New Developments
in ExaBGP
Why should YOU care ?

LINX 83
18th/19th of November 2013

Thomas Mangin
Exa Networks

exa networks

Whatever a speaker is missing in depth he will compensate for in length

Montesquieu

Another presentation

to ignore while you have fun on IRC !

Another presentation between you and BEER !

Doing BGP with OSS

Well known open source implementations of BGP

Quagga	http://bird.network.cz/
BIRD	http://www.quagga.net/

The underdog

ExaBGP	https://github.com/Exa-Networks/exabgp

Another UK born and bred

BGPFeeder	https://projects.bytemark.co.uk/projects/bgpfeeder

And the others

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

A little learning is a dangerous thing
Alexander Pope

ExaBGP ..

A “BGP swiss army knife” since 2009..

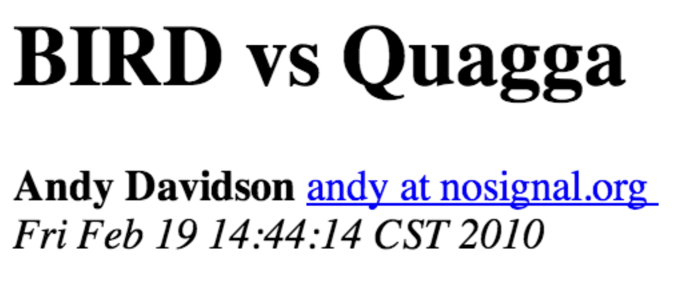
```
commit 5490f7baf5981279e2360d88c735570bc9f72532
Author: Thomas Mangin <thomas.mangin@exa-networks.co.uk>
Date:   Thu Sep 3 22:12:05 2009 +0000

initial commit [...] announce a route to a 7204 and keep the connection alive
```

Patience is bitter, but its fruit is sweet
Rousseau

ExaBGP?

NANOG Thread



BIRD vs Quagga

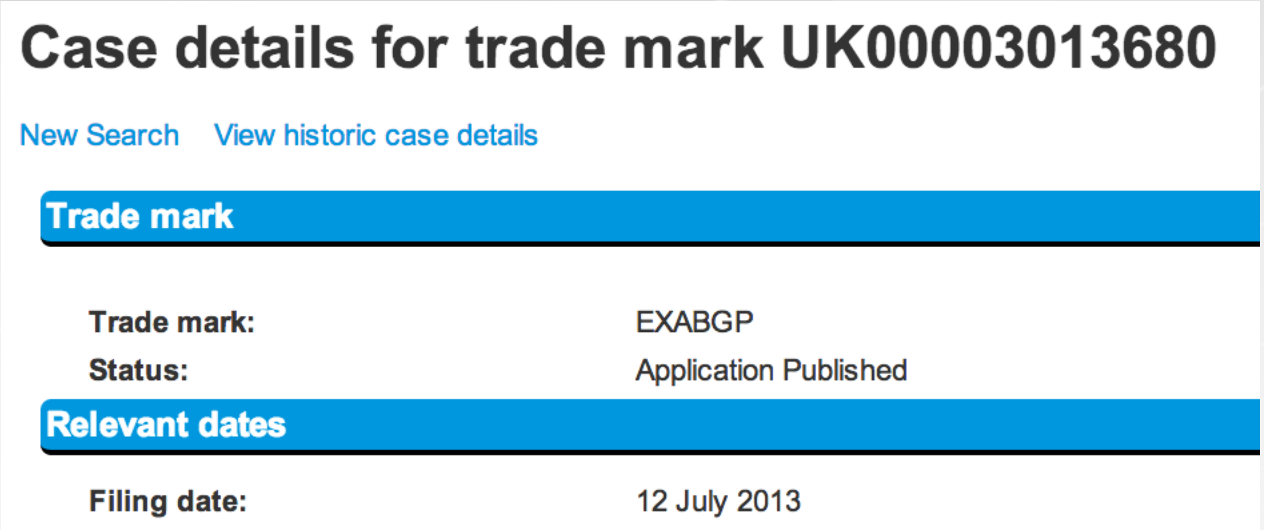
Andy Davidson andy at nosignal.org
Fri Feb 19 14:44:14 CST 2010

Andy's marketing services

[…] you might find **ExaBGP** more lightweight in this role – see **http://bgp.exa.org.uk/** – do check it out. This has an interface which will feel extremely comfortable to Juniper users.

Best wishes
Andy

Genius …



Case details for trade mark UK00003013680

New Search   View historic case details

Trade mark

Trade mark:	EXABGP
Status:	Application Published

Relevant dates

Filing date:	12 July 2013

We liked it so much we trademarked it!

