

(Ab)Using Route Servers

<elzbieta.jasinska@ams-ix.net>



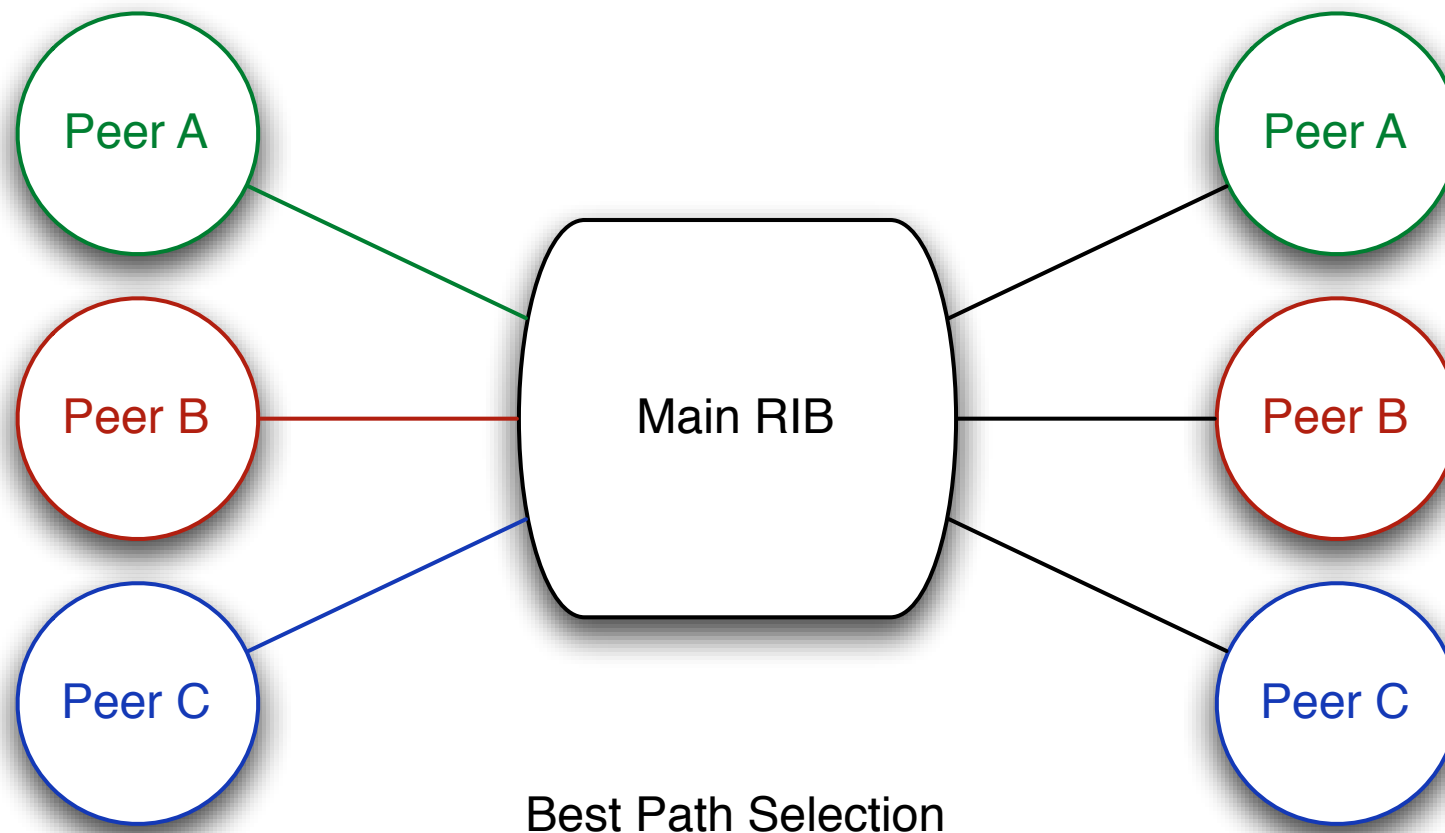
Agenda

- Route Servers
- OpenBGPD at AMS-IX
- Route Server Working Group
- Functionality and Scalability

