

Impressions

An overview of the global IPv6 routing table

Gert Döring, SpaceNet AG, Munich

September 3, 2003

RIPE 46, Amsterdam

Overview

- what's “the global IPv6 routing table”
- numbers & pictures
- observations & trends
- conclusions & recommendations
- references

Slides online at: <http://www.space.net/~gert/RIPE/R46-v6-table/>

What's "the global IPv6 routing table"?

- it's what you get when you hook up to "the IPv6 routing world" using BGP4++
- a mixture of 6bone and RIR IPv6 addresses and networks
- some ASes announce 6bone only, some RIR only, a few both
- Network structure different in US vs. EU vs. AP region
- unlike IPv4: transit agreements don't reflect business relations. Transit usually for free (in US/EU).
- unlike IPv4: most ASes do not filter anything (is changing)
- unlike IPv4: still heavily tunnel based - BGP topology does not always reflect physical topology (is improving)

Numbers - Prefixes

As of 2003/09/03: 485 prefixes in total (2003/05/13: 526)

/n	global	RIR space	6bone	6to4	(2003/05/13)
/16	1	0	0	1	(1 0 0)
/24	45	0	45	0	(47 0 47)
/28	42	0	42	0	(48 0 48)
/32	266	238	28	0	(225 195 30)
/33	4	3	0	1	(2 1 0)
/34	1	0	0	1	(2 0 1)
/35	45	45	0	0	(53 53 0)
/36	3	2	0	1	(2 1 0)
/39	1	1	0	0	(0 0 0)
/40	4	4	0	0	(6 5 1)
/41	0	0	0	0	(5 5 0)
/42-/45	3	3	0	0	(3 3 0)
/48	66	36	26	4	(81 43 35)
/52-/60	0	0	0	0	(3 0 3)
/64	4	2	2	0	(46 39 7)
/65-/128	0	0	0	0	(2 2 0)