Pride is the consolation of the weak
Vauvenargues

Let’s work on that marketing

ExaBGP

“SDN without marketing”
“SDN on commodity hardware”

ExaBGP

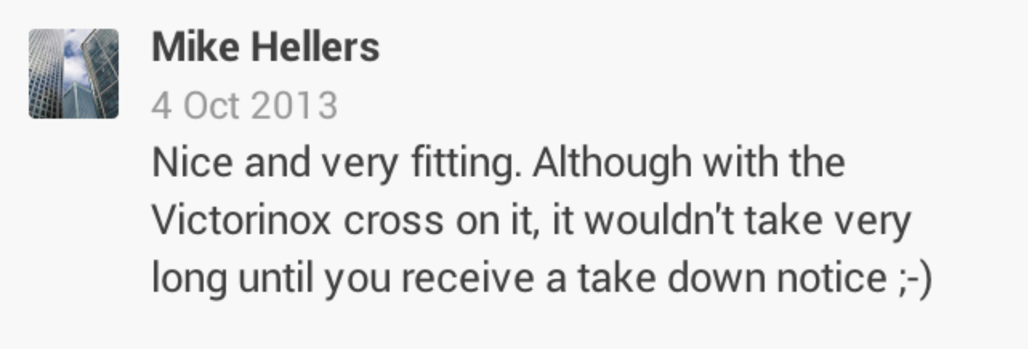
“The BGP swiss army knife of networking”

no new suggestions required

exa networks

Truth is more valuable if it takes you a few years to find it.
Renard

Thomas’ idea



Mike Hellers

4 Oct 2013

Nice and very fitting. Although with the Victorinox cross on it, it wouldn't take very long until you receive a take down notice ;-)

Thank you Mike …

I expected Malcolm to bring me this kind of bad news

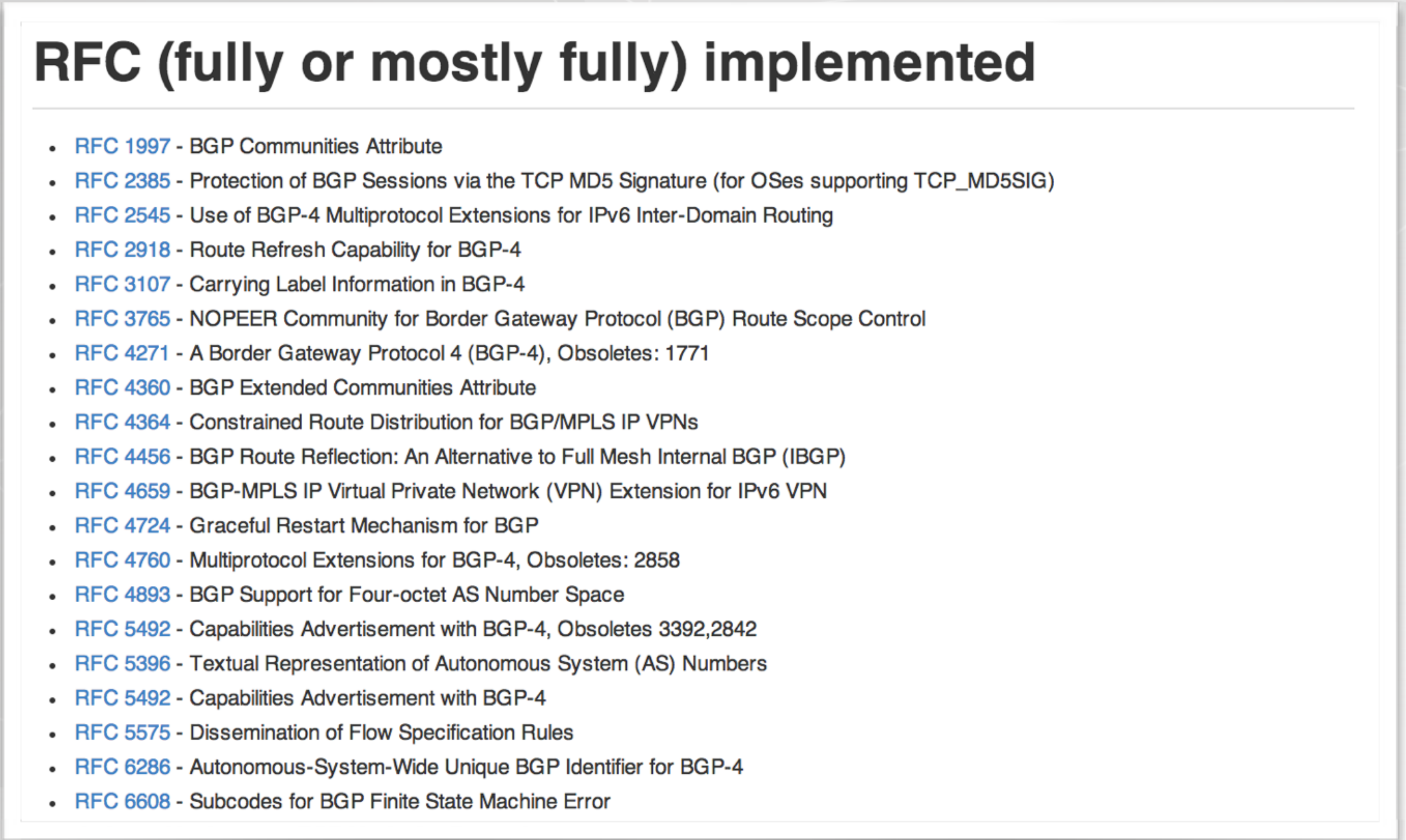
Back to square one !

