RIPE

ExaBGP

How to mess with BGP

Thomas Mangin, RIPE 69, London

RIPE

ExaBGP is not a BGP daemon

- And not a recursive acronym !
  - Overview presented yesterday on Open Source WG
  - Introduce “common” usage
- It can
  - establish or accept BGP connections
  - generate or parse BGP packets
  - be controlled from a third party program
  - send raw / parsed updates to it too
  - …

ExaBGP: think BGP Engine

- Gaming engine
  - let you write Games

- BGP engine let you write
  - let you write custom BGP applications

- Yes, you can “simply” send routes
  - using the configuration file

Programming API

- Unix PIPE, like on the shell

  #> tail -f logfile | grep pattern

  #> producer | consumer

- ExaBGP
  - forks the “consumer” for you
  - will restart it on failure (but it should never exit)
  - can accept command from it too
    (two ways communication)

JSON Example

#> ./sbin/exabgp conf --decode "0000 0022 4001 0100 4002 0602 0100 000D 1C40 0304 3209 0000 4005 0400 0000 64C0 0804 0190 00C8 0000 0001 18C8 0100 0000 0001 18C8 0101”

{'exabgp': '3.4.0',

'neighbor': {'ip': '17.0.142.1',

'message': {'**update**': {'announce': {'ipv4 unicast': {

'50.9.0.0': {

'200.1.0.0/24': {'path-information': '0.0.0.1'},

'200.1.1.0/24': {'path-information': '0.0.0.1'}}}},

'attribute': {'as-path': [3356], 'atomic-aggregate': False,

'community': [[400, 200]], 'local-preference': 100, 'origin': 'igp'}}}},

'pid': ‘64488', 'ppid': ‘55736', 'time': 1415031653,

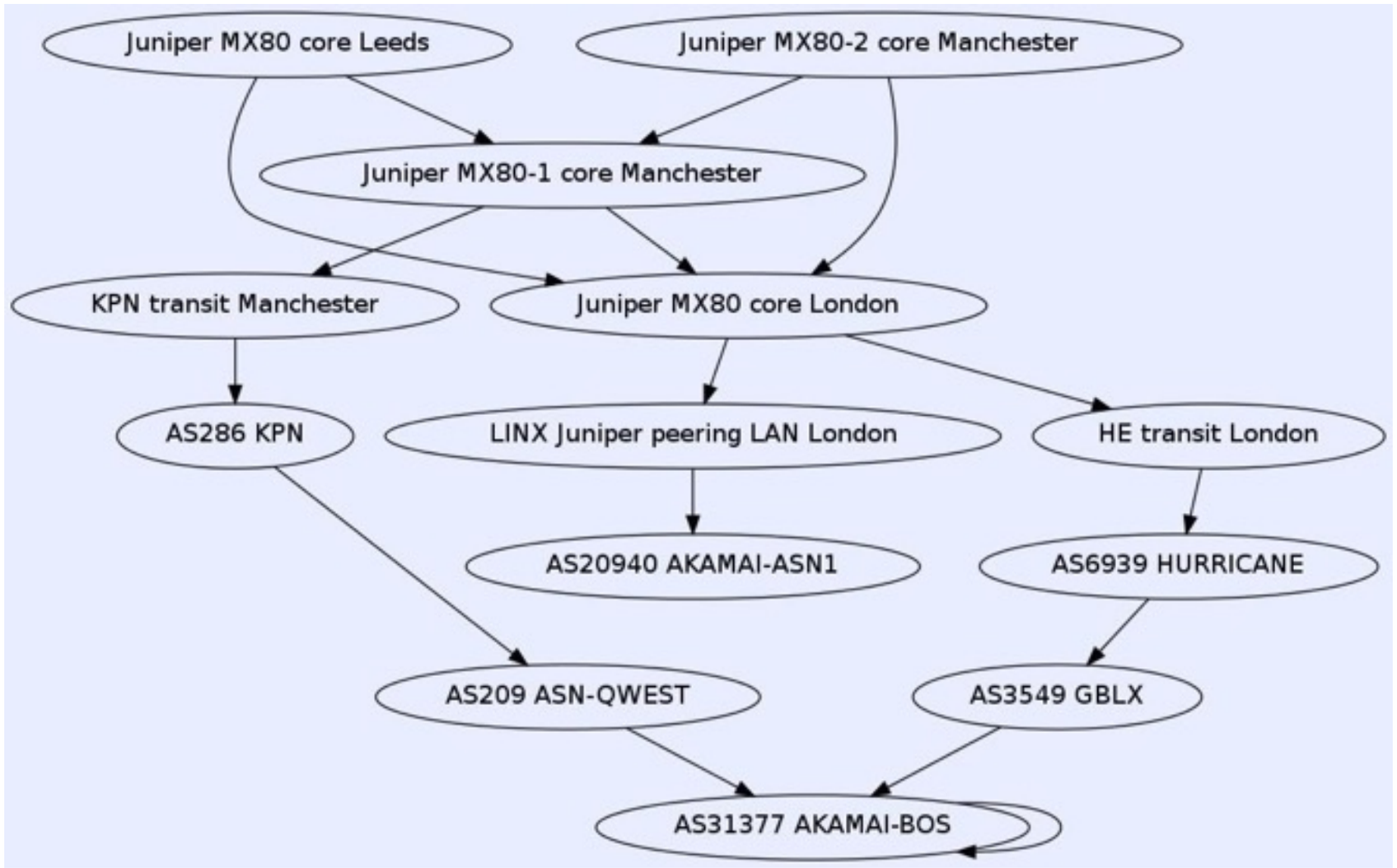
'type': ‘**update**’, 'host': ‘dhcp-26-181.ripemtg.ripe.net' }

