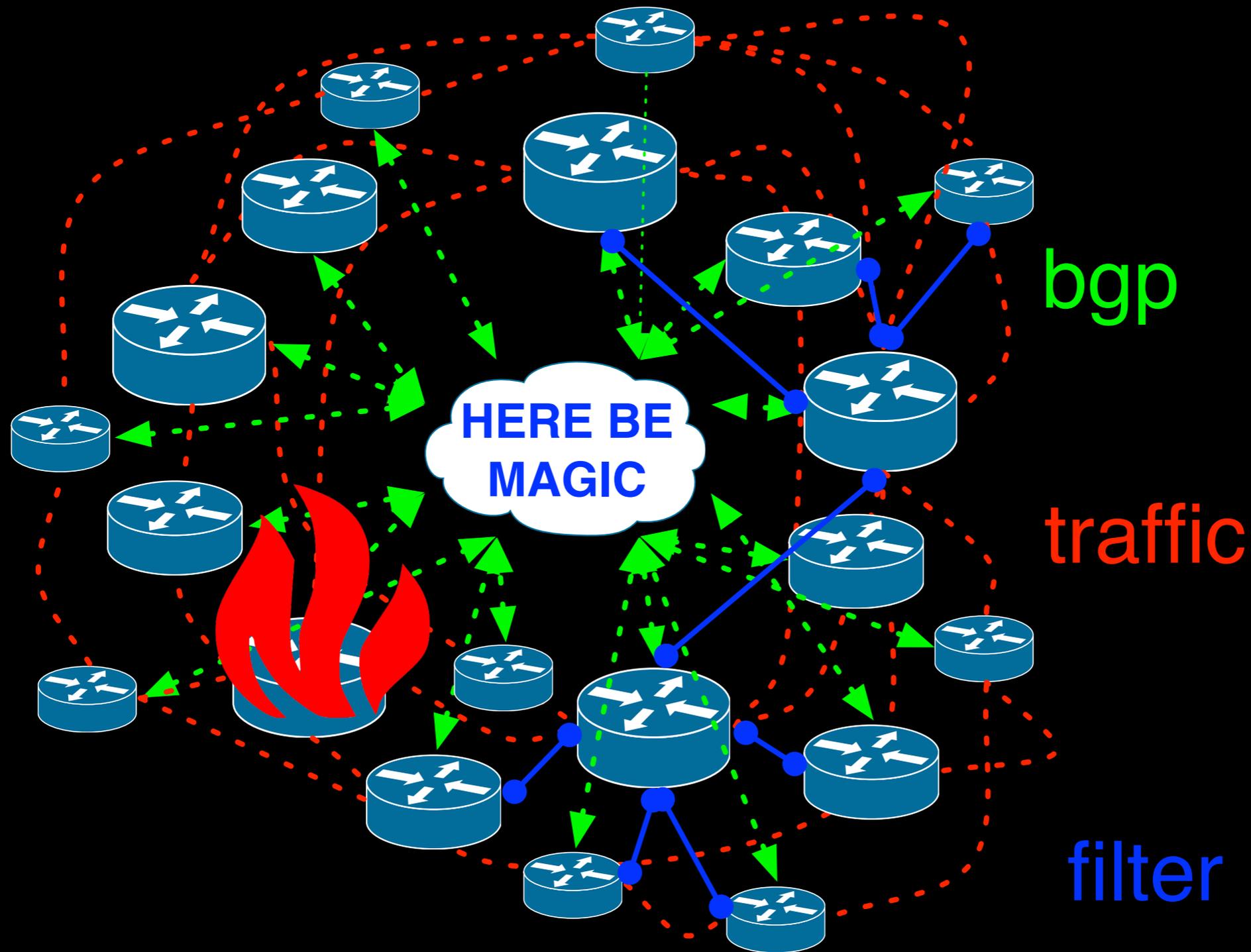
What route-servers do



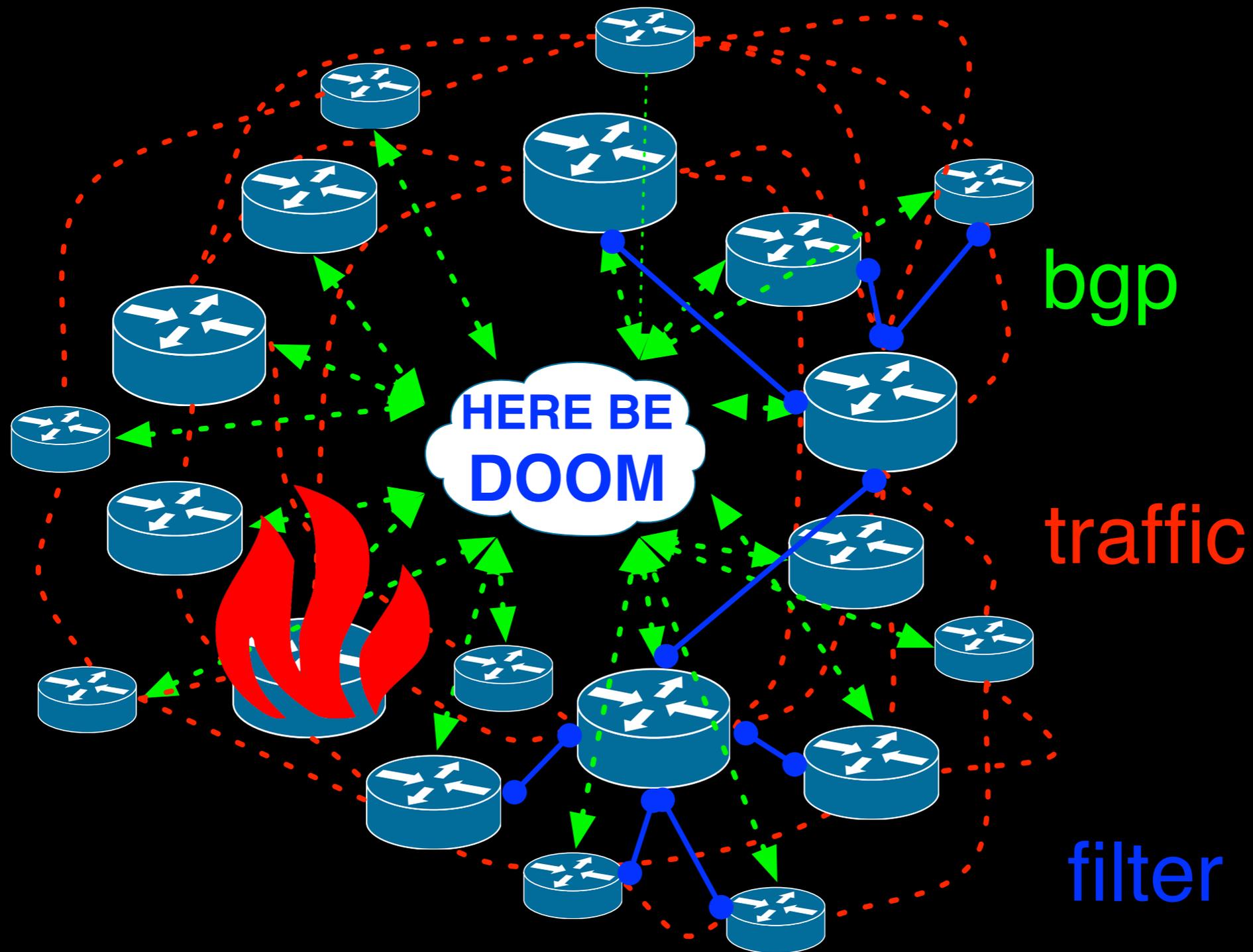
What route-servers do



What route-servers do



What route-servers do



Agenda

- What Route Servers Do
- **Situation in November 2009**
- Functional Tests
- Performance/Scalability Tests

Situation

- IXP Community **complaints** turned into a significant effort to fix situation.
- But we did not know the **new limits** of this software, and had not discovered any new **bugs**.

