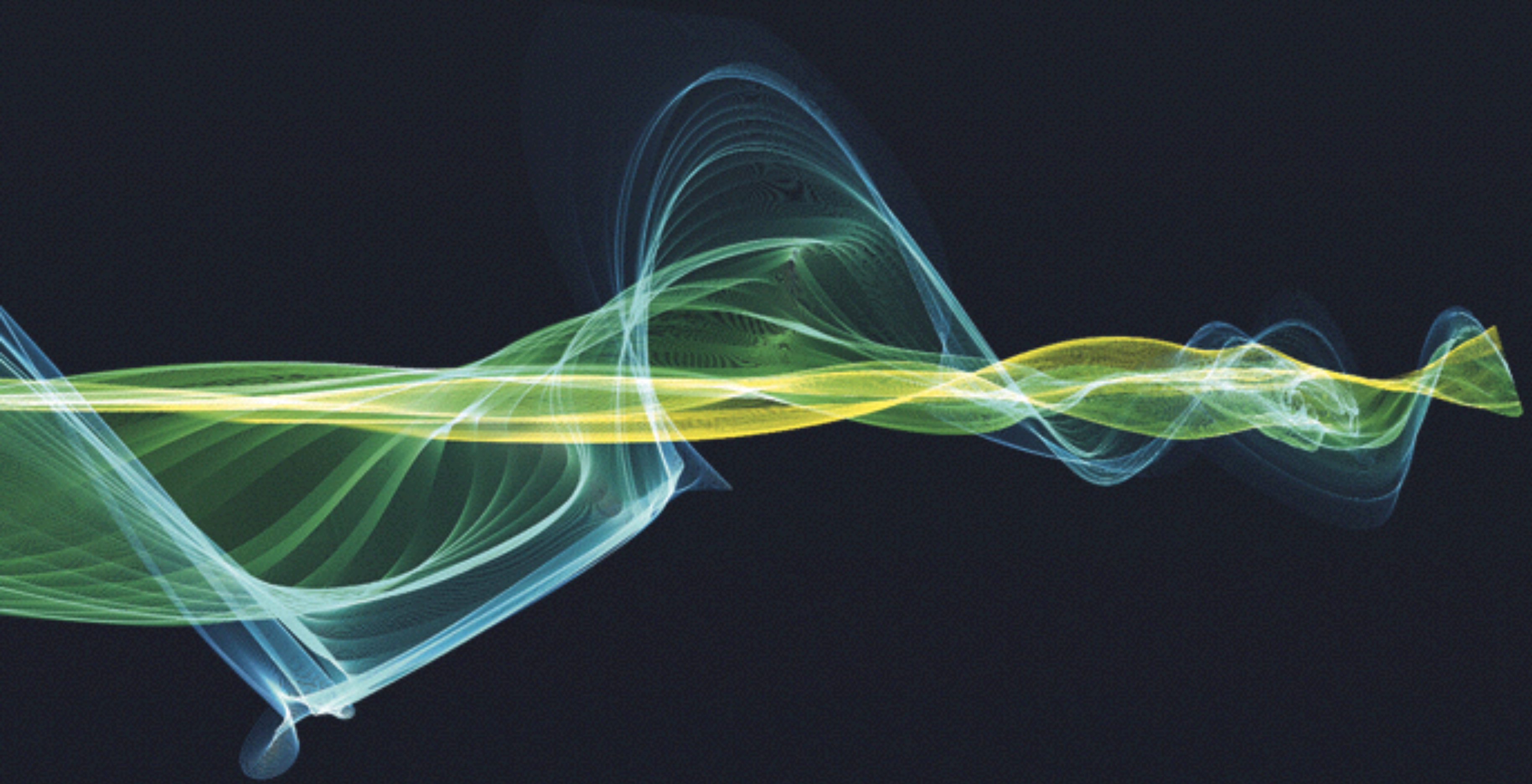
SysEleven
Hosting. Skaliert.

Protect your edge
BGP security made simple

Theo Voss

Technical Lead Network
SysEleven GmbH

DENOG8



Proud Sponsor of PeeringDB
SILVER

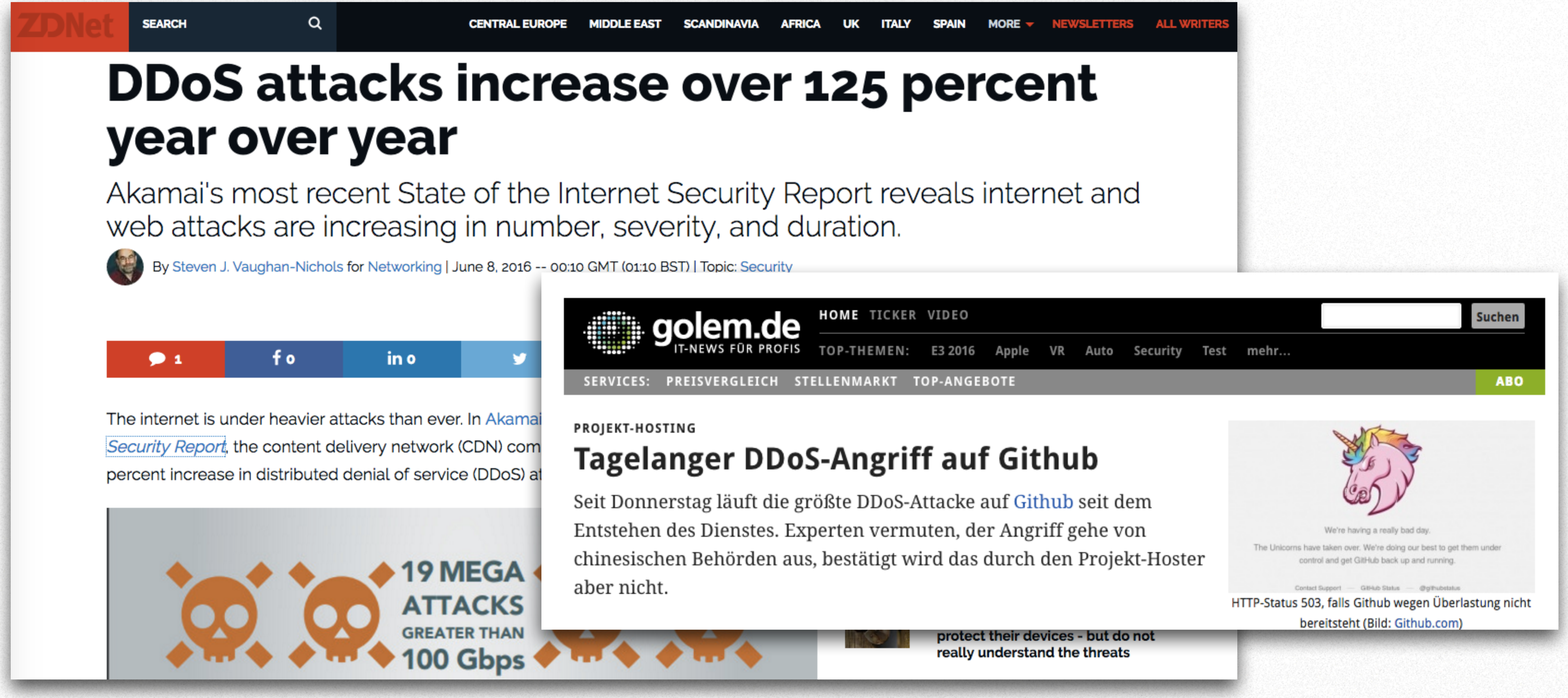
Who is SysEleven?