Programming API

- TEXT - simplest
  - works both way
    - announce/withdraw routes to peers
    - get notified of the peer activity
    - announce route 10.0.0.1/24 next-hop 10.10.1.1
- JSON - newest
  - only messages received from the peers
  - no control via JSON … yet

What can you do with it ?

- write tools to find how your routing really works
  - https://github.com/dpiekacz/gixlg



Stress Test other implementation

- Simulate large EBGP network
  - reconnection following outage on a peering LAN
  - worse case scenarios …
  - you name it …
- Launch many BGP speaker all at once

```
#> export exabgp_tcp_delay=60
#> for config in `ls *.conf`; do
./sbin/exabgp $config &
done
```

Stress Test other implementation

- Worse case : random updates

```
#!/usr/bin/env python
from random import Random
r = Random()
while True:
args,path = [],[]
for _ in range(8): args.append(r.randint(0,255))
args.extend((r.randint(0,65000),r.randint(0,65000)))
for _ in range(r.randint(1,10)): path.append(str(r.randint(1,65000)))
args.append(' '.join(path))
print "announce route %d.%d.%d.%d next-hop %d.%d.%d.%d med %d
local-pref %d as-path [ %s ] " % tuple(args)
```

Generic Attributes

- Attribute must be valid
  - no overflow testing
- Can contain anything
  - route … med 100;
  - route … attribute [0x04 0x80 0x00000064];
- Make a transitive MED?
  - route … attribute [0x04 0xC0 0x00000064];
  - 0xC0 = 0x80 (optional) + 0x40 (transitive)
- Generate attribute for new unsupported RFC
- Break your favorite vendor

HA with BGP

- Core load balanced traffic
  - multiple servers announcing the same IP
  - on service failure, stop IP announcement
  - use core routers to balance the TCP flows
  - eliminate dedicated Load balancers

http://bits.shutterstock.com/2014/05/22/stop-buying-load-balancers-and-start-controlling-your-traffic-flow-with-software/

HA with BGP

- Control services IP
  - migrate an IP on service failure
    - hosts announce services IP using BGP
    - on service failure, the announcement is removed
    - another host takes on the traffic
  - Allows to fail service between DC at Layer 3

http://vincent.bernat.im/en/blog/2013-exabgp-highavailability.html

DDOS

- ExaBGP implement Flow Spec
  - announcement and recently decoding too

- Often used to stop “simple” amplication DDOS
  - Support next-hop rewrite

- Started another project ExaDDOS to
  - automate DDOS detection and response

Please help !

- ExaBGP is a “personal” project
  - competing for time with my work at Exa, LINX, IXLeeds, my family, Jui Jitsu, and need to sleep …
- I welcome contributors
  - I scan for fork and often merge and “fix” the code provided.
  - I am happy to spend some time per mail / IM / video conf to explain the code

Other relevant information

- Other BGP implementation
  - https://github.com/Exa-Networks/exabgp/wiki/Other-OSS-BGP-implementations

- My personal email / jabber address
  - first@last.com

- All my previous ExaBGP presentations
  - http://thomas.mangin.com/data/pdf/

Questions?

(or comments)