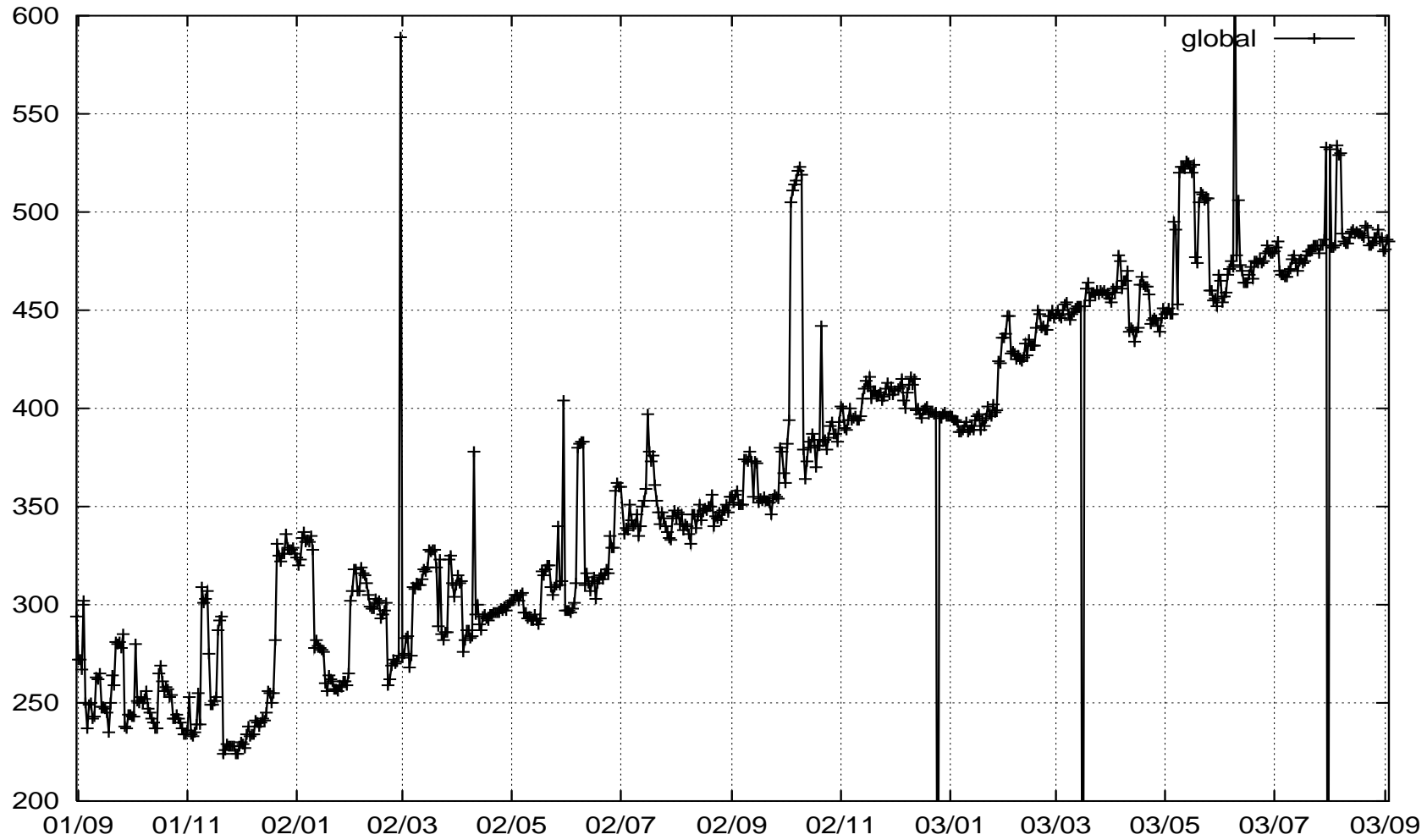
6to4 - 2002::/16

- 6to4 prefix 2002::/16 anycast prefix - *multiple* origin ASes

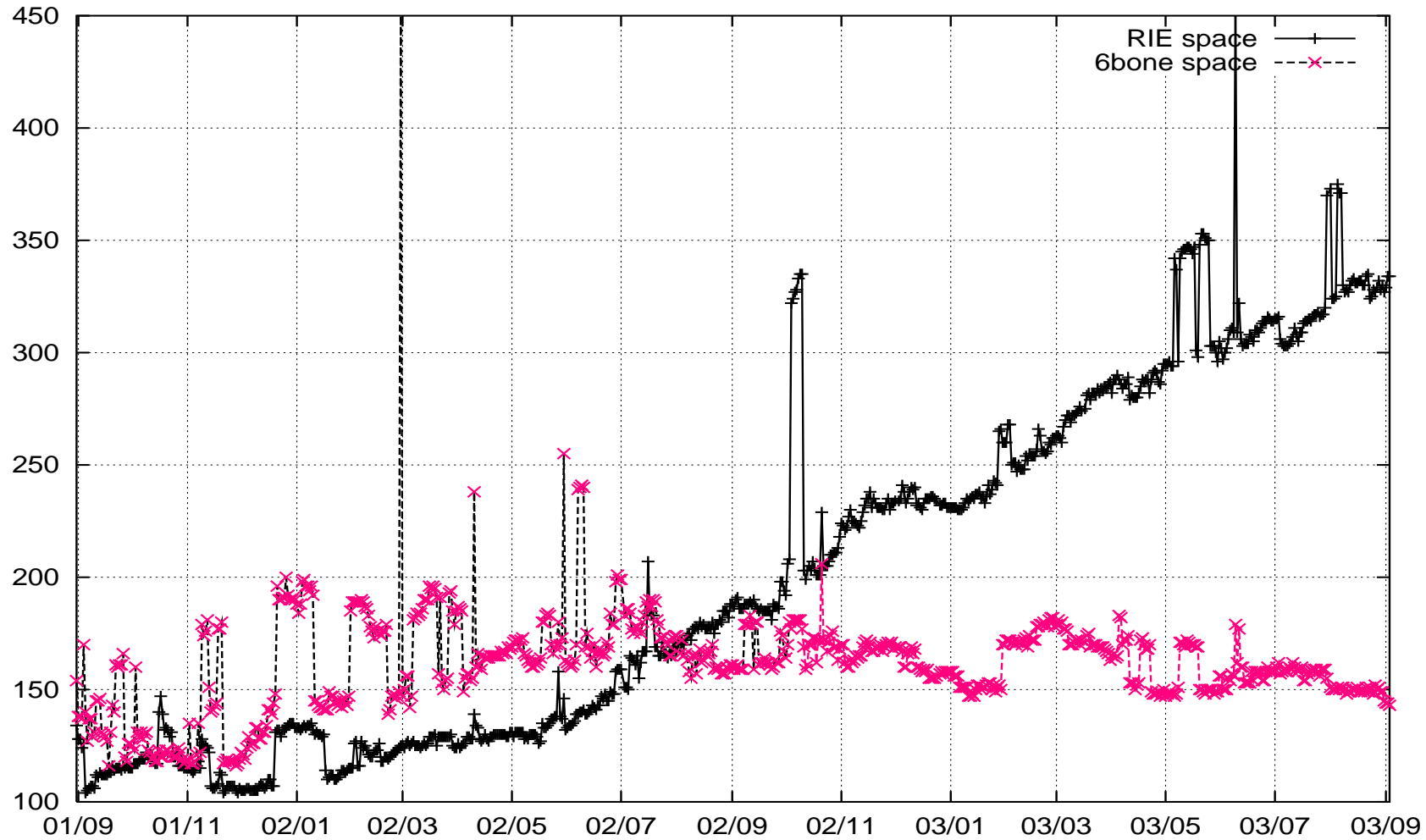
Network	Next Hop	Path
* 2002::/16	2001:608:0:3::D	5430 3549 20965 2603 1741 i
*	3FFE:8150:0:1::17	9044 559 i
*	2001:608:0:3::7	1930 i
*	2001:7F8:2:8001::2	1752 i
*>i	2001:650:F807::20BB:1	8379 i
*	2001:948:0:F00F::1	2603 1741 i
*	3FFE:C00:8023:19::1	109 i
*	2001:608:0:3::9	3320 1752 i

- anycast relay address 192.88.99.1/24, RFC3068
- some research on non-publically visible 6to4 relays by David Malone (dwmalone@maths.tcd.ie): approximately 33 relays found. Good start, but more relays would be useful.
- some more-specific pfxs seen (prohibited by RFC3056 5.2.3!)

