

Routing without collateral damage

AfriNIC #15 – Nov 23, 2011



Your Speaker Today....



Fredy Künzler

CTO & Founder

kuenzler at init7.net

www.init7.net

www.blogg.ch

www.bgp-and-beyond.com

AS13030

Twitter: @init7

Init Seven AG

Elias-Canetti-Strasse 7

CH-8050 Zürich

Switzerland



AGENDA

A Init7 / AS13030

B When using BGP - think global!

C Small steps to a smaller (cleaner) BGP table



A Init7 / AS13030

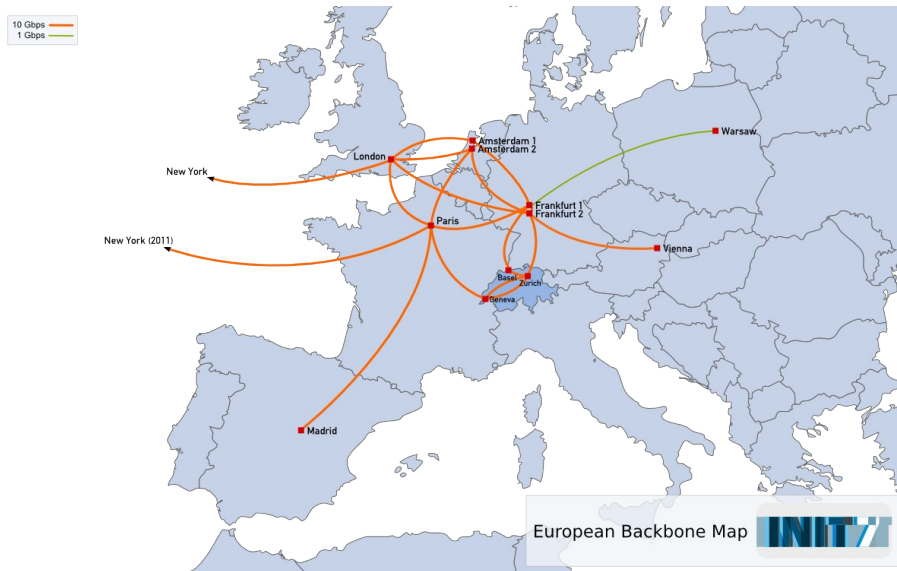
Init7 / AS13030

Who we are

- **Carrier / Internet Service Provider, based in Zurich, Switzerland**
- **Privately owned company**
- **Own international fully dual-stacked v4 and v6 backbone (AS13030), 10gig or multiple 10gig enabled**
- **Connected to 20+ internet exchanges and close to 1'000 BGP peers/ customers**



Init7 operates an international backbone



Facts

- Public Peerings at ~20 Internet Exchanges
- Open Peering policy



DISCLAIMER

These slides show experience examples of the Init7 / AS13030 backbone over various years. They may work or may not work for you. Please use the methods described with care and at your own risk. Init7 or the author cannot be held responsible for any damage occurred by using the methods described here.



B When using BGP - think global!

When using BGP...

...think global! #1

- We learned BGP4 routing, made lab tests, we even operate it in a productive environment. But most operators look at their gear only from their local perspective
- See and be seen is a different story. From the local perspective, everything looks good – but would I win a (network) beauty contest?



When using BGP...