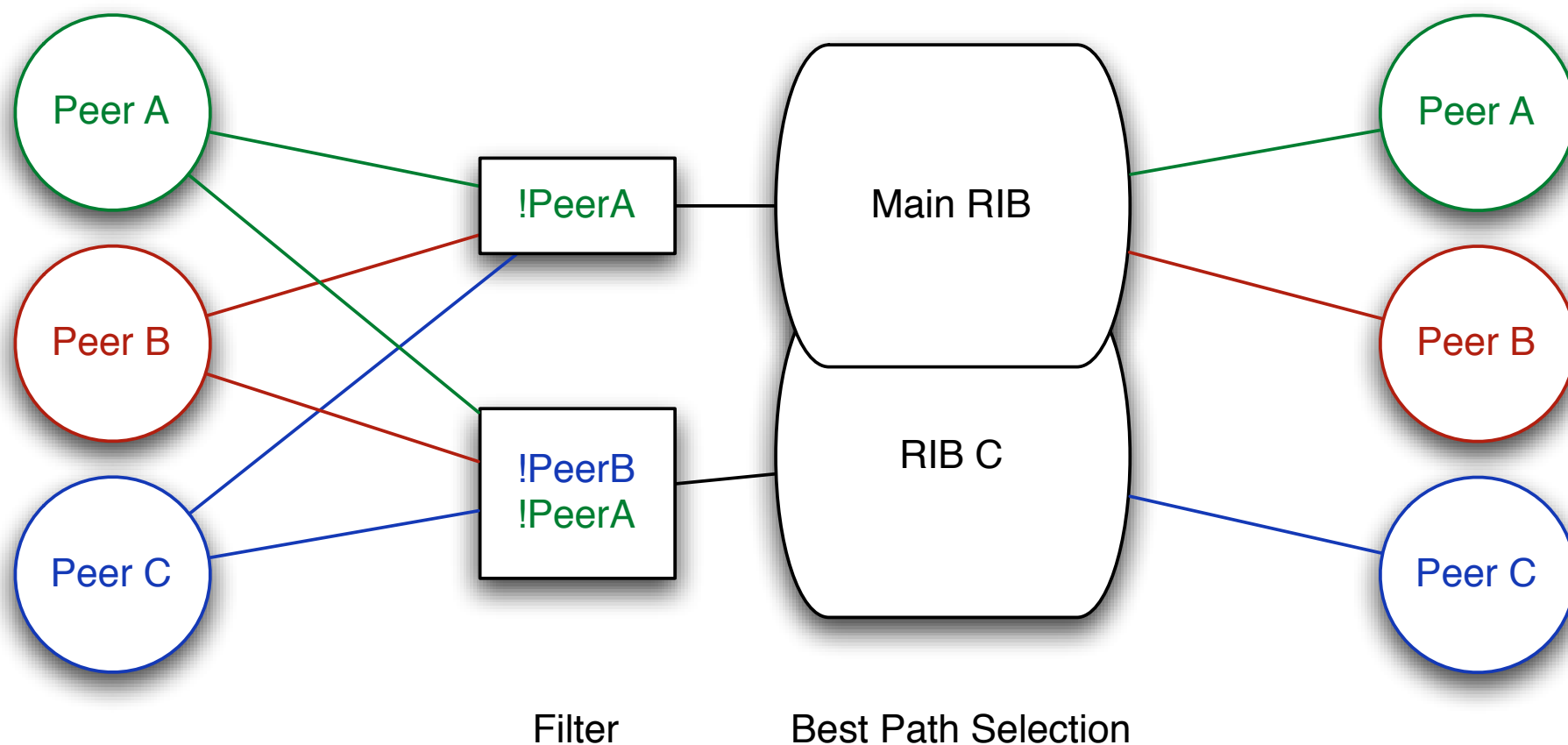
Agenda

- Route Servers
- OpenBGPD at AMS-IX
- Route Server Working Group
- Functionality and Scalability

Route Servers



Route Servers



Route Servers

- Transparent BGP node
- Easy entry point at IX
- Serves as backup
- Saves on maintaining sessions

Agenda

- Route Servers
- **OpenBGPD at AMS-IX**
- Route Server Working Group
- Functionality and Scalability

After PLNOG3....



More Issues

- Existing prefixes were not propagated into newly created RIBs
 - Fixed since October 2, 2009
- MED transparency
 - Fixed since October 5, 2009
- Restart capability introduced error in EOR markers
 - Fixed on October 12, 2009

The day has come...