So



- December 2009, tests of capacity
and an awful lot of config/tip swapping.

So



- December 2009, tests of capacity and bugs and an awful lot of config/tip swapping.

Agenda

- What Route Servers Do
- Situation in November 2009
- **Functional Tests**
- Performance/Scalability Tests

Test bugs marked as fixed

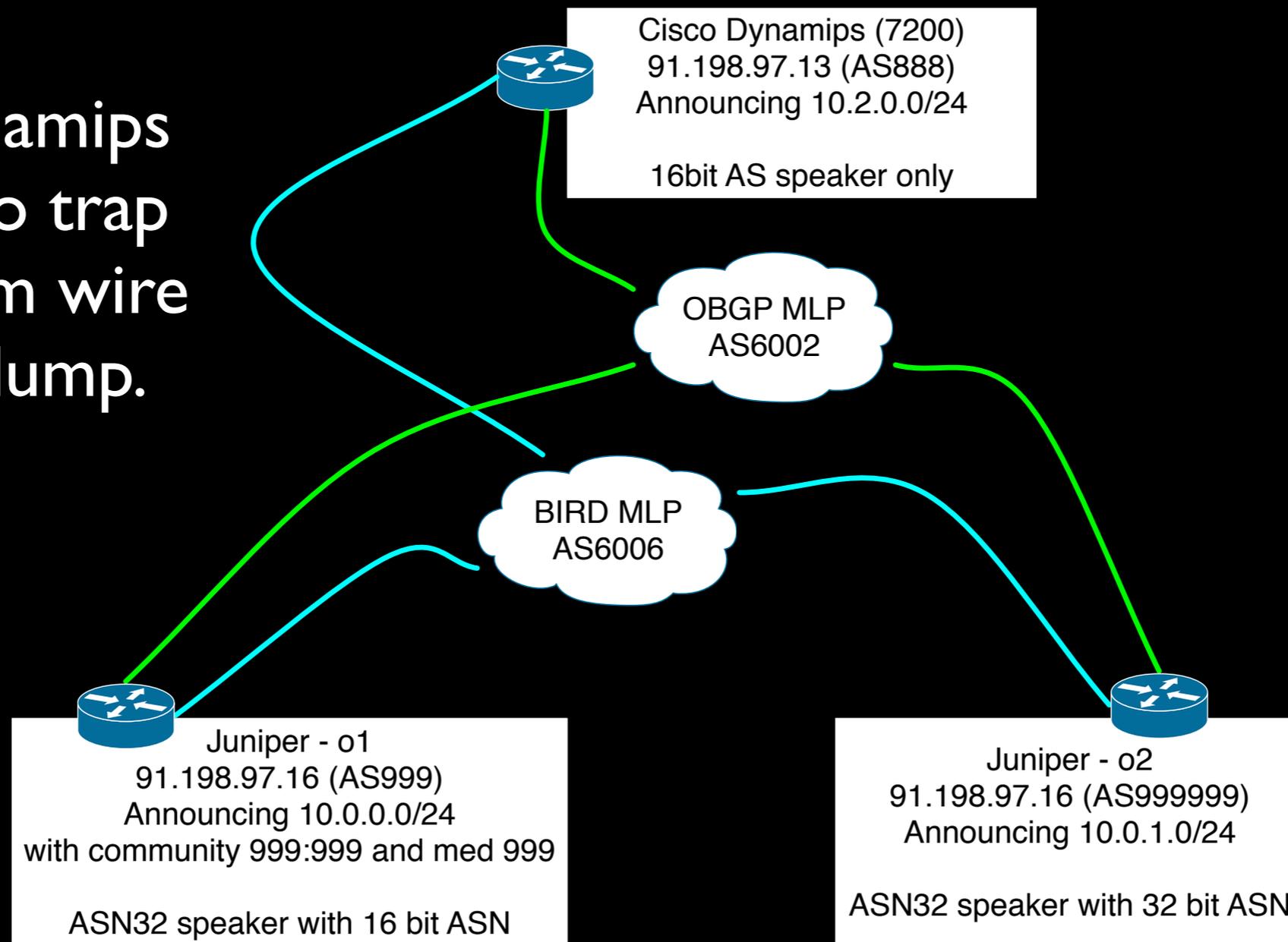
- Several bugs raised in Q3/4 2009 relating to brand new features that IXPs asked for.
- Behaviour on New RIB.
- Erronous IPv6 withdraws