...think global! #2

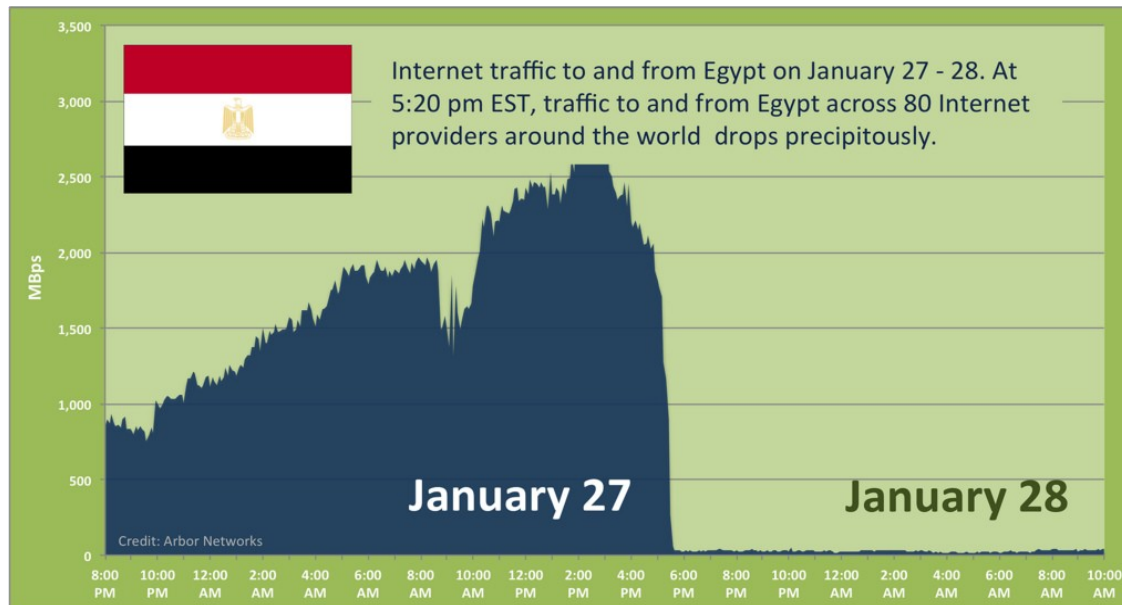
- Not even the smallest BGP4 hick-up goes away undetected. Every reboot of a BGP router, every prefix change, every new transit relation is immediately notified and noticed – worldwide



When using BGP...

...think global! #3

- Earlier this year, when several Arabic speaking countries cut themselves off from the internet, internet monitoring companies [Arbor Networks | Renesys] tried to beat each other in sending out press releases quicker – the fastest would be quoted worldwide in newspapers and television, because global recognition means a lot of free marketing!



Graphic by Arbor Networks



When using BGP...

...think global! #4

- Look into the mirror! Is your network neat and tidy?
- “The mirror” of the BGP4 table is the well known CIDR report, distributed every week in the AFNOG mailing list



When using BGP...

...think global! #5

- The CIDR report, which shows how much smaller the global BGP table could be if everybody would aggregate neatly:

Aggregation Summary

The algorithm used in this report proposes aggregation only when there is a precise match using AS path so as to preserve traffic transit policies. Aggregation is also proposed across non-advertised address space ('holes').

--- 21Nov11 ---

ASnum	NetsNow	NetsAggr	NetGain	% Gain	Description
Table	384785	225215	159570	41.5%	All Ases

Source: <http://www.cidr-report.org/as2.0/#Gains>



When using BGP...

...think global! #6

Please aggregate your prefixes! It's good for our industry and community!

- Less memory usage
 - Faster BGP conversion / less CPU cycles
 - Longer life of equipment
-

Are network operators lazy? Do network engineers know what they are doing? Are they actually caring?



There are smarter ways of traffic engineering than de-aggregation! → <http://goo.gl/A1Nu3>



C Small steps to a smaller (cleaner) BGP table

Small steps to a smaller BGP table #1

What networks often do :

- More-Specific propagation  Acceptable, when smartly & decently executed
- Massive de-aggregation  Pollution of the global BGP table! → more than 40% of the table size is rubbish...



Small steps to a smaller BGP table #2

```
*>i1.45.0.0/16      213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.16.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.20.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.24.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.28.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.32.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.36.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.40.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.44.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.48.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.52.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.56.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.60.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.64.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.68.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.72.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.76.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.80.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.84.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.88.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.92.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.96.0/22     213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.100.0/22    213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.104.0/22    213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.108.0/22    213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.112.0/22    213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.116.0/22    213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.120.0/22    213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.124.0/22    213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.128.0/22    213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.132.0/22    213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
*>i1.45.204.0/22    213.144.128.179      1      50      0 3549 1239 4837 4808 45083 i
...
```

We don't need
hundreds or
thousands of
identical paths...

... it's pollution!



Small steps to a smaller BGP table #3

Possible reasons for BGP de-aggregation:

- 'No-export' community not set
- 'neighbor x.x.x.x send-community' not set
- lack of knowledge
- “Best [worst] practice consulting” out in the wild – who actively promotes de-aggregation?!



Small steps to a smaller BGP table #4

... evangelize aggregation!



If everybody would convince customers / fellow network engineers / peers to get rid of the de-aggregated prefixes, the whole community would gain!



If you have any questions, please contact me...



Fredy Künzler

CTO & Founder

kuenzler at init7.net

www.init7.net

www.blogg.ch

www.bgp-and-beyond.com

AS13030

Twitter: @init7

Init Seven AG

Elias-Canetti-Strasse 7

CH-8050 Zürich

Switzerland