➔ October 13, 2009

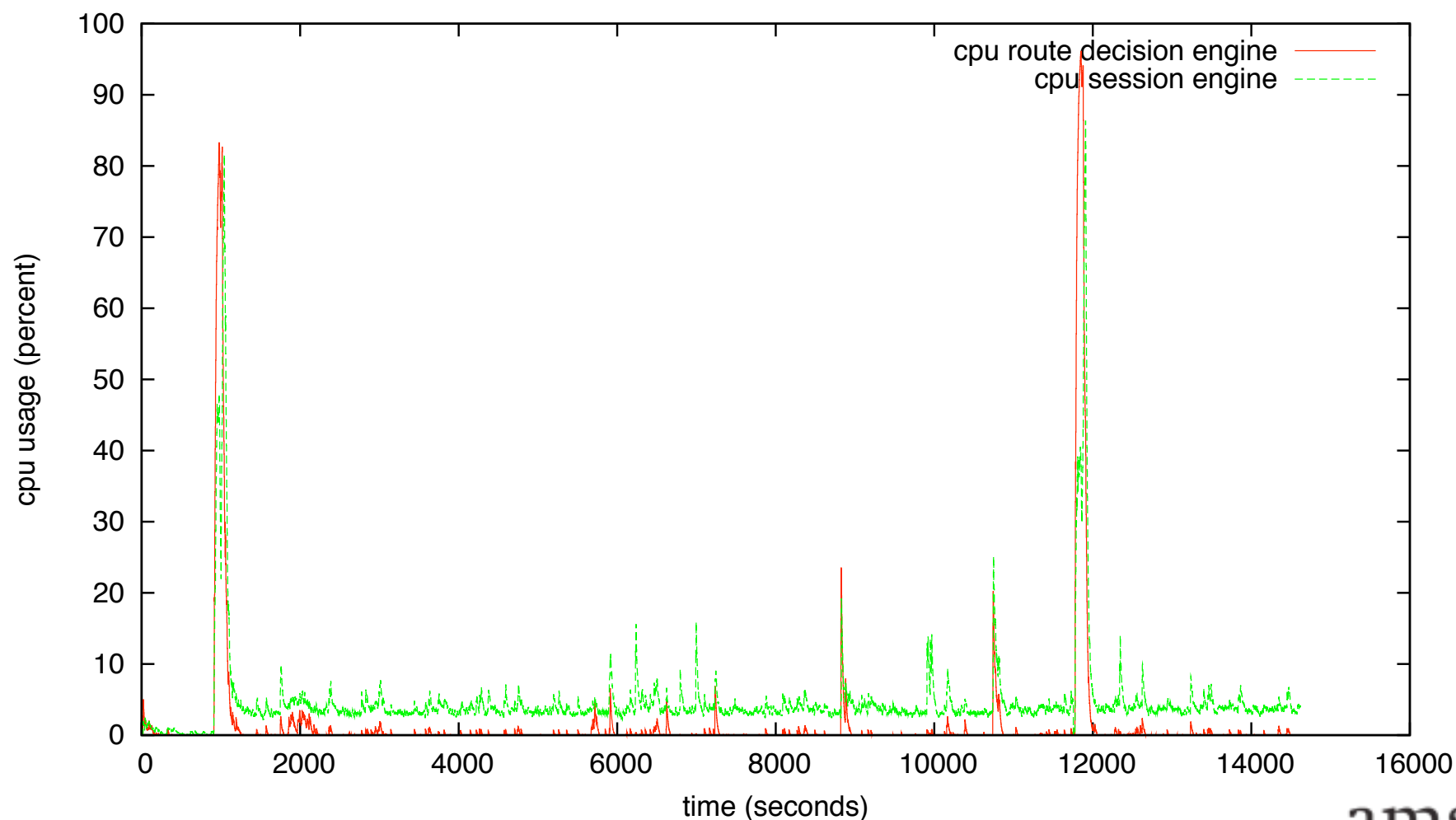
- Replace one of the Quagga boxes with OpenBGPD
- Initial startup went well, but... a hand full of peers closed the session with...

```
Oct 13 14:08:05 prefix bgpd[5121]: neighbor aaa.bbb.ccc.ddd: received  
notification: error in UPDATE message, attribute list error  
Oct 13 14:08:05 prefix bgpd[5121]: neighbor aaa.bbb.ccc.ddd: state change  
Established -> Idle, reason: NOTIFICATION received
```

- Those are tripped prefix limits!
- Cisco doesn't seem to feel like sending the correct error code