Real knife by Victorinox AG

I have always believed that to succeed in life, it is necessary to appear to be mad and to act wisely

Montesquieu

Any Good ?



RFC (fully or mostly fully) implemented

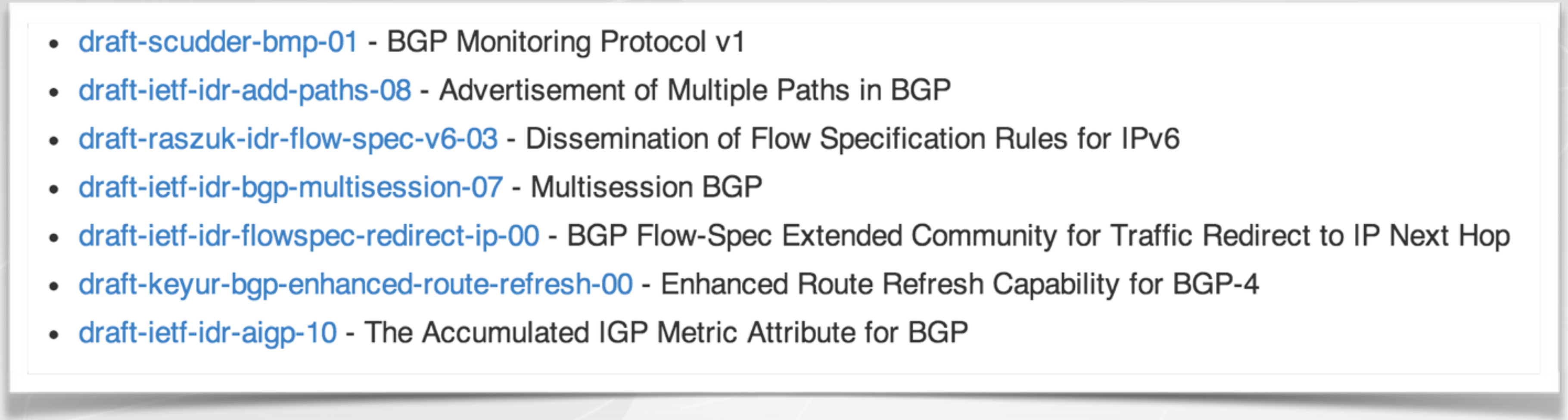
- RFC 1997 - BGP Communities Attribute
- RFC 2385 - Protection of BGP Sessions via the TCP MD5 Signature (for OSes supporting TCP_MD5SIG)
- RFC 2545 - Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing
- RFC 2918 - Route Refresh Capability for BGP-4
- RFC 3107 - Carrying Label Information in BGP-4
- RFC 3765 - NOPEER Community for Border Gateway Protocol (BGP) Route Scope Control
- RFC 4271 - A Border Gateway Protocol 4 (BGP-4), Obsoletes: 1771
- RFC 4360 - BGP Extended Communities Attribute
- RFC 4364 - Constrained Route Distribution for BGP/MPLS IP VPNs
- RFC 4456 - BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)
- RFC 4659 - BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN
- RFC 4724 - Graceful Restart Mechanism for BGP
- RFC 4760 - Multiprotocol Extensions for BGP-4, Obsoletes: 2858
- RFC 4893 - BGP Support for Four-octet AS Number Space
- RFC 5492 - Capabilities Advertisement with BGP-4, Obsoletes 3392,2842
- RFC 5396 - Textual Representation of Autonomous System (AS) Numbers
- RFC 5492 - Capabilities Advertisement with BGP-4
- RFC 5575 - Dissemination of Flow Specification Rules
- RFC 6286 - Autonomous-System-Wide Unique BGP Identifier for BGP-4
- RFC 6608 - Subcodes for BGP Finite State Machine Error