Filtering

- **V4 and V6 inbound Prefix filtering** (fierce debate about whether route-servers should actually do this!)
- **Max prefix limit**
- **V4 and V6 outbound Prefix filtering** (required if multiple RIBs in use. Vyatta Bug.)
- all worked

Transparency Tests

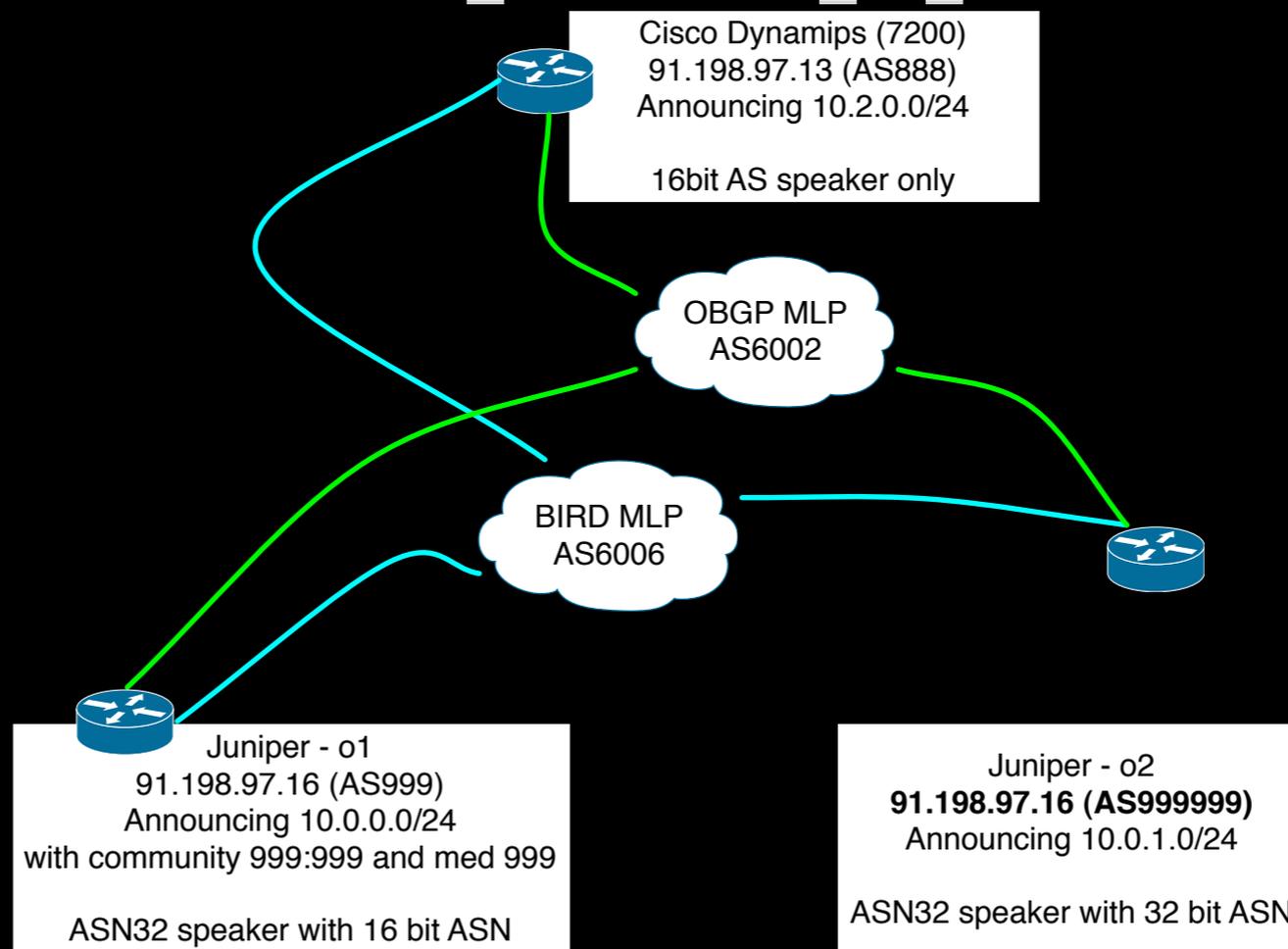
Olive/Dynamips allow us to trap packet from wire with tcpdump.