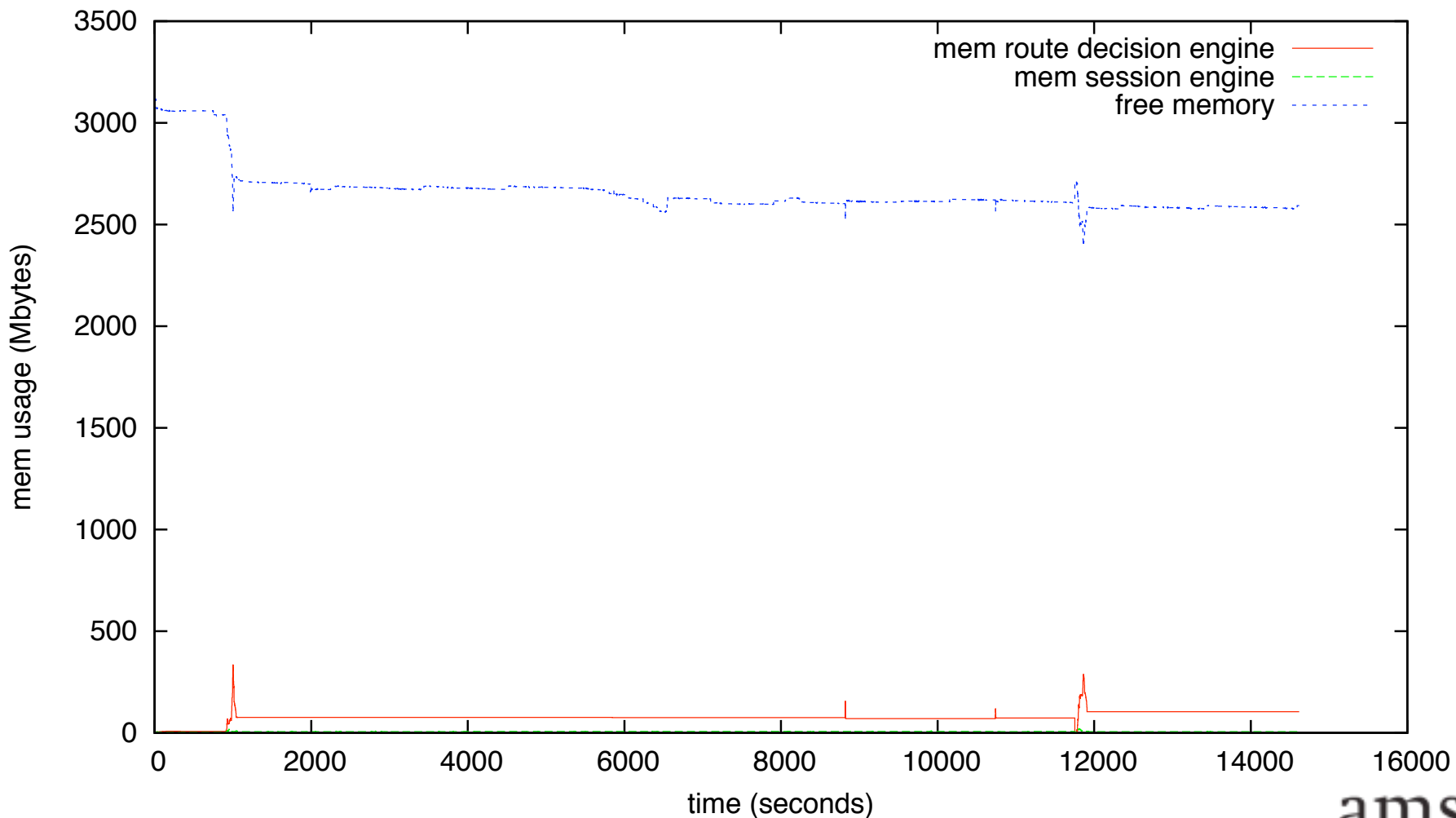
Performance - CPU

OpenBGPd cpu usage
single/multiple-rib; ~320 sessions
production
let's see if that works...
prefix.noc.ams-ix.net



Performance - Mem

OpenBGPD mem usage
single/multiple-rib; ~320 sessions
production
let's see if that works...
prefix.noc.ams-ix.net



MD5 Issue

- A hand full peers didn't come up at all, in fact OpenBGPD didn't even report a single received packet

```
13:29:20.422205 00:12:1e:da:6b:f0 > 00:1e:4f:36:ca:9f, ethertype IPv4  
(0x0800), length 62: aaa.bbb.ccc.ddd.53070 > 195.69.144.255.179: S  
1655430005:1655430005(0) win 16384 <mss 1460,sackOK,eol>  
13:29:23.621485 00:12:1e:da:6b:f0 > 00:1e:4f:36:ca:9f, ethertype IPv4  
(0x0800), length 62: aaa.bbb.ccc.ddd.53070 > 195.69.144.255.179: S  
1655430005:1655430005(0) win 16384 <mss 1460,sackOK,eol>
```

- At startup wrongly configured MD5 passwds didn't quite go away
- Workaround: `ipseccctl -F`
- Fixed in OpenBSD 4.6

Summary