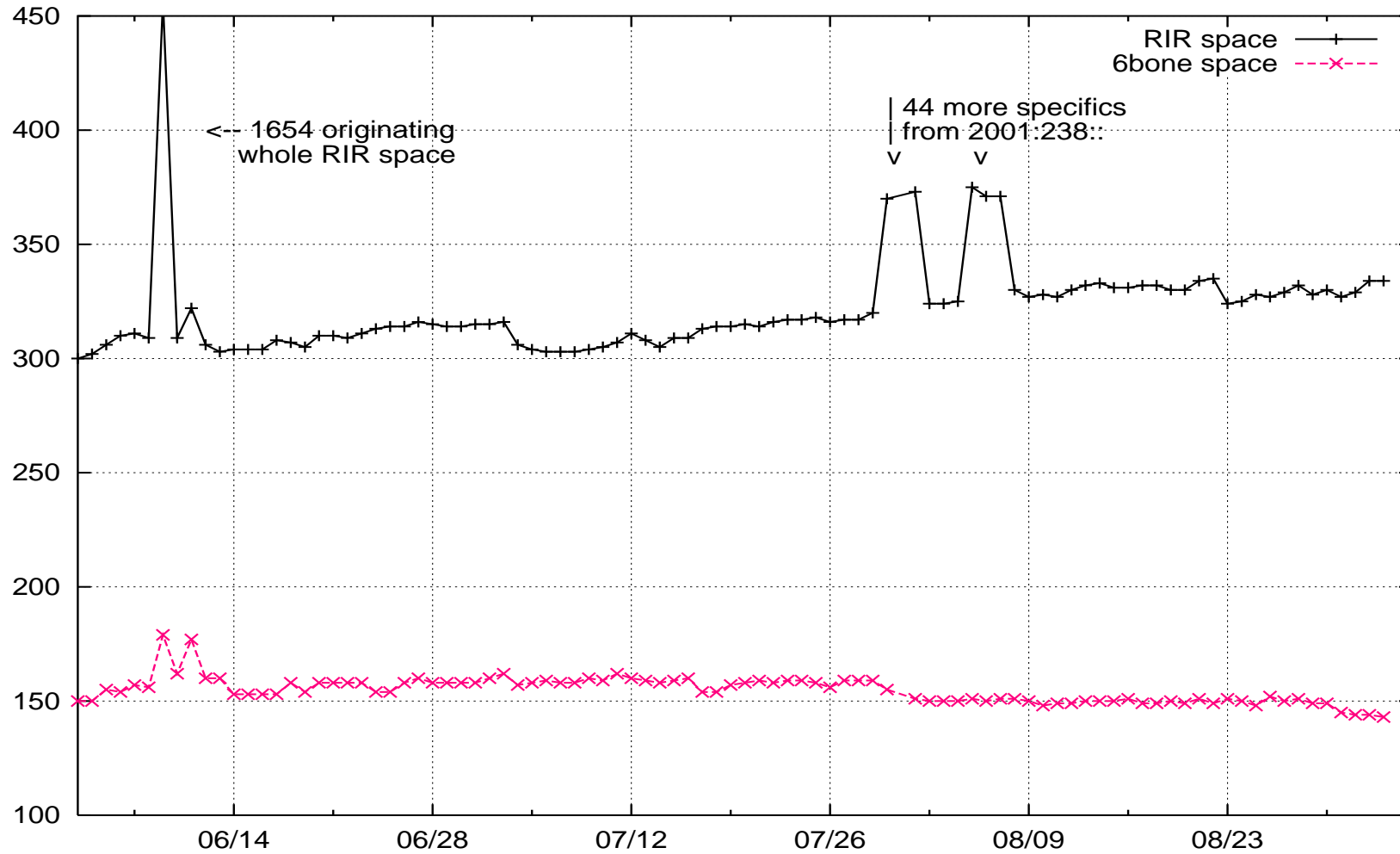
Graphics: Total Prefixes - 24 months



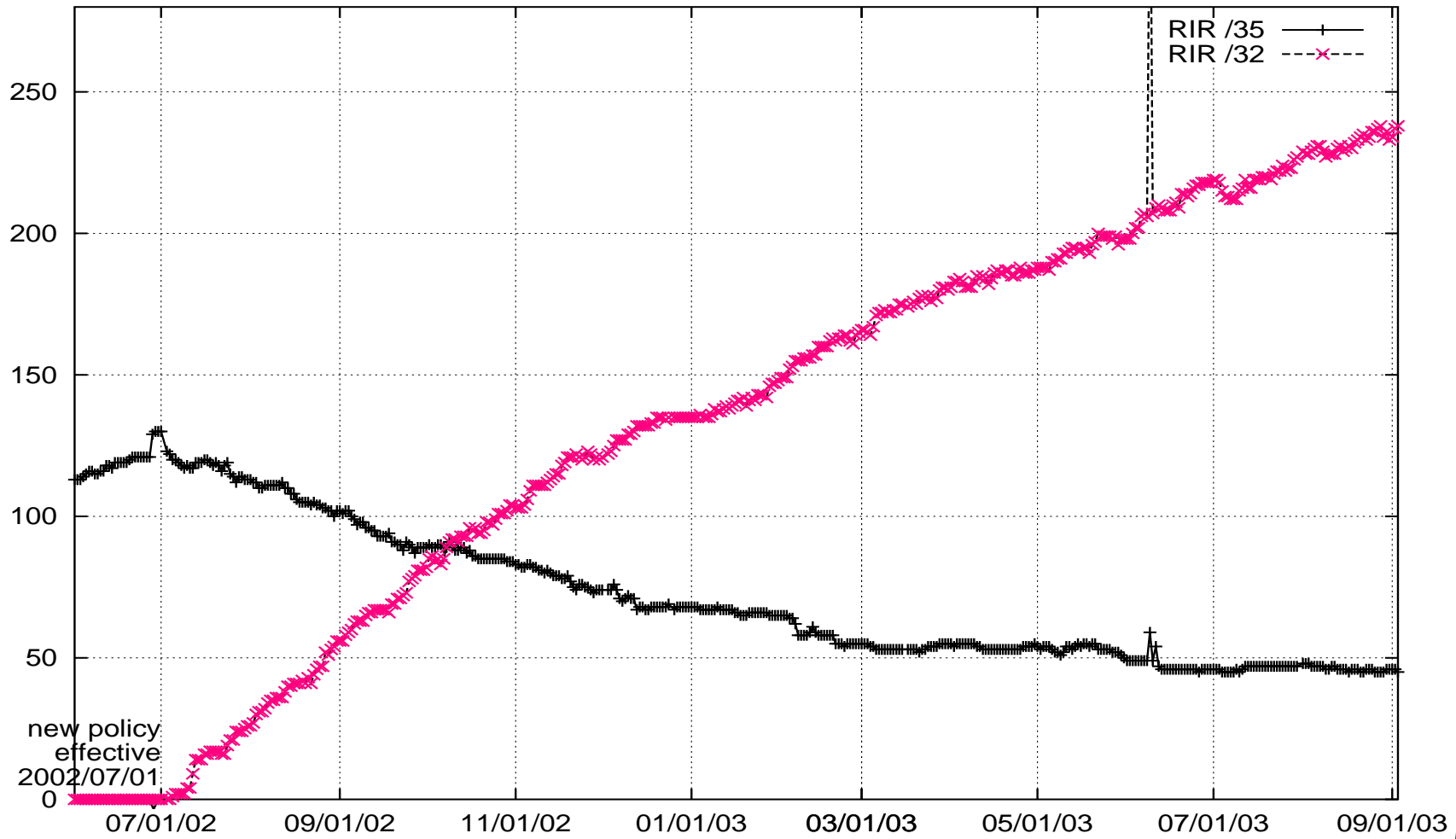
Graphics: RIR vs. 6Bone Prefixes - 24 months