Managed Hoster & Upstream-Provider

300+ customers, 10 Points-of-Presence

t3n
we love technology
Wooga
HOTSPLOTS
ahd
your business.
DailyDeal
GRAVIS
EyeEm
McFIT
Einfach gut aussehen.
golem.de
IT-NEWS FÜR PROFIS
VERIVOX.DE
Da sparen Sie mit Sicherheit
HeLi.NET
neofonie*
Data. Technology. Services.
N26
+Babbel
HelloFresh
AMORELIE
traffics

Current Situation



DDoS attacks increase over 125 percent year over year

Akamai's most recent State of the Internet Security Report reveals internet and web attacks are increasing in number, severity, and duration.

By Steven J. Vaughan-Nichols for Networking | June 8, 2016 -- 00:10 GMT (01:10 BST) | Topic: Security

The internet is under heavier attacks than ever. In Akamai

Security Report, the content delivery network (CDN) com

percent increase in distributed denial of service (DDoS) at

protect their devices - but do not really understand the threats

PROJEKT-HOSTING

Tagelanger DDoS-Angriff auf Github

Seit Donnerstag läuft die größte DDoS-Attacke auf Github seit dem Entstehen des Dienstes. Experten vermuten, der Angriff gehe von chinesischen Behörden aus, bestätigt wird das durch den Projekt-Hoster aber nicht.

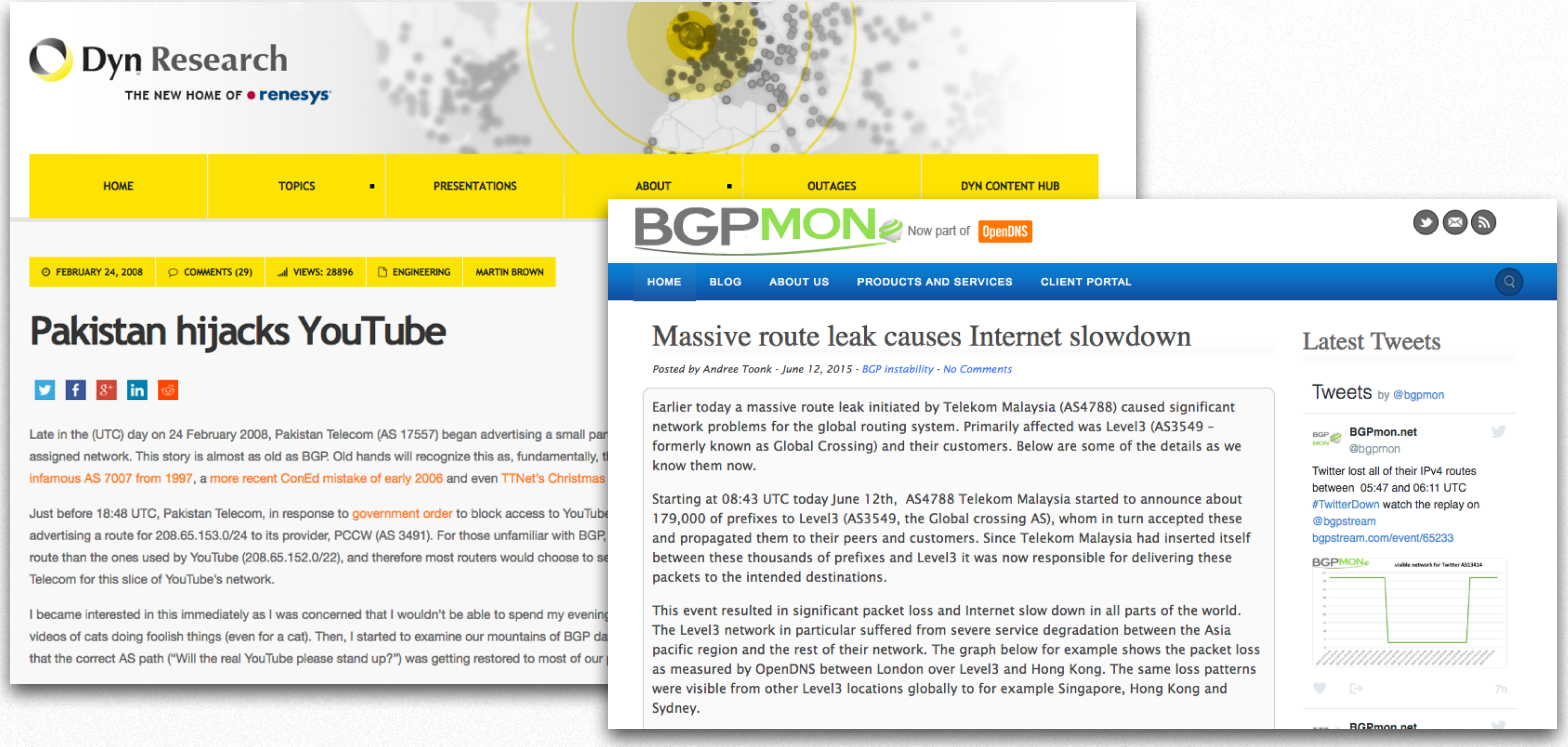
We're having a really bad day.