- OpenBGPD withstands issues which cause Quagga to fail
 - More prefixes due to more liberal filtering default
- ➔ November 12, 2009
- Replaced second route server with OpenBGPD
 - Bug causing route decision engine to crash on config reload still existent

Agenda

- Route Servers
- OpenBGPD at AMS-IX
- **Route Server Working Group**
- Functionality and Scalability

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

- Route Servers
- OpenBGPD at AMS-IX
- Route Server Working Group
- **Functionality and Scalability**

Functional Testing



AS4 / 32 Bit ASN

- All three implementations support AS4
- All three versions tested as of 04 Dec 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 Jan 2009 and Jan 2010
- Running a port of a route server is ill-advised and can leave a bad taste in your mouth!

Scalability



Testing

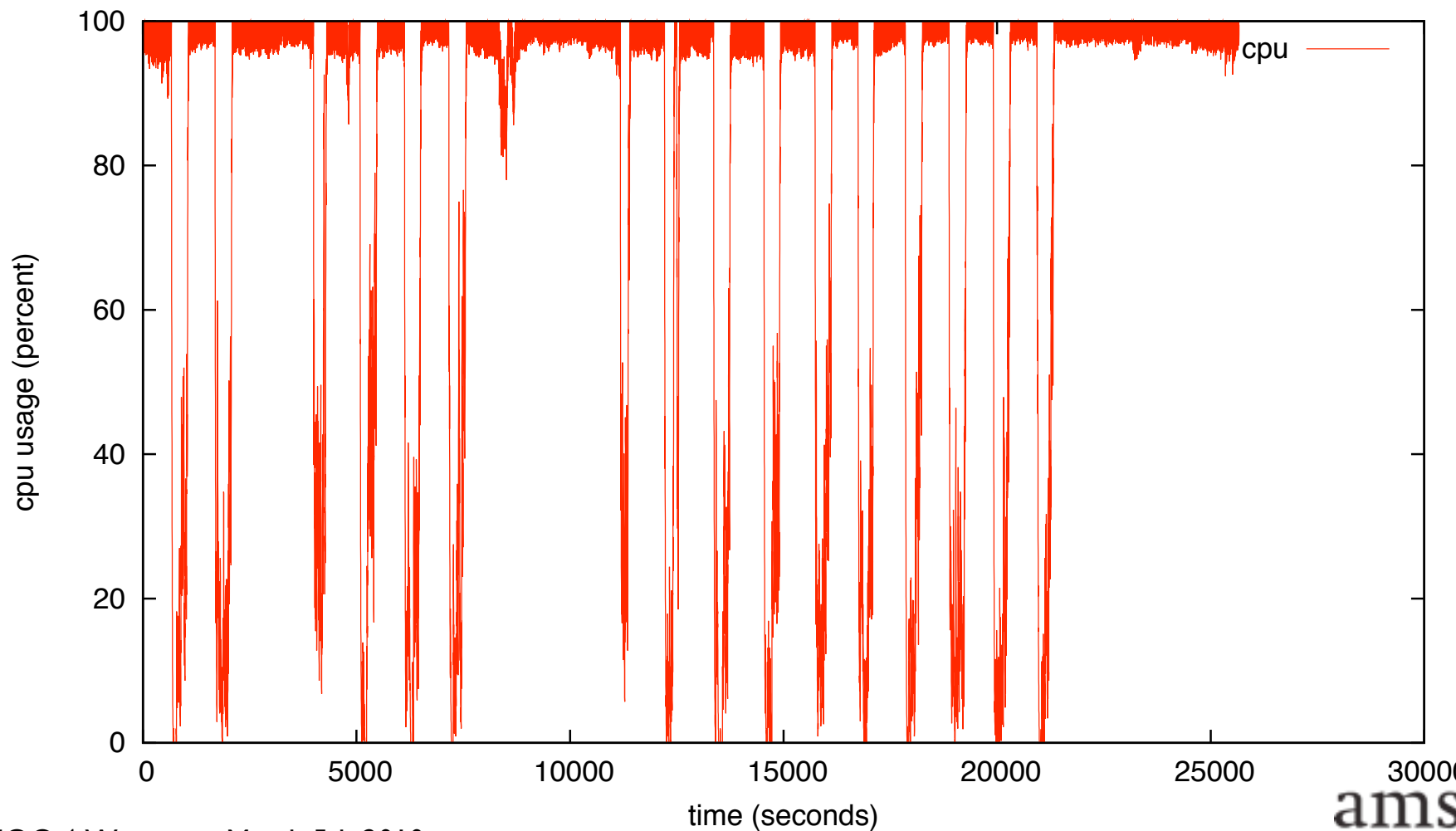
- 100 sessions, set up from Ixia
- 500 or 1000 prefixes per session
- Additional random flapping