ASN32

Correct **Compatibility Mode**

AS23456 in AS_PATH.NEW_AS_PATH attribute added.

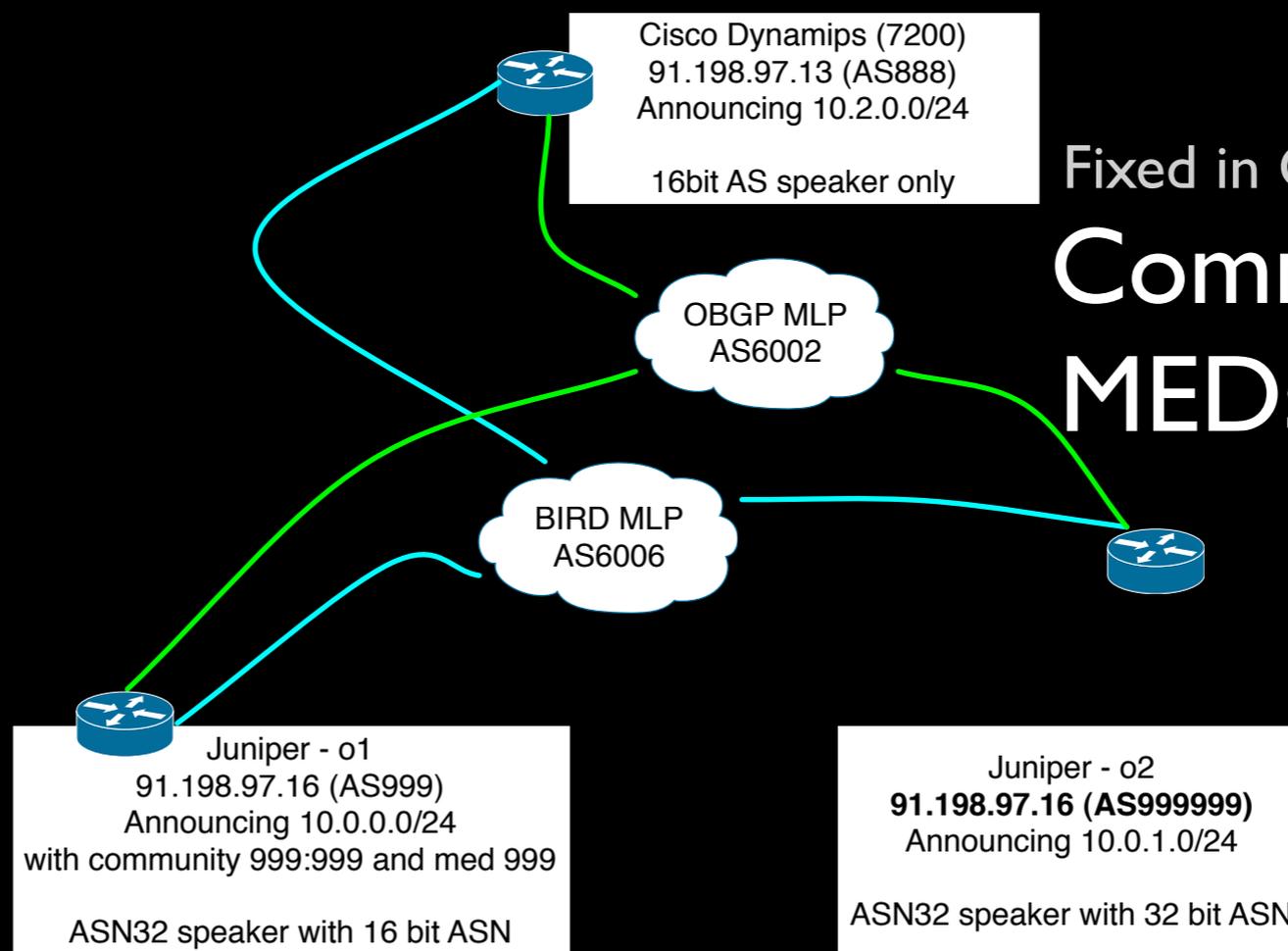


Correct **32 bit ASN** properties

(AS4_PATH, adjacent to AS9999999)

MEDs & General attribute transparency

Common handling of attributes across RS families to allow IXPs to run multiple RS & send equivalent BGP messages.