Graphics: RIR vs. 6Bone Prefixes - 3 months



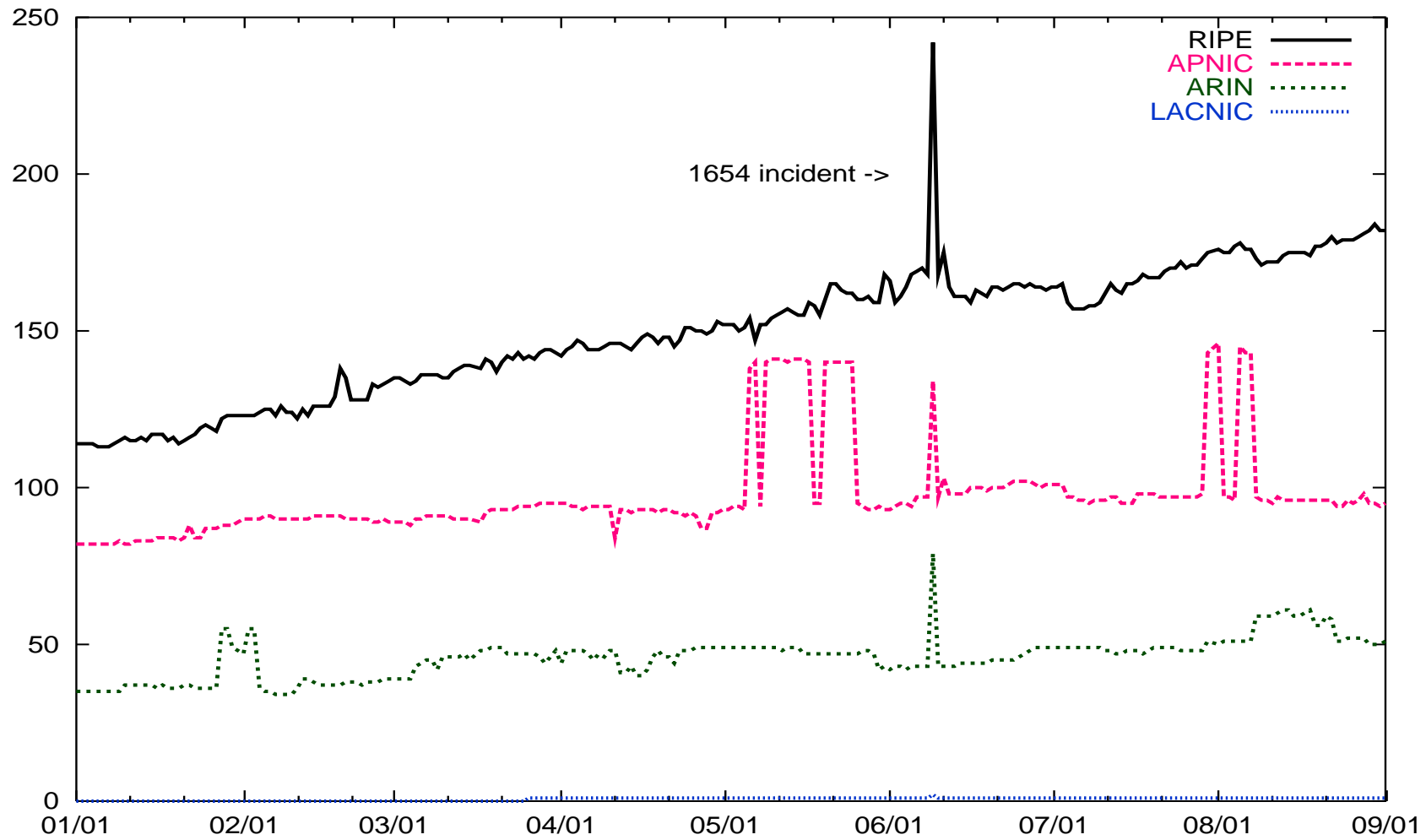
Graphics: RIR /35s vs. /32s



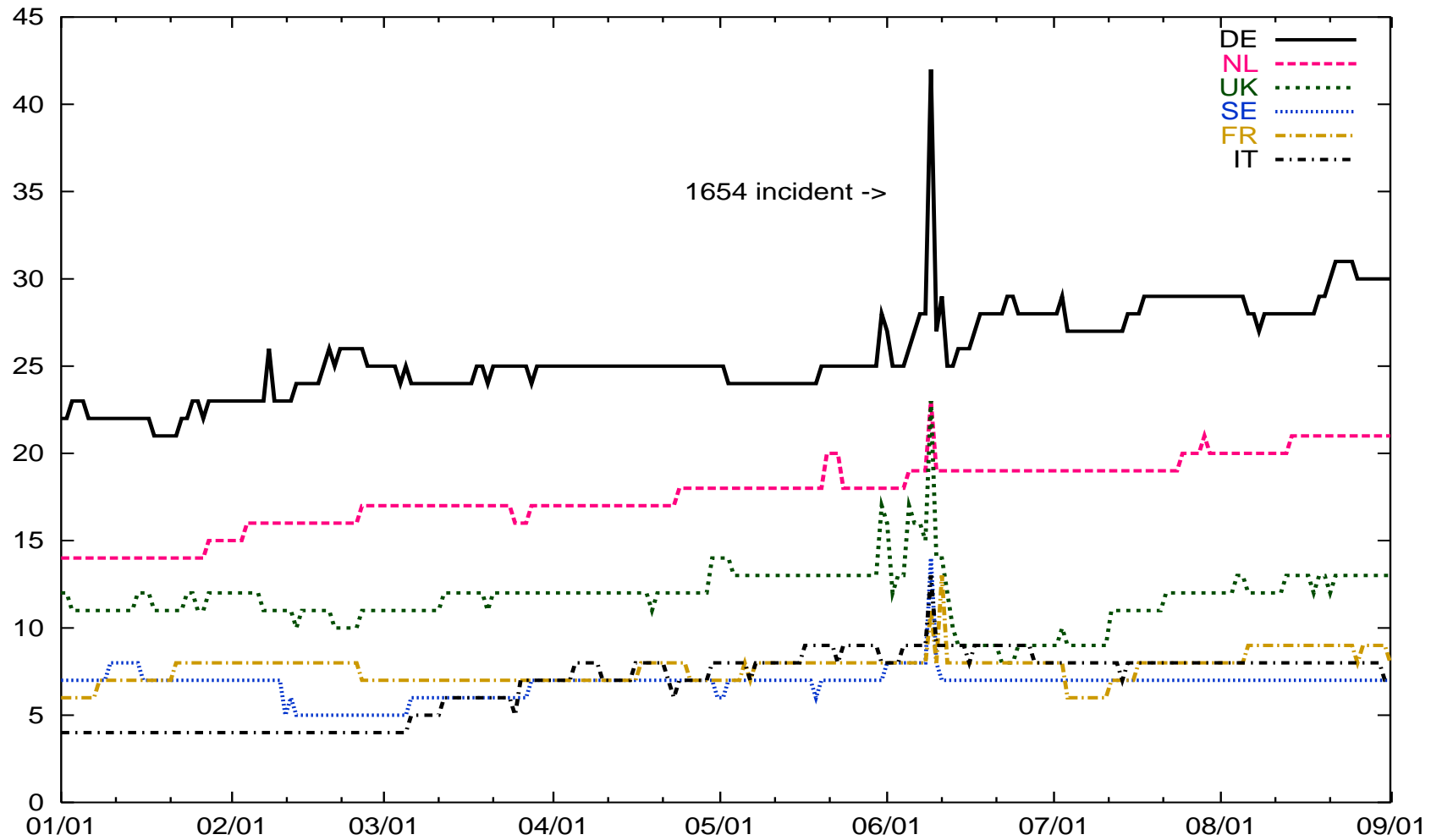
Numbers: RIRs, ASes, ...