Nothing is more humiliating than to see idiots succeed in enterprises we have failed at

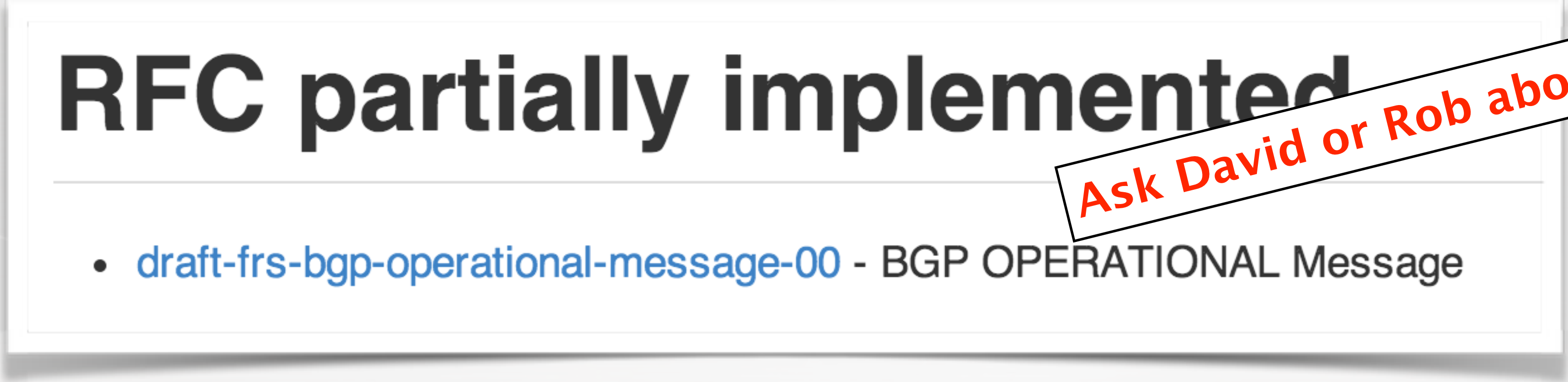
Flaubert

Up to date ?

Oh yeah baby!



- draft-scudder-bmp-01 - BGP Monitoring Protocol v1
- draft-ietf-idr-add-paths-08 - Advertisement of Multiple Paths in BGP
- draft-raszuk-idr-flow-spec-v6-03 - Dissemination of Flow Specification Rules for IPv6
- draft-ietf-idr-bgp-multisession-07 - Multisession BGP
- draft-ietf-idr-flowspec-redirect-ip-00 - BGP Flow-Spec Extended Community for Traffic Redirect to IP Next Hop
- draft-keyur-bgp-enhanced-route-refresh-00 - Enhanced Route Refresh Capability for BGP-4
- draft-ietf-idr-aigp-10 - The Accumulated IGP Metric Attribute for BGP