Fixed in OpenBGPd Q4 2009
Communities and MEDs unchanged

IPv6

- Run a current version of all route servers for best v6 support

MD5

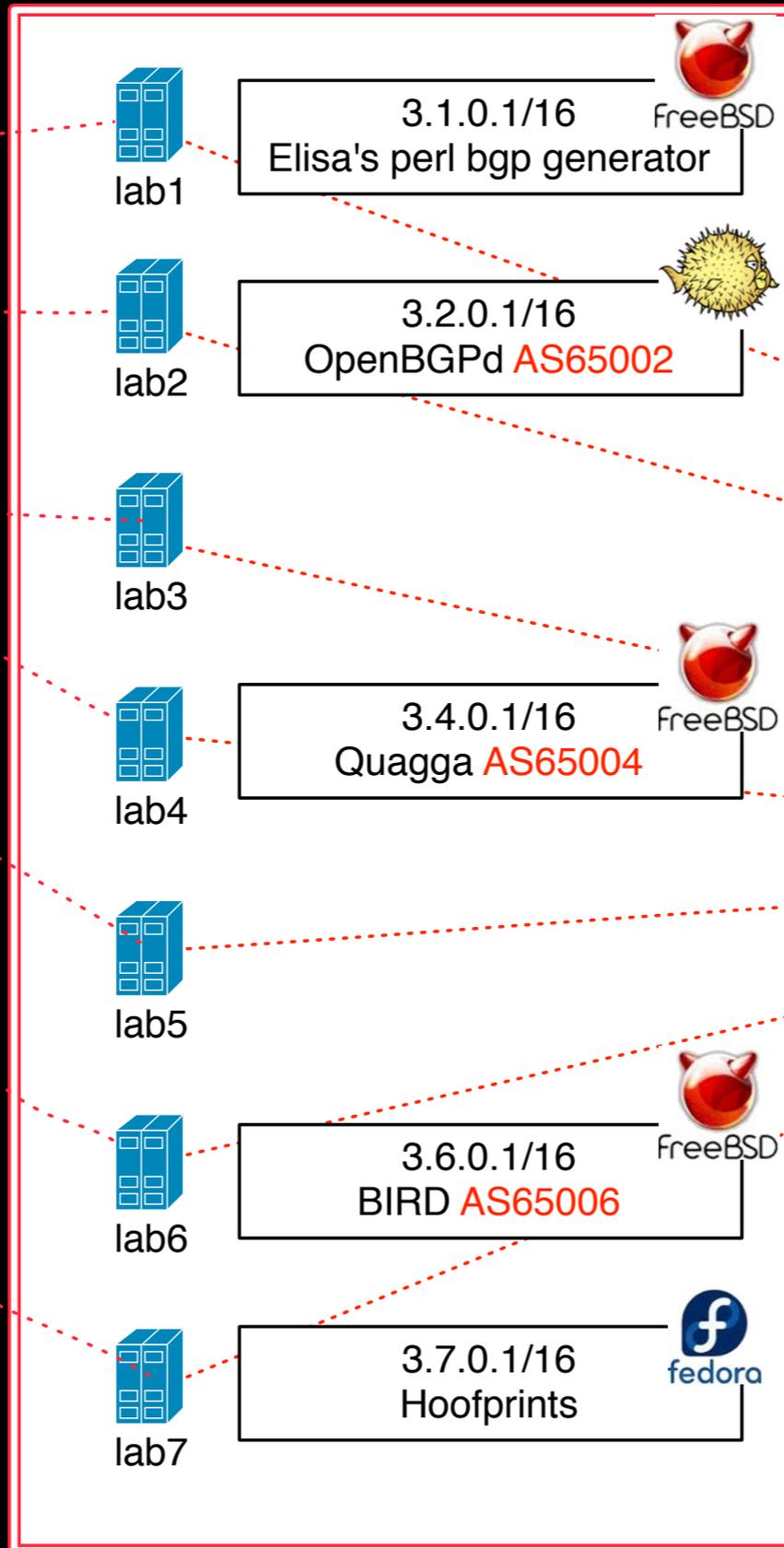
- Problem with sessions continuing to expect tcp-md5 once configuration is removed
- Not yet explained - possibly OS fault
- Recommend that participants do not use md5 on route-server sessions

Agenda

- What Route Servers Do
- Situation in November 2009
- Functional Tests
- **Performance/Scalability Tests**

Sessions originate from
1.6.1.2 (AS10002).
The IP address and AS
number increase by one