- 429 LIR blocks out of 2001:: $/16$ allocated by RIRs:
ARIN 76(63), APNIC 112(103), RIPE 238(193), LACNIC 3(2)
as of 2003/09/03 (2003/05/13)
- plus some IXP and other microallocations
- IPv6 assignment to K.root (2001:07FD:: $/32$, not visible) and
M.root (2001:0DC3:: $/32$, announced since 09/01)
- 271(235) allocations visible
- allocations take up 283(248) routes: 45(53) $/35$ s, 238(195) $/32$ s
16(18) allocations visible as $/32$ and $/35$
- total unique ASes in the IPv6 BGP table: 371(335)
- new IPv6 block for RIPE since 2003/07/30: **2001:1600:: $/23$**
- first RIR $/31$ allocation ever on 2003/09/02 (NL-LIBERTEL)

Graphics: prefixes by RIR region



Graphics: prefixes by country (RIPE)

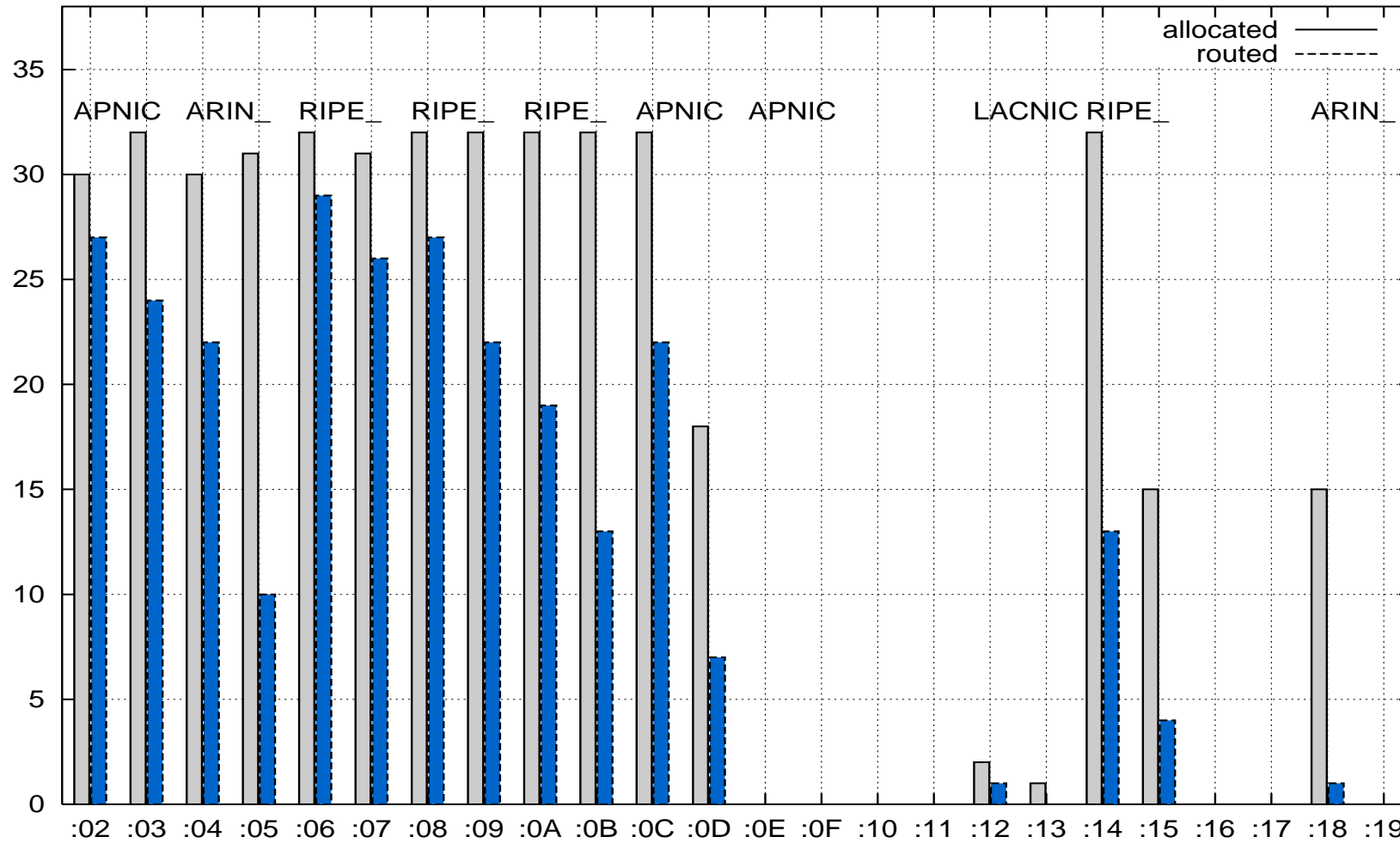


Numbers: new /32s and /35s since 2003/05/13

dropped due to not-so-much usefulness

(but tell me if you want it back)

Graphics: Allocated vs. Routed



Interesting Observations (1) - AS 1654 incident

Network	Next Hop	Path
*> 2001:210::/35	2001:608:0:3::9 3FFE:8150:0:1::17	3320 9112 2847 1654 i 9044 513 9112 2847 1654 i
*> 2001:230::/32	2001:608:0:3::9	3320 9112 2847 1654 i
*> 2001:280::/32	2001:608:0:3::9	3320 9112 2847 1654 i
*> 2001:2D8::/32	2001:608:0:3::9	3320 9112 2847 1654 i
*> 2001:300::/32	2001:608:0:3::9	3320 9112 2847 1654 i
...		
*> 2001:538::/32	2001:608:0:3::9	3320 9112 2847 1654 i
*> 2001:540::/32	2001:608:0:3::9	3320 9112 2847 1654 i
*> 2001:548::/32	2001:608:0:3::9	3320 9112 2847 1654 i
*> 2001:550::/32	2001:608:0:3::9	3320 9112 2847 1654 i
*> 2001:558::/32	2001:608:0:3::9	3320 9112 2847 1654 i
*> 2001:560::/32	2001:608:0:3::9	3320 9112 2847 1654 i
*> 2001:568::/32	2001:608:0:3::9	3320 9112 2847 1654 i

1654 announced 'every allocated prefix' - quite a lot have not been in the table before or afterwards (178 pfxs seen at 5539).

Actually *two independent networks* used AS1654.

Interesting Observations (2) - Ghost Busting

```

Network      Path
*> 2001:650::/32 3561 i
*           3274 790 790 3561 i
*           5430 13285 786 1752 3561 i
*           4555 6830 5511 3561 i
*           1752 3561 i
*           9044 10566 1930 3561 i