The Unicorns have taken over. We're doing our best to get them under control and get GitHub back up and running.

Contact Support — GitHub Status — @githubstatus

HTTP-Status 503, falls Github wegen Überlastung nicht bereitsteht (Bild: Github.com)

Current Situation



FEBRUARY 24, 2008 COMMENTS (29) VIEWS: 28896 ENGINEERING MARTIN BROWN

Pakistan hijacks YouTube

Late in the (UTC) day on 24 February 2008, Pakistan Telecom (AS 17557) began advertising a small par assigned network. This story is almost as old as BGP. Old hands will recognize this as, fundamentally, t infamous AS 7007 from 1997, a more recent ConEd mistake of early 2006 and even TTNet's Christmas

Just before 18:48 UTC, Pakistan Telecom, in response to government order to block access to YouTube advertising a route for 208.65.153.0/24 to its provider, PCCW (AS 3491). For those unfamiliar with BGP, route than the ones used by YouTube (208.65.152.0/22), and therefore most routers would choose to se Telecom for this slice of YouTube's network.

I became interested in this immediately as I was concerned that I wouldn't be able to spend my evening videos of cats doing foolish things (even for a cat). Then, I started to examine our mountains of BGP da that the correct AS path ("Will the real YouTube please stand up?") was getting restored to most of our

Massive route leak causes Internet slowdown

Posted by Andree Toonk - June 12, 2015 - BGP instability - No Comments

Earlier today a massive route leak initiated by Telekom Malaysia (AS4788) caused significant network problems for the global routing system. Primarily affected was Level3 (AS3549 – formerly known as Global Crossing) and their customers. Below are some of the details as we know them now.

Starting at 08:43 UTC today June 12th, AS4788 Telekom Malaysia started to announce about 179,000 of prefixes to Level3 (AS3549, the Global crossing AS), whom in turn accepted these and propagated them to their peers and customers. Since Telekom Malaysia had inserted itself between these thousands of prefixes and Level3 it was now responsible for delivering these packets to the intended destinations.

This event resulted in significant packet loss and Internet slow down in all parts of the world. The Level3 network in particular suffered from severe service degradation between the Asia pacific region and the rest of their network. The graph below for example shows the packet loss as measured by OpenDNS between London over Level3 and Hong Kong. The same loss patterns were visible from other Level3 locations globally to for example Singapore, Hong Kong and Sydney.

Latest Tweets

Tweets by @bgpmon

BGPmon.net @bgpmon

Twitter lost all of their IPv4 routes between 05:47 and 06:11 UTC #TwitterDown watch the replay on @bgpstream bgpstream.com/event/65233

The majority of ISPs in the world still filter on max-prefix limits at most and hope for the best.

Is filtering not easy enough?

HowTo’s or BCPs missing?

Routing Policy

We filter..

- Bogon ASNs
- Bogon prefixes
- IXP networks
- Own networks
- Prefix length
- Invalid prefixes

```
term REJECT-BOGON-ASN from as-path-group BOGON-ASN
term REJECT-BOGON-ASN then reject

term REJECT-BOGON-PREFIXES from prefix-list-filter BOGON-PREFIXES orlonger
term REJECT-BOGON-PREFIXES then reject

term REJECT-SYS11-PREFIXES from prefix-list-filter SYSELEVEN-NETWORKS orlonger
term REJECT-SYS11-PREFIXES then reject

term REJECT-IXP-NETWORKS from prefix-list-filter IXP-NETWORKS orlonger
term REJECT-IXP-NETWORKS then reject

term FILTER-PREFIX-LENGTH-1 from route-filter 0.0.0.0/0 prefix-length-range /0-/8
term FILTER-PREFIX-LENGTH-1 then reject

term FILTER-PREFIX-LENGTH-2 from route-filter 0.0.0.0/0 prefix-length-range /25-/32
term FILTER-PREFIX-LENGTH-2 then reject

term RPKI_REJECT_INVALID from community SYS11_ORIGIN_RPKI_INVALID
term RPKI_REJECT_INVALID then reject
```

No Dynamic prefix filter generator

Autogen

- Reads AS-SETs from file
- Generates XML „prefix-list“
- Applied via NETCONF
- Executed every night

router;type;lclpref;metric;enabled;import;export;passive;rpki;addr;email;ipv;peer-name;peer-ip;asn;as-set;md5;prefix-limit

router ; DECIX ; 110 ; 90 ; Y ; Y ; Y ; Y ; N ;;; 4 ; YAHOO ; 80.81.192.115 ; 10310 ; **AS-YAHOO** ;