Quagga

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

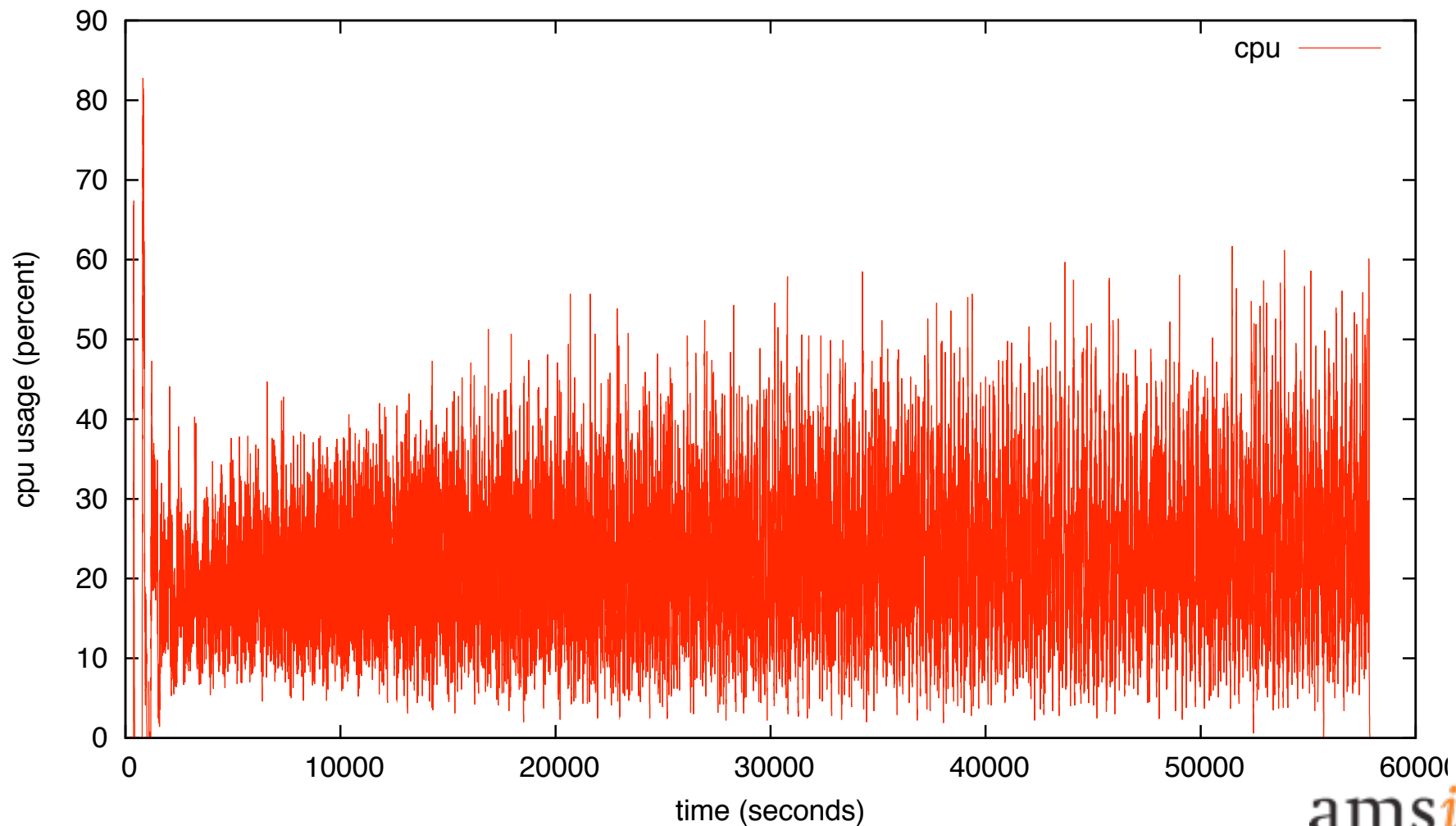
Quagga CPU

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



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



OpenBGPd CPU

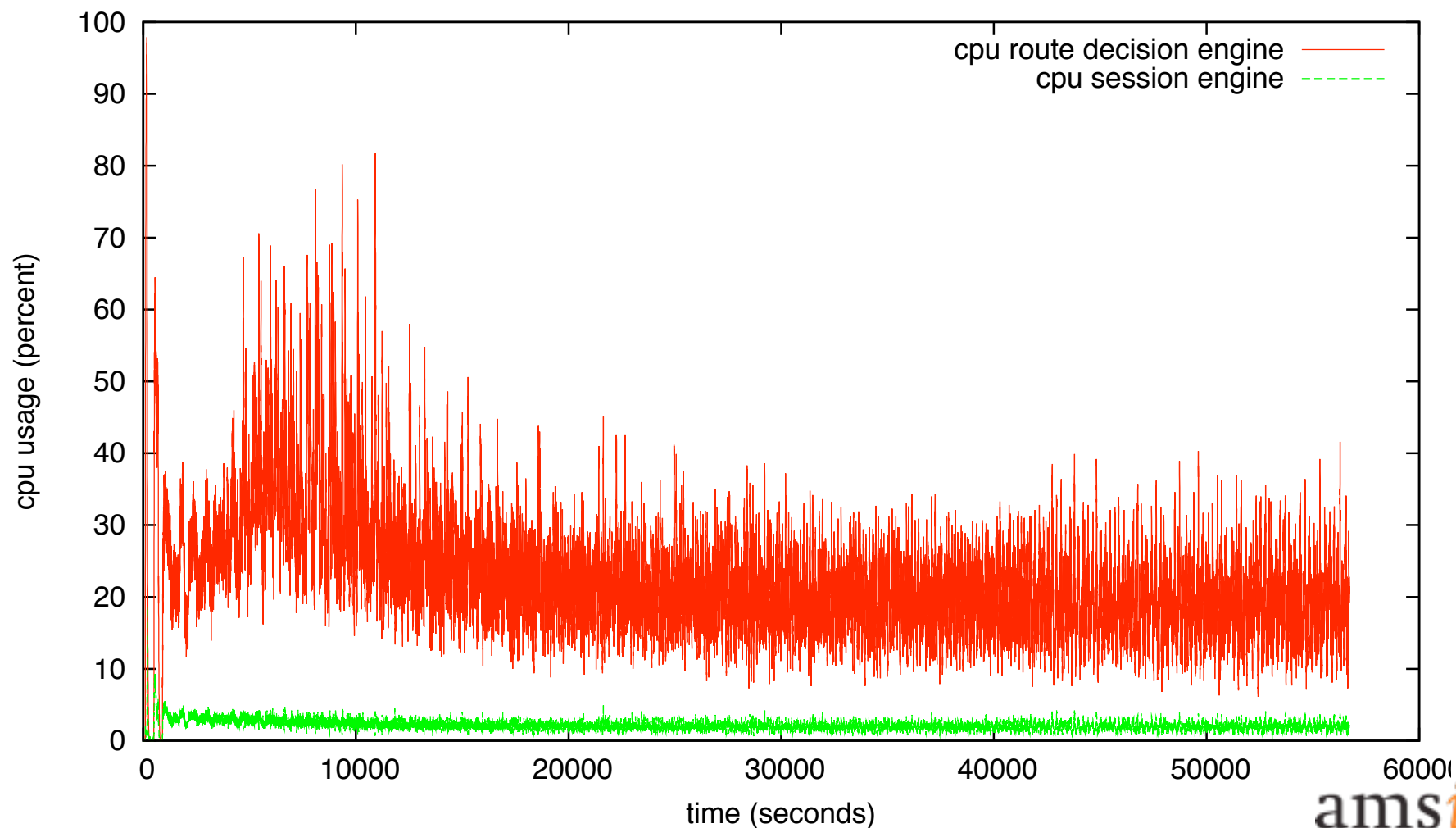
OpenBGPd cpu usage
multiple-rib; 100 sessions