*           1930 20965 11537 6939 3257 3561 i
*           6939 3257 3561 i

*>i2001:650::/35 3257 6762 3263 6939 145 3561 ?
*           3274 790 790 6830 4589 3257 6762 3263 6939 145 3561 ?
*           5430 3549 2500 2497 3257 6762 3263 6939 145 3561 ?
*           4555 6830 4589 3257 6762 3263 6939 145 3561 ?
*           1752 6830 4589 3257 6762 3263 6939 145 3561 ?
*           109 6175 2497 3257 6762 3263 6939 145 3561 ?
*           9044 559 3303 15717 3257 6762 3263 6939 145 3561 ?
*           1930 20965 11537 3425 293 6175 2497
*           3257 6762 3263 6939 145 3561 ?
*           3320 6830 4589 3257 6762 3263 6939 145 3561 ?

```

Ghosts = BGP withdrawal bug, caused by old and buggy software.

Long paths stay *mostly unchanged* in the table for weeks.

Don't confuse with BGP count-to-infinity (= paths change quickly).

Interesting Observations (3) - Count-to-Infinity

Start:

```
* 2001:638::/35      2001:7F8::2A8:0:1      680 i
*                   3FFE:C00:8023:19::1    109 6939 3257 680 i
*                   2001:470:1FFF:2::      6939 3257 680 i
*                   3FFE:8150:0:1::17      9044 559 6680 680 i
```

After withdrawal: (snapshots between 08:30 and 08:47)

```
13129 20646 8560 680 i
3549 6939 4725 1752 6830 4589 680 i
1752 3320 9112 8664 13110 6939 3257 680 i
3320 9112 8664 13110 6939 14277 4725 1752 6830 4589 680 i
...
3561 2497 2500 5511 3320 9112 8664 13110 6939 14277 4725 1752 6830 4589 680 i
...
8472 8903 16091 513 9044 5623 5609 15589 3320 9112 8664 13110 6939 +
14277 4725 1752 6830 4589 680 i
...
9044 5623 5609 22 11537 145 12199 237 3748 17832 7623 1237 17579 3425 +
293 3320 9112 8664 13110 6939 14277 4725 1752 6830 4589 680 i
% Network not in table
...
9044 5623 5609 33 25396 25396 25396 25396 25396 25396 25396 15703 3265
8954 790 790 790 790 790 790 209 8002 2516 7660 2915 2713 2042 4774 +
2497 2500 5511 2200 20965 11537 145 12199 237 3748 17832 7623 1237 +
17579 3425 293 3320 9112 8664 13110 6939 14277 4725 1752 6830 4589 680 i
```

Path lengths of BGP path buildup after withdrawal hints at high percentage of ASes giving transit to (unsuspecting) third parties.