term PEERING from prefix-list-filter **4-AS-YAHOO** orlonger; then accept

Autogen

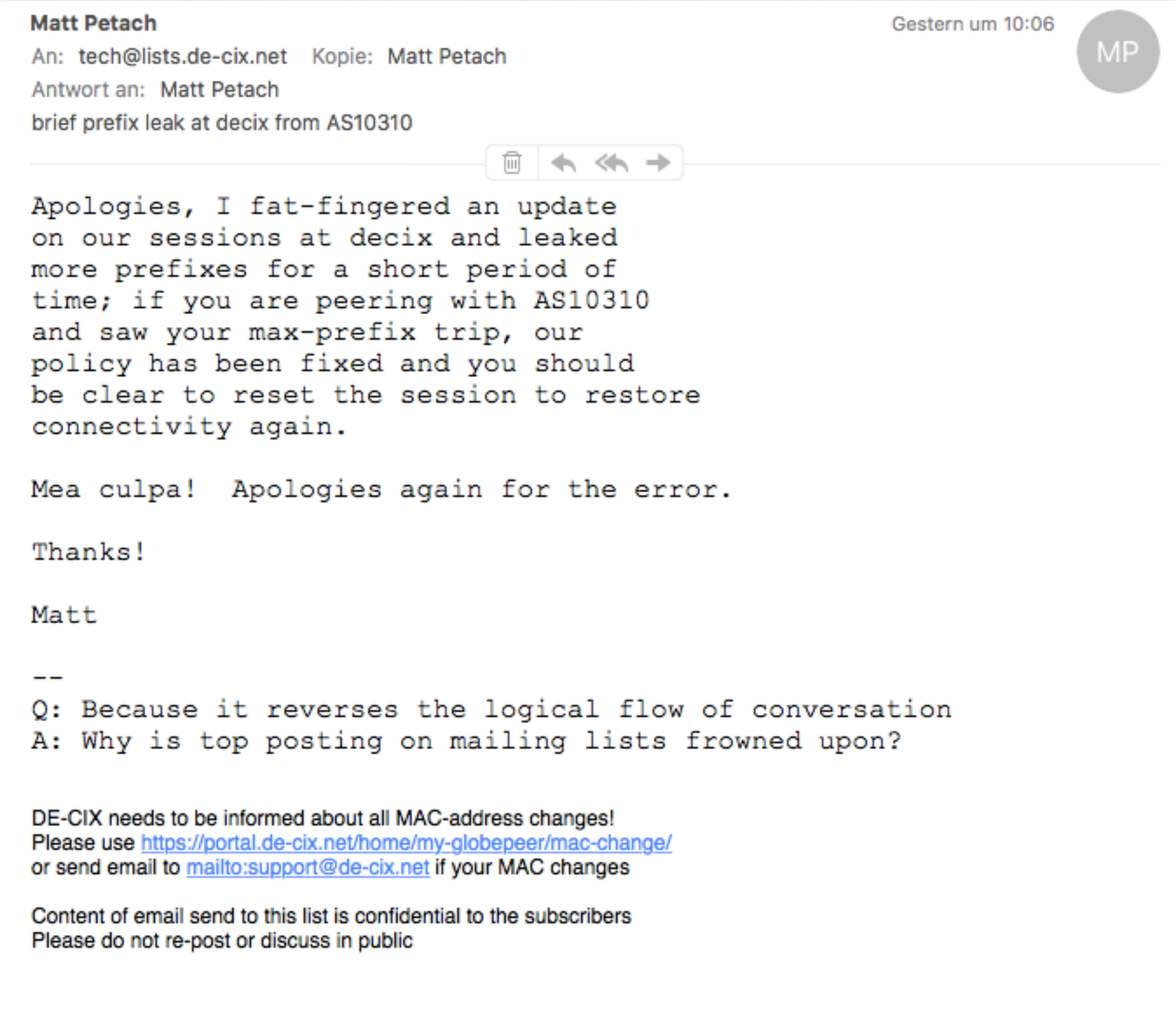
SysEleven
Hosting. Skaliert.

- Reads AS
- Generate
- Applied vi
- Executed

router;type;lclpref; -set;md5;prefix-limit

router ; DECIX ; 10 ; **AS-YAHOO** ;

term P accept



Matt Petach Gestern um 10:06
An: tech@lists.de-cix.net Kopie: Matt Petach
Antwort an: Matt Petach
brief prefix leak at decix from AS10310

```
Apologies, I fat-fingered an update
on our sessions at decix and leaked
more prefixes for a short period of
time; if you are peering with AS10310
and saw your max-prefix trip, our
policy has been fixed and you should
be clear to reset the session to restore
connectivity again.

Mea culpa!  Apologies again for the error.

Thanks!

Matt

--
Q: Because it reverses the logical flow of conversation
A: Why is top posting on mailing lists frowned upon?
```

DE-CIX needs to be informed about all MAC-address changes!
Please use https://portal.de-cix.net/home/my-globepeer/mac-change/
or send email to mailto:support@de-cix.net if your MAC changes

Content of email send to this list is confidential to the subscribers
Please do not re-post or discuss in public

Autogen / bgpq3

- Prefix-filter generator
- Extracts prefixes from route-objects
- Default IRR: RADB
- Supports Cisco & Juniper

https://github.com/snar/bgpq3

Autogen / aggregate

EVERYBODY LOVES AGGREGATION!

```
# apt-get install aggregate
```

https://github.com/job/aggregate6

Autogen

Generates Juniper XML:

```
echo "<configuration ><groups>
        <name>AUTOGEN-$ip_version</name><apply-flags><omit/></apply-flags>
          · <policy-options replace=\"replace\">"
             for a in $objects; do
          ·         echo "<prefix-list replace=\"replace\"><name>$ip_version-$a</name>"
                    /usr/bin/bgpq3 -h whois.syseleven.net $a | awk '{print $5}' | aggregate -q
                    while read line; do
          ·                echo "<prefix-list-item>$line</prefix-list-item>"
          ·         done
          ·         echo "</prefix-list>"
             done
echo "</policy-options></groups></configuration>"
```

Autogen / NETCONF

SysEleven
Hosting. Skaliert.