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



OpenBGPD

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

OpenBGPd Mem

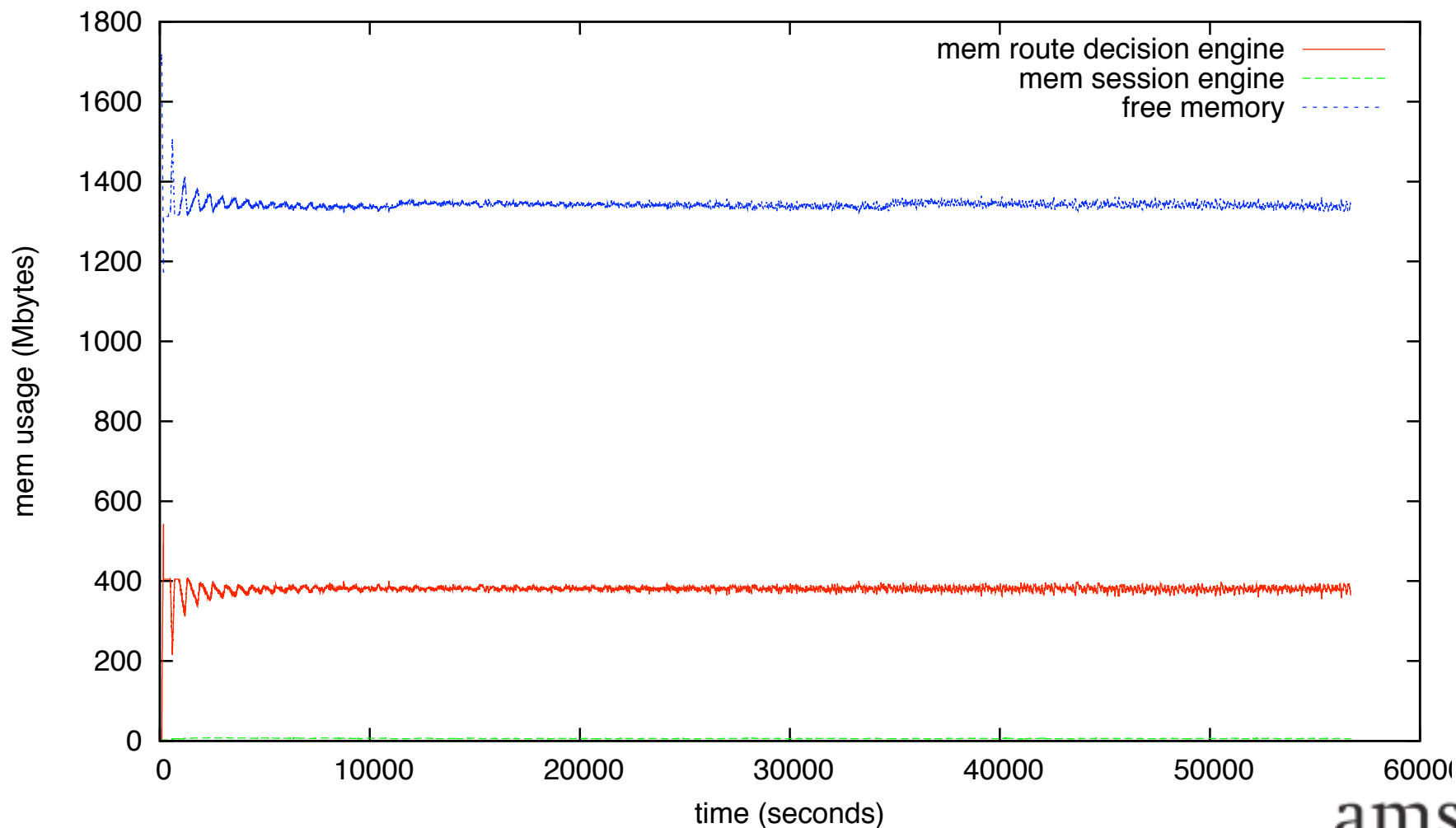
OpenBGPd mem usage
multiple-rib; 100 sessions