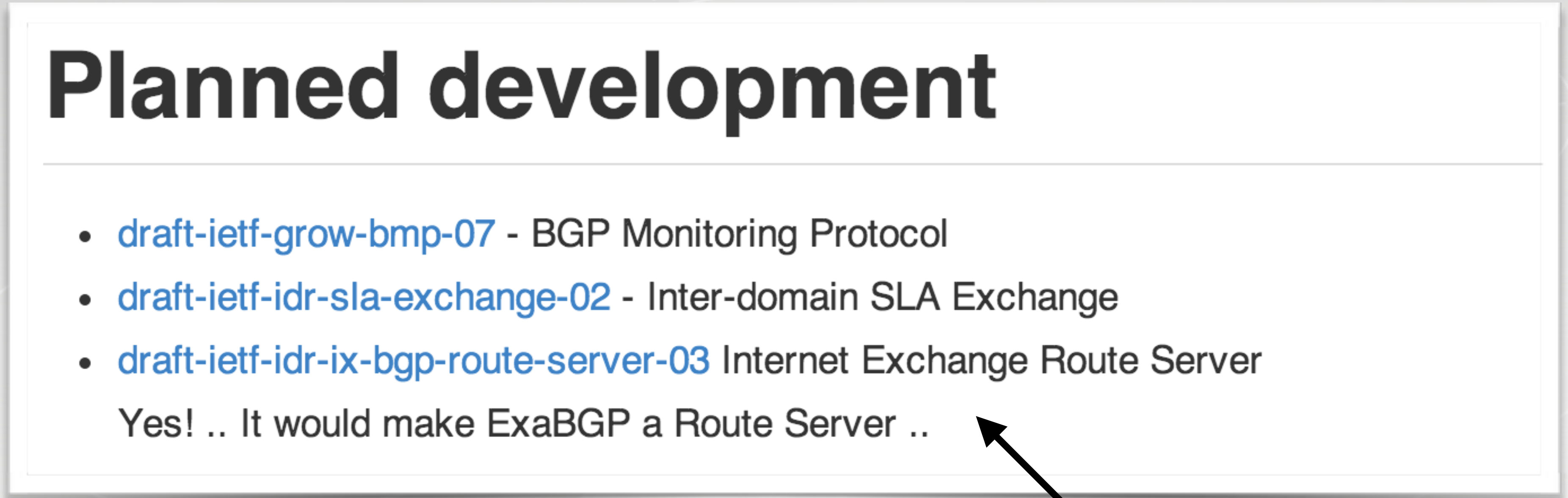
RFC partially implemented

- draft-frs-bgp-operational-message-00 - BGP OPERATIONAL Message

Ask David or Rob about it …

I love fools' experiments. I am always making them.
Charles Darwin

What next?



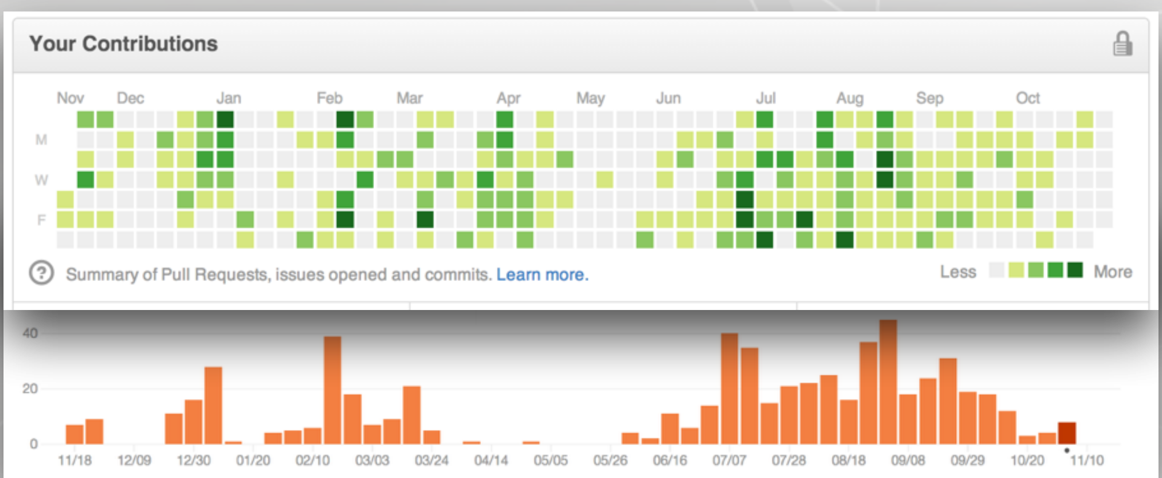
Planned development

- draft-ietf-grow-bmp-07 - BGP Monitoring Protocol
- draft-ietf-idr-sla-exchange-02 - Inter-domain SLA Exchange
- draft-ietf-idr-ix-bgp-route-server-03 Internet Exchange Route Server

  Yes! .. It would make ExaBGP a Route Server ..

I will focus on that…
later .. way later in the talk

For when?



Your Contributions

Summary of Pull Requests, issues opened and commits. Learn more.