Interesting Observations (4) - more-specific leaks

Network	Next Hop	Path
* 2001:238::/32	2001:478:FFFF::1	4555 6939 3257 17419 i
* i	2001:7F8::CB9:0:1	3257 17419 i
*	2001:470:1FFF:2::	6939 3257 17419 i
*	2001:7F8:2:8001::2	1752 2914 17419 i
*>i	2001:608:0:3::D	5430 3549 17419 i
...		
*> 2001:238::/64	2001:470:1FFF:2::	6939 6939 17715 17419 i
*> 2001:238:0:24::/64	2001:470:1FFF:2::	6939 6939 17715 17419 i
* 2001:238:100::/41	2001:478:FFFF::1	4555 6939 6939 17715 17419 i
*>	2001:470:1FFF:2::	6939 6939 17715 17419 i
*> 2001:238:200::/64	2001:470:1FFF:2::	6939 6939 17715 17419 i
* 2001:238:200::/41	2001:478:FFFF::1	4555 6939 6939 17715 17419 i
*>	2001:470:1FFF:2::	6939 6939 17715 17419 i
*> 2001:238:600::/64	2001:470:1FFF:2::	6939 6939 17715 17419 i
*> 2001:238:800::/64	2001:470:1FFF:2::	6939 6939 17715 17419 i
* 2001:238:882::/48	2001:478:FFFF::1	4555 6939 6939 17715 17419 i
*>	2001:470:1FFF:2::	6939 6939 17715 17419 i
*> 2001:238:900::/64	2001:470:1FFF:2::	6939 6939 17715 17419 i
*> 2001:238:A00::/64	2001:470:1FFF:2::	6939 6939 17715 17419 i

17419/17715 repeatedly leaking more-specifics, likely unintentional.

Overall, people are aggregating (and filtering!) pretty well.

Interesting Observations (5) - Invalid AS numbers

Network	Next Hop	Path
*> 2001:468:501:A00::/56	2001:360:1:2::1	1221 3856 64999 i
*> 2001:468:501:ABE::/64	2001:360:1:2::1	1221 3856 64999 ?
*> 2001:468:501:ABF::/64	2001:360:1:2::1	1221 3856 64999 ?
* 3FFE:2C03::/32		
*	2001:360:1:2::1	1221 109 109 4618 3836 9681 17419 17715 64734 2012 i
*>	3FFE:C00:8023:19::1	109 4618 3836 9681 17419 17715 64734 2012 i
*> 3FFE:200:3E::/48	2001:7F8::CB9:0:1 0	3257 6762 3263 65001 1275 4 ?

again: prolonged leaks of private ASns to the global table :-)

Long-standing offenders AS45333 and AS45328 are gone (changed to 11340/disappeared). *Thank you!*

News (?)

- 6bone (3FFE:...) going away, end date: 2006/06/06
- no martian (bogon) networks seen since 2002/10/21
- private/unallocated AS numbers reappearing ?!
- people build native commercial IPv6 backbones across Europe (Tiscali, Easynet, others?)
- is there progress concerning tunnel cleanup and bogon filtering??
- more people actually look at traceroutes and fix things
- overall structure really improving, towards production quality (to be defined as: IPv6 path is no worse than the IPv4 path)
- US region catching up on allocations, but still lacking far behind on actually advertised routes

Where to go from here?

- more work needed on filtering recommendations
- more work on “routing BCP” recommendations (→ routing wg)
- still **much** cleanup work to do (“bad” tunnels, filters, unsolicited transit relations)
- bug your upstream providers to offer native IPv6 upstream
- have an eye on traceroute(6)s to find out which ways packets are travelling, and resolve stupid paths if possible
- consider de-peering non-useful peers (bad tunnels)
- *talk* to your peers and help them fix their stuff

IPv6 routing recommendations

- MIPP project recommendations:
 - no peerings over 'bad' tunnels (high RTTs / 3rd parties)
 - apply incoming prefix filters to peers
 - filter private ASn and overly long paths
- do not give unrestricted IPv6 transit to peers unless asked to
- do not take IPv6 transit from too many upstreams
- avoid taking your single upstream over intercontinental tunnel

References

- Merit 6bone routing report:
<http://www.merit.edu/mail.archives/html/6bone-routing-report/>
- List of IPv6 blocks allocated by the RIRs:
<http://www.ripe.net/rs/ipv6/ipv6allocs.html>
- MIPP (minimum peering policy) project:
<http://ip6.de.easynet.net/ipv6-minimum-peering.txt>
- Ghost Route Hunter: <http://www.sixxs.net/tools/grh/>
- IPv6 sample prefix list page
<http://www.space.net/~gert/RIPE/ipv6-filters.html>
- Slides are available at:
<http://www.space.net/~gert/RIPE/R46-v6-table/>

Questions?

`gert@space.net`