Ixia XM12
1.0.0.1/8



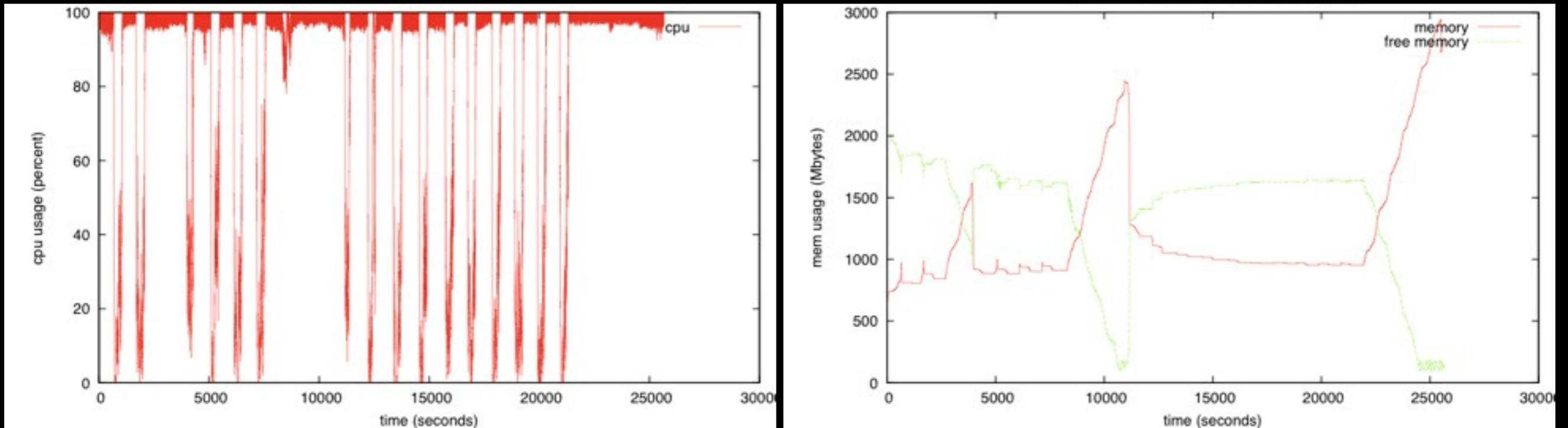
Sessions originate from
3.1.1.0 (AS9000).
The IP address and AS
number increase by one

Force10 E300
Vlan 20 - 10.0.0.0/8 Mgmt network
Vlan 21 - Public IP addressing

Performance