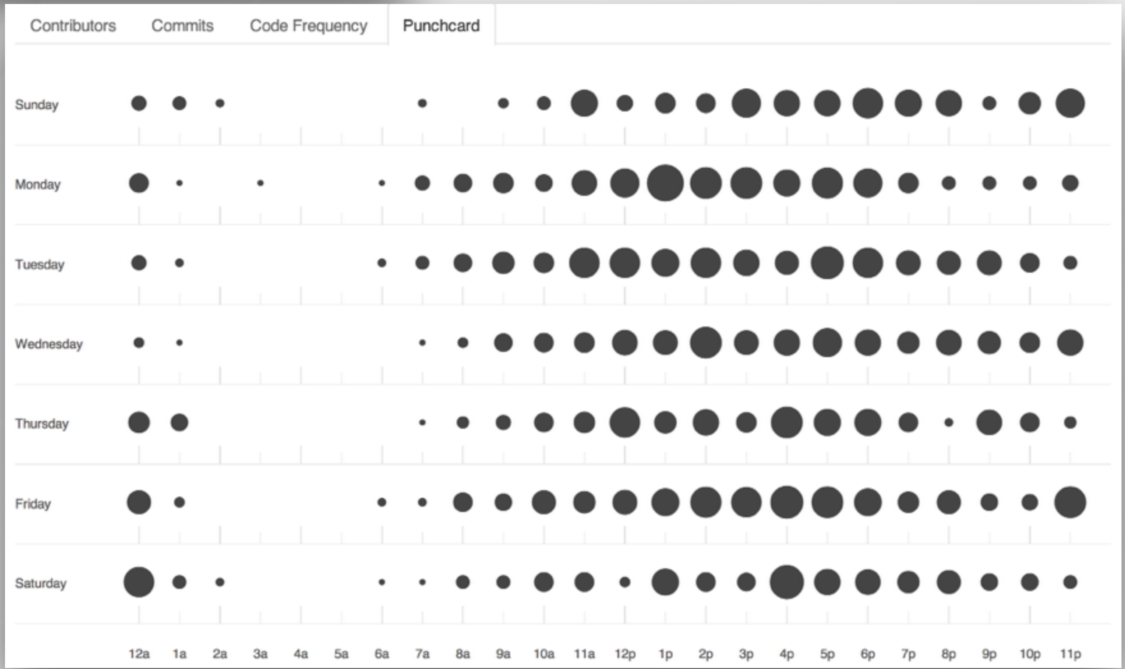
I am taking a small break…

This is my “hobby”
be kind I have **three** jobs

A hobby which gets

- Heidi complaining
- My colleagues too
  (I can ignore these)

Therefore **ExaBGP Users are NOT allowed to complain**!



What's the expected use?

NOC usage ..

DDOS RTBH : prevents bad traffic from reaching its destination
Flow Spec : RTBH on steroid, firewall rules deployed using BGP
Interception : Legal requirements (IWF,…)
SDN : over 200k routes updates every 5 minutes ..

DevOps usage ..

Service IPs : servers mobility using extra/32 with BGP
Anycast : the same IP at different locations (CDN, DNS, ...)

IX usage ..

Collector : at IXLeeds
Route Server : future development needed

Be regular and orderly in your life, so that you may be violent and original in your work
Flaubert

Easy to install?

Use GitHub

```
> wget https://github.com/Exa-Networks/exabgp/archive/3.2.17.tar.gz
> tar zxvf 3.2.17.tar.gz
> cd exabgp-3.2.17
> ./sbin/exabgp —help
```

Use your distribution (often older code)

```
> apt-get install exabgp        # Debian / Ubuntu
> pacman -S exabgp              # ArchLinux
> port install exabgp           # OS X / FreeBSD
> emerge exabgp                 # Gentoo (soon? Thank you Tony)
```

exa networks

Be regular and orderly in your life, so that you may be violent and original in your work

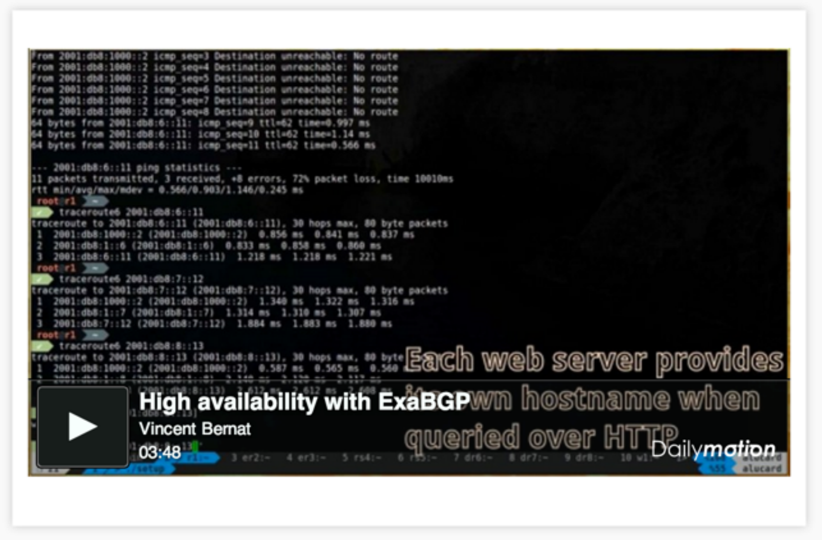
Flaubert

Easy to use?