- Juniper NETCONF client
- edit_configuration.pl for JunOS 14+
- Reads xml-formatted configuration

```
/usr/bin/perl edit_configuration.pl -l $user -p $pass -m ssh $xmlfile $target:22
```

https://github.com/juniper/netconf-perl

Autogen / Challenges

- RPKI/max-prefix for peers with 10k+ prefixes
- Using ASN if no AS-SET exists
- Install own mirror instead of using RADB

whois.syseleven.net

SysEleven
Hosting. Skaliert.

- Running on IRRd v3.0.8
- RIPE, RADB, BBOI, LEVEL3, NTTCOM, ARIN, ALTDB
- Using downsized RIPE database

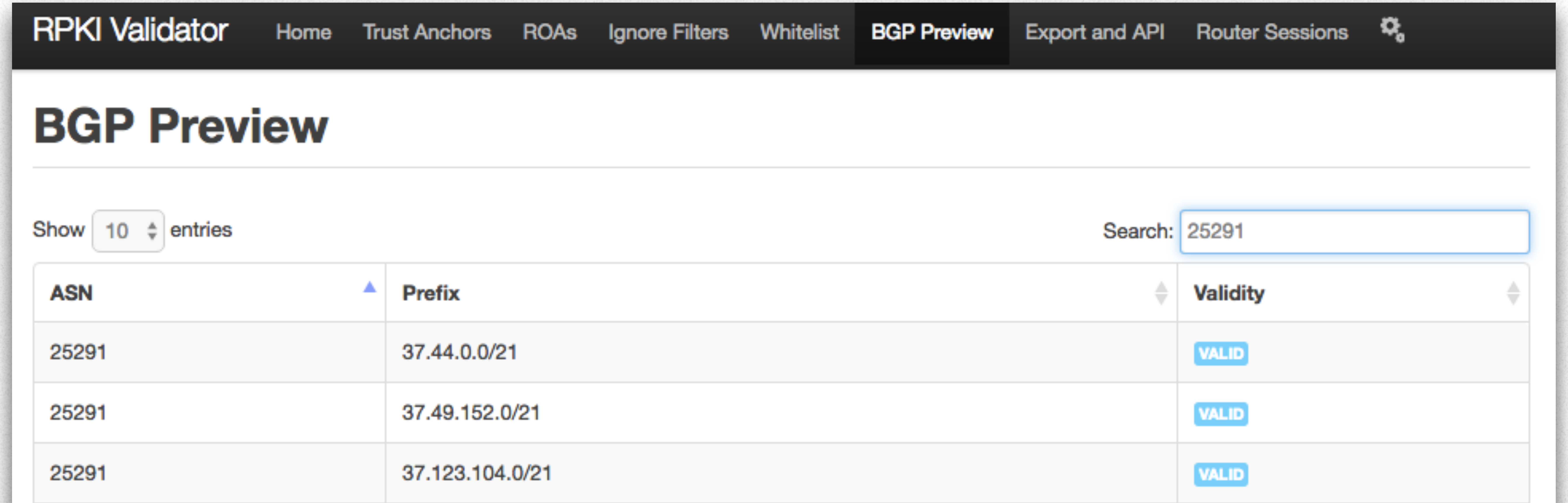
https://github.com/irrdnet/irrd

https://launchpad.net/~syseleven-platform/+archive/ubuntu/irrd

RPKI

- RIPE validator v2.23 used
- Please create ROAs via LIR Portal

https://github.com/RIPE-NCC/rpki-validator/



BGP Preview

Show 10 entries

Search: 25291

ASN	Prefix	Validity
25291	37.44.0.0/21	VALID
25291	37.49.152.0/21	VALID
25291	37.123.104.0/21	VALID

RPKI

Modes configured per peer:

- **`MODERATE`** Reject invalid announcements
- **`STRICT`** Accept only valid announcements

router;type;localpref;metric;enabled;import;export;passive;rpki;localaddr;email;ip-version;peer-name;peer-ip;asn;as-set;md5

router ; UPSTREAM ; 100 ; 100 ; Y ; Y ; Y ; N ; **{M,S}** ;;; 4 ; LEVEL3 ; 212.*.*.* ; 3356 ;;

RPKI

Configuration on JunOS:

```
tvoss@router> show configuration routing-options validation
group RPKI {
  session 151.252.**.** {
    refresh-time 300;
    hold-time 600;
    port 8282;
    local-address 37.123.**.**;
  }
  session 37.44.**.** {
    refresh-time 300;
    hold-time 600;
    port 8282;
    local-address 37.123.**.**;
  }
}
```

```
tvoss@router> show configuration policy-options policy-statement 4-DOWNSTREAM-IN
term RPKI-VALIDATION-VALID {
  from validation-database valid;
  then {
    validation-state valid;
    community add SYS11_ORIGIN_RPKI_VALID;
  }
}
term RPKI-VALIDATION-INVALID {
  from validation-database invalid;
  then {
    validation-state invalid;
    community add SYS11_ORIGIN_RPKI_INVALID;
  }
}

tvoss@router> show configuration policy-options policy-statement 4-CUSTOMER-IN
term RPKI_REJECT_INVALID {
  from community SYS11_ORIGIN_RPKI_INVALID;
  then reject;
}
```

RPKI / Challenges

- 10k+ invalid routes rejected
- Biggest polluter: a certain Tier1
- Disputable possibility of censorship

```
tvoss@router> show route receive-protocol bgp CERTAIN-TIER1 table inet.0 hidden | count
Count: 3765 lines*
tvoss@router> show route receive-protocol bgp TELIA-CARRIER table inet.0 hidden | count
Count: 0 lines*
```

* 1 line subtracted for header information

RPKI / Challenges

- If validator dies, invalid announcements accepted
- Setup a second validator

```
tvoss@router> show validation session
Session            State    Flaps    Uptime             #IPv4/IPv6 records
37.44.**.**        Up       0        1w3d 05:47:59      24999/3591
151.252.**.**      Up       0        1w3d 06:04:23      24999/3591
```

It’s not only about filtering

Denial of Service

SysEleven
Hosting. Skaliert.