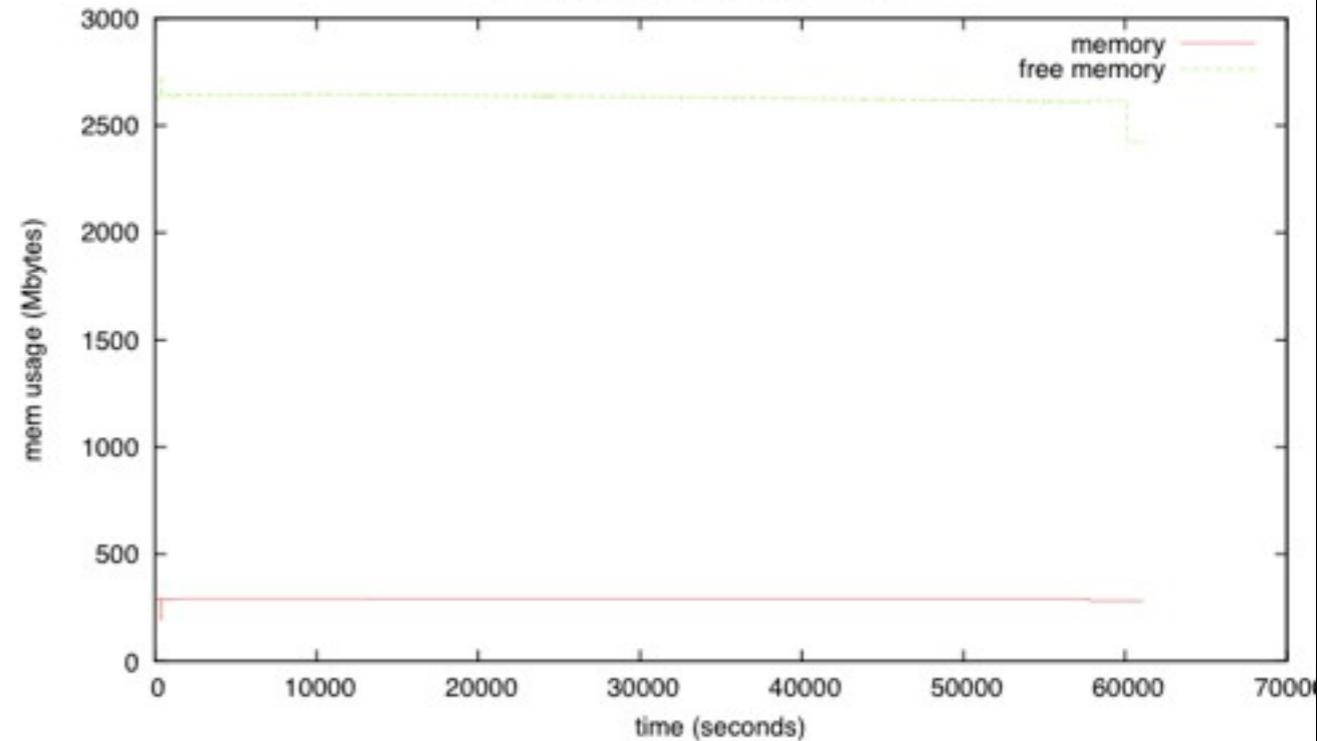
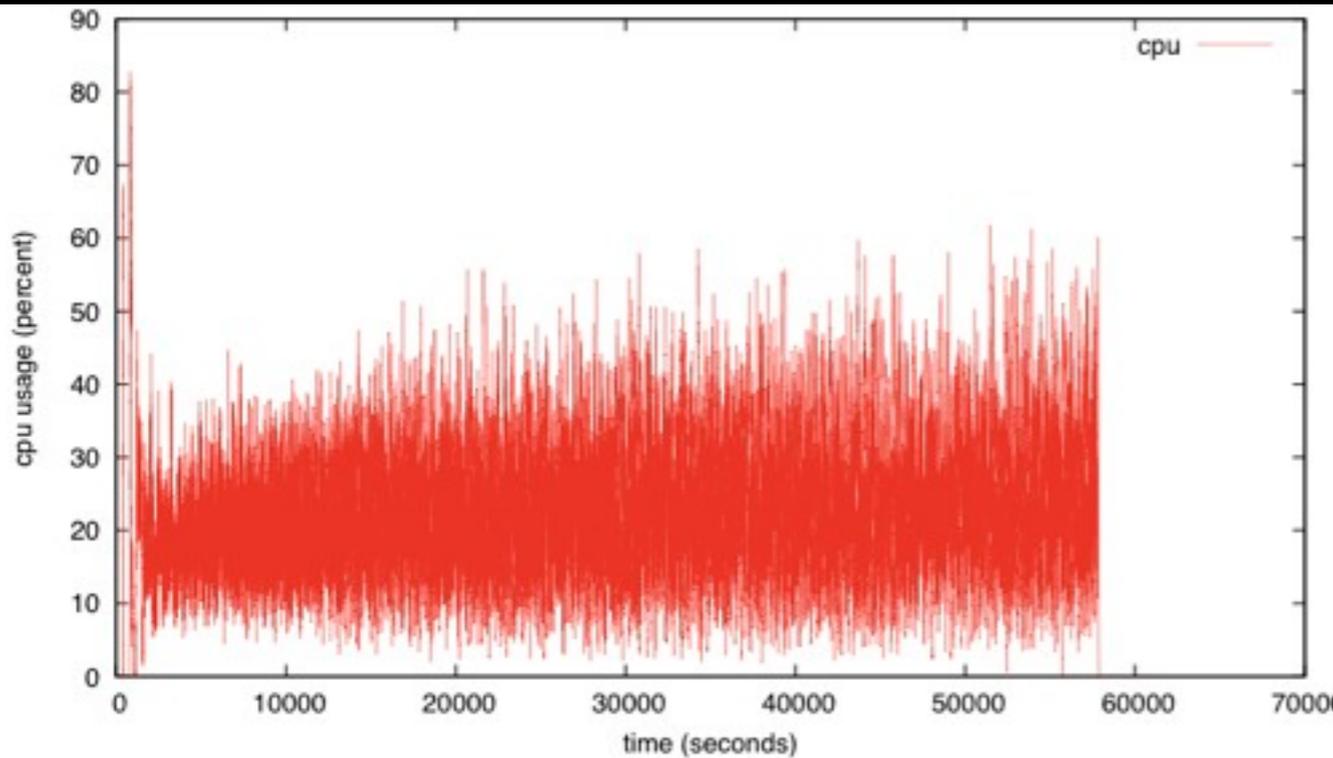
- Compressing a week of tests into **general themes**, that best demonstrate:
 - General behaviour **under load**
 - **Capacity** (sessions, prefixes...)