Not as easy as it could be

No real documentation

Help welcome...



```
From 2001:db8:1000::2 icmp_seq=3 Destination unreachable: No route
From 2001:db8:1000::2 icmp_seq=4 Destination unreachable: No route
From 2001:db8:1000::2 icmp_seq=5 Destination unreachable: No route
From 2001:db8:1000::2 icmp_seq=6 Destination unreachable: No route
From 2001:db8:1000::2 icmp_seq=7 Destination unreachable: No route
From 2001:db8:1000::2 icmp_seq=8 Destination unreachable: No route
64 bytes from 2001:db8:6::11: icmp_seq=9 ttl=62 time=0.997 ms
64 bytes from 2001:db8:6::11: icmp_seq=10 ttl=62 time=1.14 ms
64 bytes from 2001:db8:6::11: icmp_seq=11 ttl=62 time=0.566 ms

--- 2001:db8:6::11 ping statistics ---
11 packets transmitted, 3 received, +8 errors, 72% packet loss, time 10010ms
rtt min/avg/max/mdev = 0.566/0.903/1.146/0.245 ms
 root  r1  ~
 ✓  traceroute6 2001:db8:6::11
traceroute to 2001:db8:6::11 (2001:db8:6::11), 30 hops max, 80 byte packets
 1  2001:db8:1000::2 (2001:db8:1000::2)  0.856 ms  0.841 ms  0.837 ms
 2  2001:db8:1::6 (2001:db8:1::6)  0.833 ms  0.858 ms  0.860 ms
 3  2001:db8:6::11 (2001:db8:6::11)  1.218 ms  1.218 ms  1.221 ms
 root  r1  ~
 ✓  traceroute6 2001:db8:7::12
traceroute to 2001:db8:7::12 (2001:db8:7::12), 30 hops max, 80 byte packets
 1  2001:db8:1000::2 (2001:db8:1000::2)  1.340 ms  1.322 ms  1.316 ms
 2  2001:db8:1::7 (2001:db8:1::7)  1.314 ms  1.310 ms  1.307 ms
 3  2001:db8:7::12 (2001:db8:7::12)  1.884 ms  1.883 ms  1.880 ms
 root  r1  ~
 ✓  traceroute6 2001:db8:8::13
traceroute to 2001:db8:8::13 (2001:db8:8::13), 30 hops max, 80 byte
 1  2001:db8:1000::2 (2001:db8:1000::2)  0.587 ms  0.565 ms  0.560
```

Each web server provides
wn hostname when
queried over HTTP

The community is stepping up !

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

DDOS http://media.frnog.org/FRnOG_18/FRnOG_18-6.pdf

Be regular and orderly in your life, so that you may be violent and original in your work

Flaubert

I can hear Martin Levy ask “Does it supports IPv6 ”

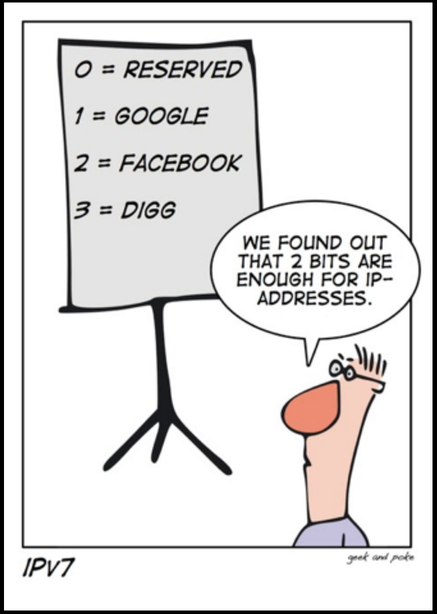
Native **IPv6**
Themed slide

IPv4	Neighbours	yes
IPv6	Neighbours	yes
IPv4	Prefixes (and MPLS)	yes
IPv6	Prefixes (MP NLRI)	yes
IPv4	Flow Spec (RFC 5575)	yes
IPv6	Flow Spec (draft)	yes *

* I do not know any vendors supporting it yet …

As you can never fully please Martin, I admit …

RFC 5701 – IPv6 Address Specific BGP Extended Community Attribute no



Usage RTBH

Tell your provider to stop sending you traffic for some IPs

Announce some more specific routes (/32, /29, …) part of your network
and TAG the route with communities
so it can be filtered (dropped by your upstream edge routers)
Traffic is dropped before it is billed

Many Talks (NANOG, APRICOT, ...) on the topic and an RFC (5635)
> google RTBH or REMOTELY TRIGGERED BLACKHOLE

The goal is to bypass the transit provider NOC and reduce response time when under duress

Each ISP implements it differently ..
level3 > whois –h whois.ripe.net AS3356 | grep –B1 –A15 –i blakhole

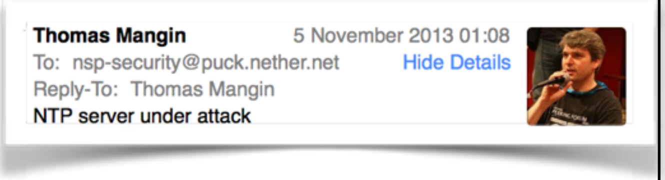
exa networks

It is dangerous to be right in matters on which the established authorities are wrong
Voltaire