SysEleven’s challenge:

- DDoS smaller than 100 Gbps
- 99% volumetric attacks
- 99% stupid attacks

Denial of Service

SysEleven's approach:

- Increased upstream capacity
- Moved all ports into LAGs
- Installed FastNetMon
- Enabled FlowSpec

Denial of Service / FastNetMon

SysEleven
Hosting. Skaliert.

- DDoS attack detection
- User-defined thresholds
- Collects NetFlow, sFlow, IPFIX data
- Support for Graphite, InfluxDB, ExaBGP

https://github.com/pavel-odintsov/fastnetmon

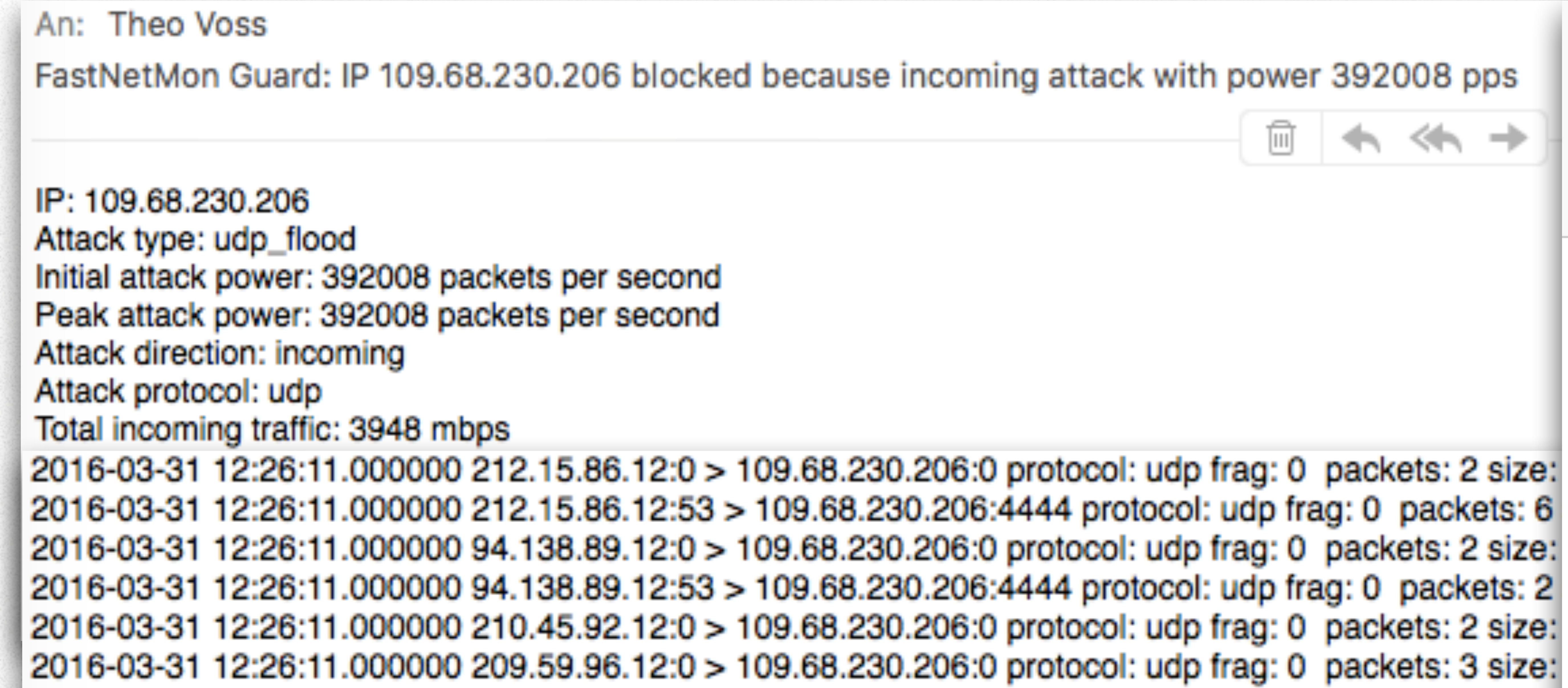
Detection is good
Mitigation is better

Denial of Service / FlowSpec

- FlowSpec (RFC5575) enabled
- Filters propagated by BGP
- Rate-Limit possible
- Upstream sessions are FlowSpec enabled
- Communities for advertising/exporting