IXIA vs. OpenBGPd

1000 prefixes per session

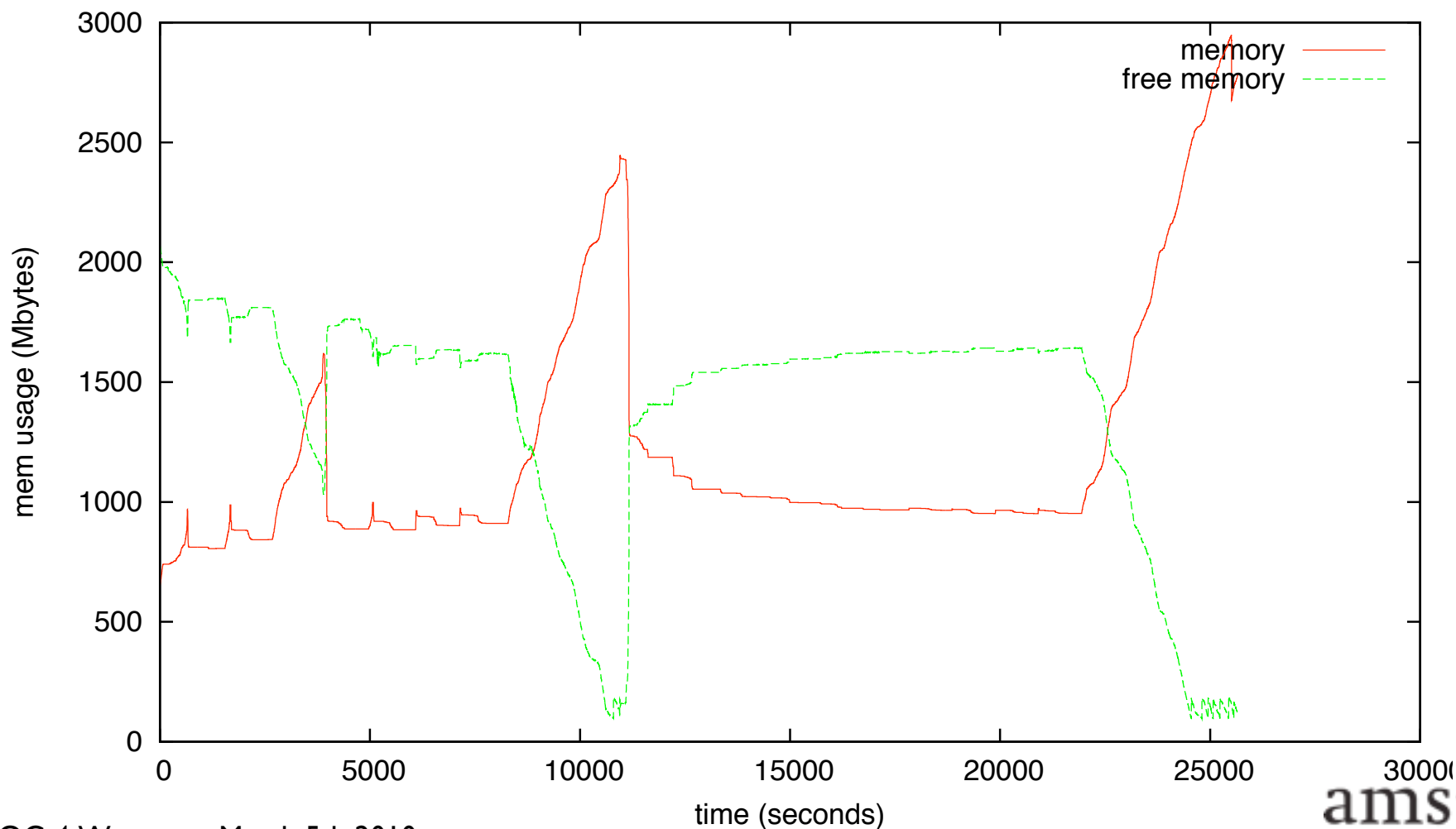
lab2.paix.net

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

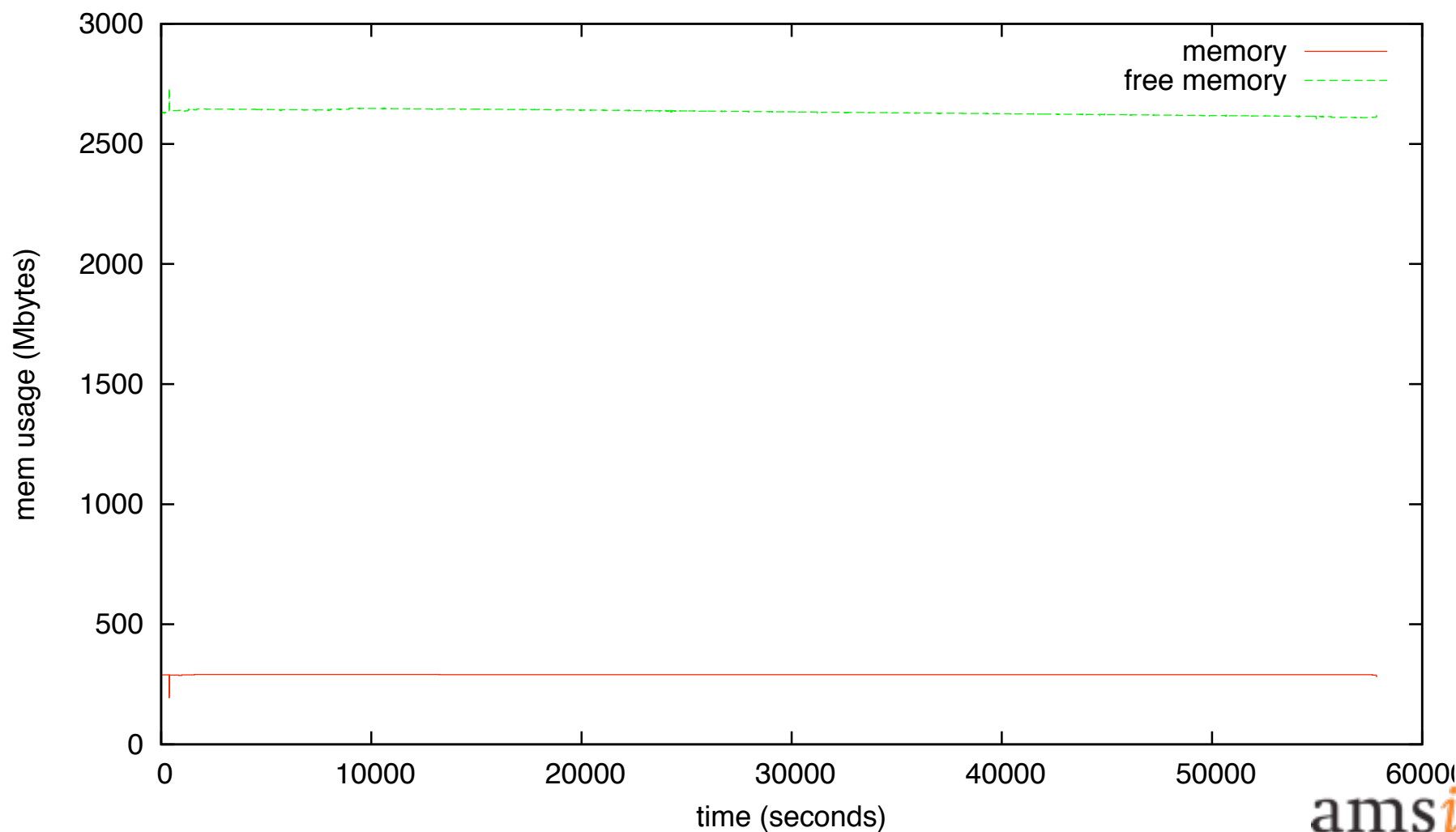


BIRD

- 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



Thank you!

Questions?

<elzbieta.jasinska@ams-ix.net>