Flow Routes

Control the filtering Yourself, do not disconnect the target

```
group ddos {
  local-as 30740;
  peer-as 30740;
  router-id 82.219.0.1;
  local-address 82.219.0.1;
  graceful-restart 5;
  family {
    ipv4 unicast;
    ipv4 flow;
  }
 flow {
    route drop-ddos-ntp2 {
      match {
        destination 82.219.4.31/32;
        destination-port >123 <123;
        protocol udp;
      }
      then {
        discard;
      }
    }
  }
  neighbor 82.219.0.2 {
    description “nothing at those IP";
  }
  neighbor 82.219.0.3 {
    description “no point attacking them";
  }
}
```



Thomas Mangin 5 November 2013 01:08

To: nsp-security@puck.nether.net

Reply-To: Thomas Mangin

NTP server under attack

Firewall rules via BGP
RFC 5575

Juniper and Alcatel
Cisco coming in 2014
for IOS-XR and XE
Ask Cisco for more info

ExaBGP is the only OSS application to support FlowSpec

```
thomas@mx-80> show route table inetflow.0

inetflow.0: 3 destinations, 3 routes (3 active, 0 holddown, 0 hidden)
Restart Complete
+ = Active Route, - = Last Active, * = Both

82.219.4.31,*,proto=17,dstport>=124&<=65535,>=0&<=122/term:2
                 *[BGP/170] 4d 13:48:20, localpref 100, from 82.219.5.101
                    AS path: I
                    Fictitious
[...]
thomas@mx-80> show firewall filter __flowspec_default_inet__
```

The secret of business is to know something that nobody else knows
Aristotle Onassis

exa networks

Designed to be scripted

```
neighbor 127.0.0.1 {
        router-id 1.2.3.4;
        local-address 127.0.0.1;
        local-as 1;
        peer-as 1;
        graceful-restart;

        process announce-routes {
                run ./api-add-remove.run;
        }
```

```
#!/usr/bin/env python

import sys, time

messages = [
'announce route 1.1.0.0/24 next-hop 101.1.101.1',
'announce route 1.1.0.0/25 next-hop 101.1.101.1',
'withdraw route 1.1.0.0/24 next-hop 101.1.101.1',
]

while messages:
        message = messages.pop(0)
        sys.stdout.write( message + '\n')
        sys.stdout.flush()
        time.sleep(1)

while True:
        time.sleep(1)
```

```
> ./sbin/exabgp ./api-add-remove.conf
```

Use ANY scripting language
perl, python, lua, go, bash, …

An example on the wiki with
SHELL PIPE ..

for examples, look into
/dev/runtest
“the test suite”

Used in prod as SDN
by at least one large network

Use for **DDOS** mitigation
by **MANY networks**

Used by vendor
For BGP interrop testing !

Their is two rules for success in business, one do not tell all you know, …

Some bad joke site

ExaBGP as a Route Server

Why only now?

- ExaBGP started as a route injector, not a BGP daemon
  - It is single threaded using windows 3.1 like multi-tasking
  - The code was blocking when sending routes
  - Fixed this summer with version 3.2
  - Hundreds of hours of work
- Most of the IX effort already on Quagga and BIRD (more mature)

How much work is required ?

- ExaBGP already works as route collector
  - only tested on a small scale (IXLeeds)
  - need some more control features (for debugging)
  - but it SHOULD scale

Divide and Conquer
Julius Caesar

ExaBGP as a Route Server

Why would it be better?

- Much simpler code to understand (python)
- Much easier to hack (adding draft RFC in hours now)
- Can still be improved though

Can take benefit of multiple cores easily

- ExaBGP does NOT have a LOCAL RIB
- The RIB can be implemented as a different process
- The RIB does not even have to be on the server
- Possible madness with things like ZeroMQ :-)
- Possible to have one BGP daemon per switch
- Possible to detect L2 loss and change announcement

ExaBGP is single threaded but can use multiple cores

- FreeBSD and Linux 3.9 SO_REUSE_PORT
- Allows to split TCP flows to different process
- All listening on the same port

No change required to current ExaBGP
(but some improvement would help)

Divide and Conquer
Julius Caesar

exa networks

Last words… perhaps!

Please HELP!

I could do with …

more contributors

need **help with documentation**

Otherwise, just **let me know if you use it**…

Any ‘it works’ mail is always appreciated

Need to tidy some code

JSON generation

Configuration format parsing (started)

More ..

LINX agreed to let me use their IXIA to see how it performs

and compare the result with BIRD

who would be interested in seeing the results?

I am surprised! you are reading those quotes!

Thomas Mangin

exa networks

Questions?

Thank you for your kindness on IRC ..

thomas.mangin@exa-networks.co.uk https://github.com/thomas-mangin/exabgp/

Judge a man by his questions rather than by his answers
Voltaire