Denial of Service / Attack

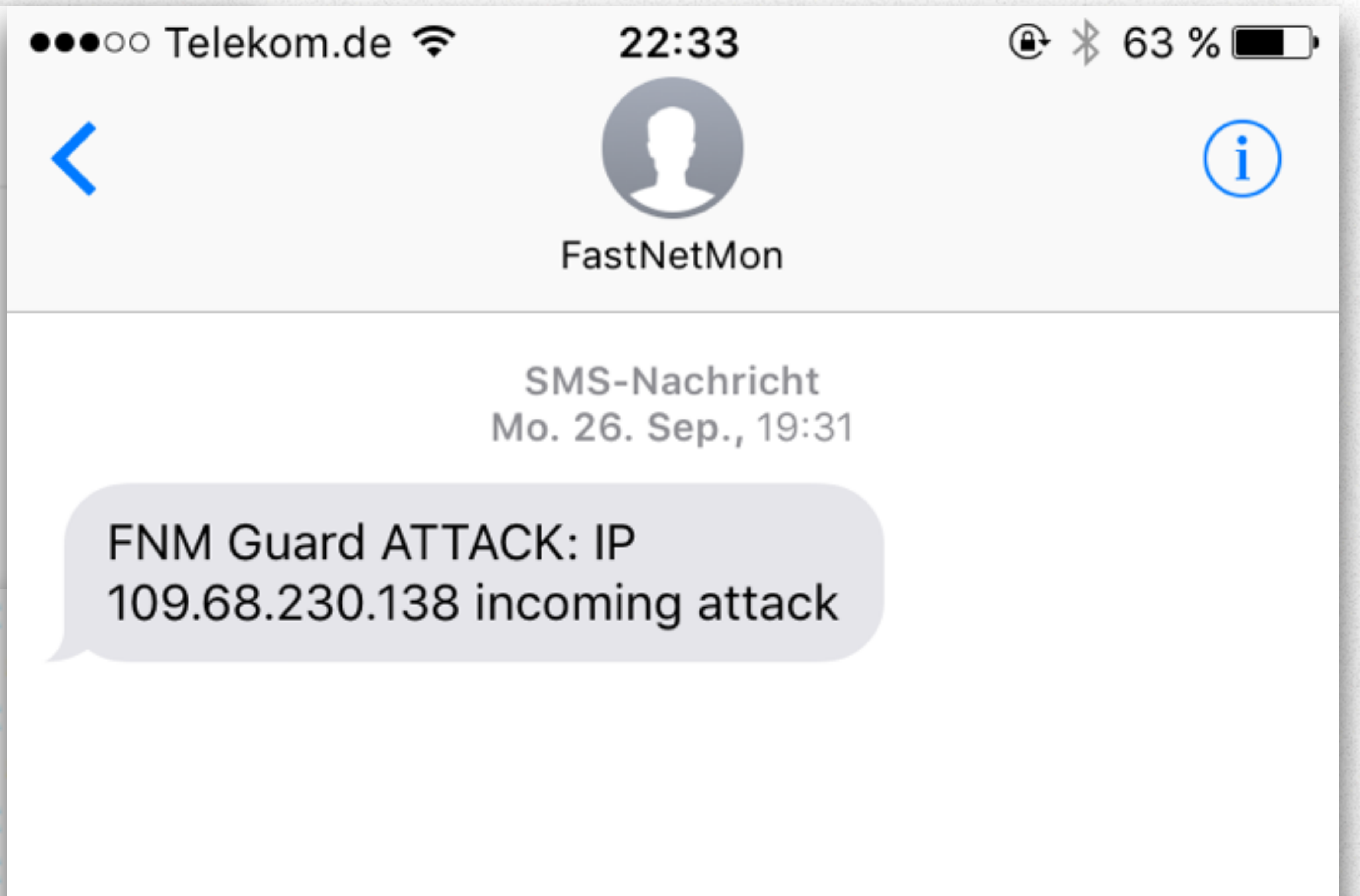
- Detection and mitigation in less then 2 minutes
- Script triggered: /usr/… /notify_about_attack.{sh,py}
- SMS via 3rd-party API to NOC engineer on duty



An: Theo Voss

FastNetMon Guard: IP 109.68.230.206 blocked because incoming attack with power 392008 pps

```
IP: 109.68.230.206
Attack type: udp_flood
Initial attack power: 392008 packets per second
Peak attack power: 392008 packets per second
Attack direction: incoming
Attack protocol: udp
Total incoming traffic: 3948 mbps
2016-03-31 12:26:11.000000 212.15.86.12:0 > 109.68.230.206:0 protocol: udp frag: 0  packets: 2 size:
2016-03-31 12:26:11.000000 212.15.86.12:53 > 109.68.230.206:4444 protocol: udp frag: 0  packets: 6
2016-03-31 12:26:11.000000 94.138.89.12:0 > 109.68.230.206:0 protocol: udp frag: 0  packets: 2 size:
2016-03-31 12:26:11.000000 94.138.89.12:53 > 109.68.230.206:4444 protocol: udp frag: 0  packets: 2
2016-03-31 12:26:11.000000 210.45.92.12:0 > 109.68.230.206:0 protocol: udp frag: 0  packets: 2 size:
2016-03-31 12:26:11.000000 209.59.96.12:0 > 109.68.230.206:0 protocol: udp frag: 0  packets: 3 size:
```