100 Sessions, 500 Pfx, Flapping at random



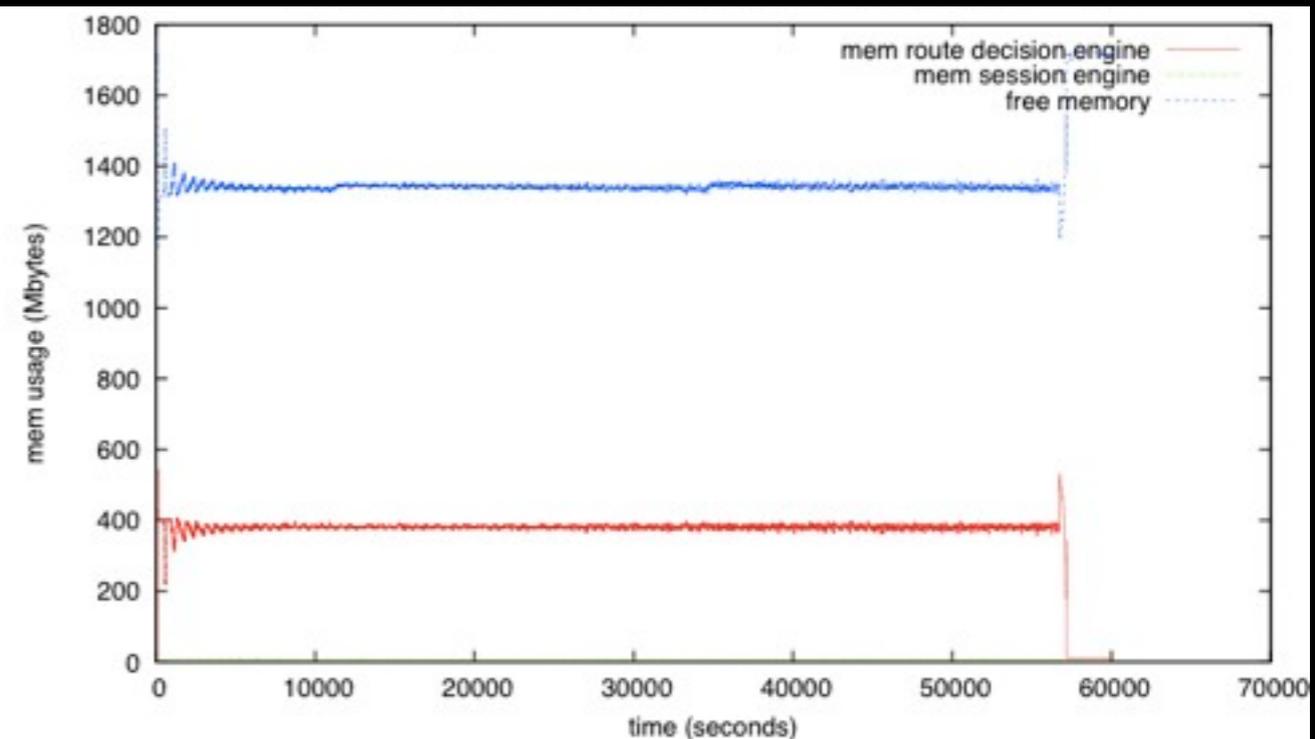
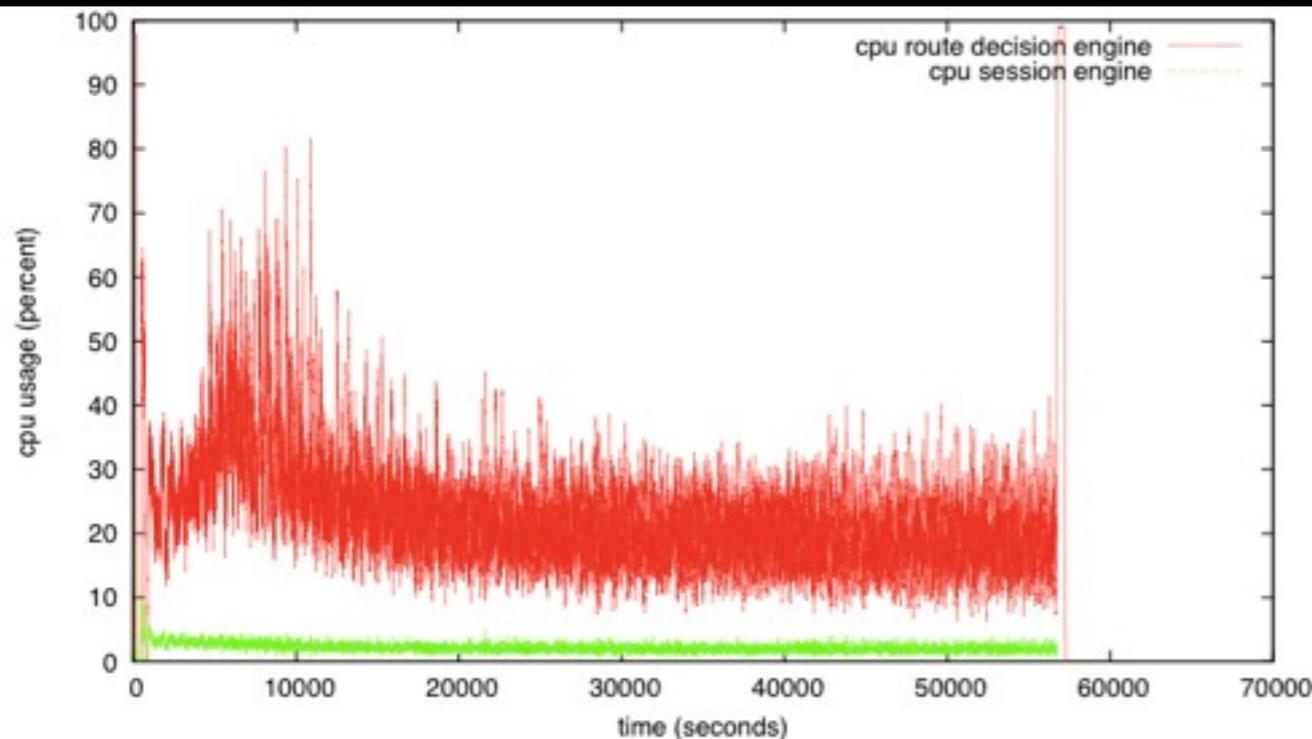
- Quagga. Strange memory profile and high CPU load with flapping sessions.
- Unstable, unusable with flapping participants.

100 Sessions, 500 Pfx, Flapping at random



- BIRD.

100 Sessions, 500 Pfx, Flapping at random



- OpenBGPd. Note - two results because two threads.

Prefixes

- Limited really by RAM
- More prefixes, more convergence time
- Memory limit is prefixes * RIBs
- OpenBSD 1 GB limit per thread (32 Bit)

Sessions

- BIRD struggles by 1,000 sessions
- OpenBGP starts 1,000 sessions fine, but adding 1,000 RIBs/Prefixes = RAM starved
- Not just describing OS limits, e.g. file desc
- Not a problem for all but the largest IXPs

Conclusions

- Where's the **doom** ?
- Where's the **gloom** ?
- Where are the **BUGS** ?!!!
A red flame icon consisting of a large flame above four smaller flames, positioned below the word 'BUGS' in the third bullet point.

Conclusions

- Where's the doom ?
- Where's the gloom ?
- Where are the BUGS ?!!
- There are new, more scalable route-server daemons available today.
- The 'vendors' did fix our bugs
- ISPs tempted to connect should do so

Any Questions ?

Any Answers ?

Andy Davidson

<andy@lonap.net>