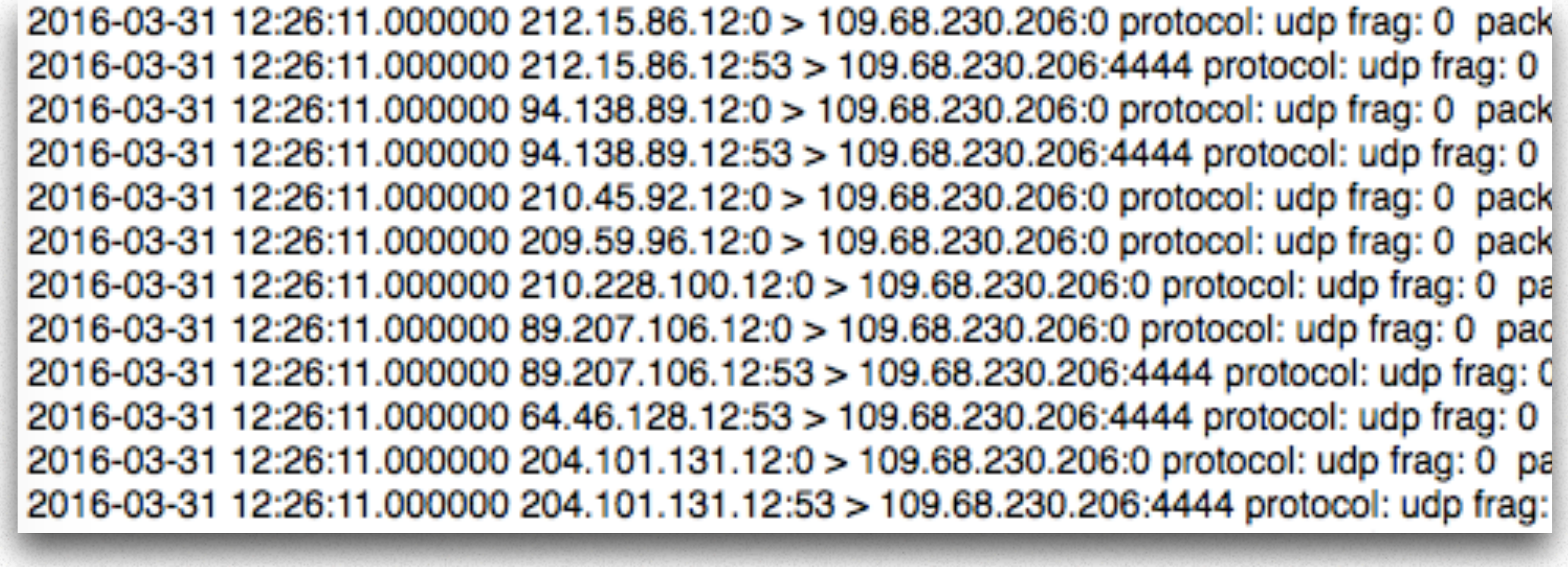


SMS-Nachricht
Mo. 26. Sep., 19:31

FNM Guard ATTACK: IP 109.68.230.138 incoming attack

Denial of Service / Attack

Information from FNM capture used:



```
2016-03-31 12:26:11.000000 212.15.86.12:0 > 109.68.230.206:0 protocol: udp frag: 0  pack
2016-03-31 12:26:11.000000 212.15.86.12:53 > 109.68.230.206:4444 protocol: udp frag: 0
2016-03-31 12:26:11.000000 94.138.89.12:0 > 109.68.230.206:0 protocol: udp frag: 0  pack
2016-03-31 12:26:11.000000 94.138.89.12:53 > 109.68.230.206:4444 protocol: udp frag: 0
2016-03-31 12:26:11.000000 210.45.92.12:0 > 109.68.230.206:0 protocol: udp frag: 0  pack
2016-03-31 12:26:11.000000 209.59.96.12:0 > 109.68.230.206:0 protocol: udp frag: 0  pack
2016-03-31 12:26:11.000000 210.228.100.12:0 > 109.68.230.206:0 protocol: udp frag: 0  pa
2016-03-31 12:26:11.000000 89.207.106.12:0 > 109.68.230.206:0 protocol: udp frag: 0  pac
2016-03-31 12:26:11.000000 89.207.106.12:53 > 109.68.230.206:4444 protocol: udp frag: 0
2016-03-31 12:26:11.000000 64.46.128.12:53 > 109.68.230.206:4444 protocol: udp frag: 0
2016-03-31 12:26:11.000000 204.101.131.12:0 > 109.68.230.206:0 protocol: udp frag: 0  pa
2016-03-31 12:26:11.000000 204.101.131.12:53 > 109.68.230.206:4444 protocol: udp frag:
```

```
tvoss@router# show | compare
[edit routing-options flow]
+    route 109.68.230.206/32 {
+        match {
+            destination 109.68.230.206/32;
+            protocol udp;
+            port [ 0 4444 ];
+        }
+        then {
+            community ANNOUNCE_UPSTREAM;
+            discard;
+    }
```

Denial of Service / Attack

- FlowRoute propagated internally and upstream
- More-specific route announced upstream

```
inetflow.0: 1 destinations, 1 routes (1 active, 0 holddown, 0 hidden)
109.68.230.206,*,proto=17,port=0,=4444/term:1 (1 entry, 1 announced)
        *BGP    Preference: 170/-101
                Next hop type: Fictitious
                Announcement bits (1): 0-Flow
                Communities: traffic-rate:0:0
                Accepted
                Validation state: Accept, Originator: 37.44.7.60
                Via: 109.68.230.0/24, Active
```

Denial of Service / FastNetMon

- FastNetMon v1.13 can do blackholing
- Don’t try to use FlowSpec, wait for v2.0

- GoBGP used for FlowSpec in v2.0
- Ratelimit/discard in case of attack

Summary

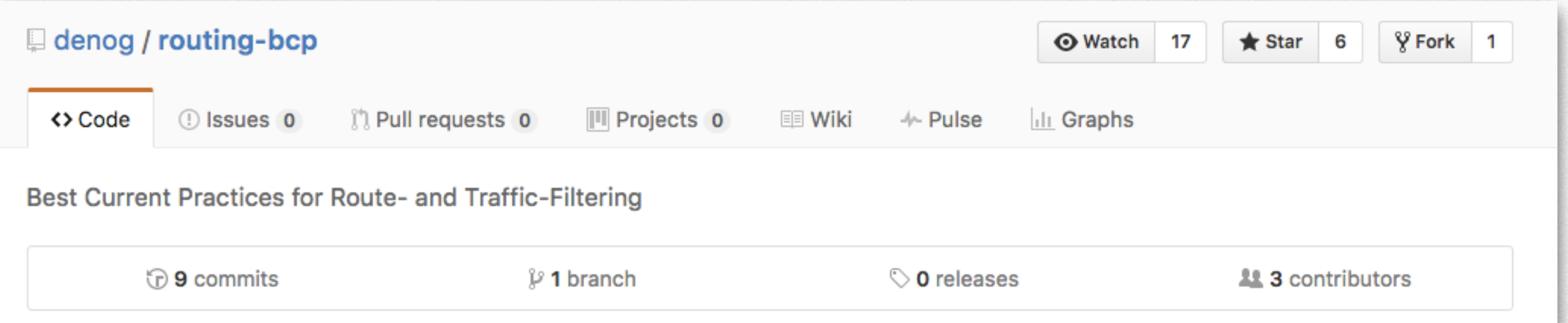
Self-Made-Filters + Open-Source-Tools

- Budget friendly
- Less incidents
- Does the job! :-)

Routing BCP

- Everybody invited to submit his routing policies
- Volunteers wanted to compile draft BCP

https://github.com/denog/routing-bcp



Best Current Practices for Route- and Traffic-Filtering

9 commits | 1 branch | 0 releases | 3 contributors