Impressions

An overview of the global IPv6 routing table

Gert Döring, SpaceNet AG, Munich

14. May 2003

RIPE 45, Barcelona

Overview

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

Slides online at: http://www.space.net/~gert/RIPE/R45-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 2003/05/13: 526 prefixes in total (2003/01/27: 399)

/n	global	RIR space	6bone	6to4	(2003/01/27)
/16	1	0	0	1	(1 0 0)
/24	47	0	**47**	0	(48 0 48)
/28	48	0	**48**	0	(50 0 50)
/32	225	**195**	**30**	0	(169 142 26)
/33	2	1	0	1	(1 0 0)
/34	2	0	1	1	(2 0 1)
/35	53	**53**	0	0	(66 66 0)
/36	2	1	0	1	(2 1 0)
/40	6	5	1	0	(6 5 1)
/41	5	5	0	0	(0 0 0)
/42-/45	3	3	0	0	(1 1 0)
/48	81	43	35	3	(38 19 17)
/52-/60	3	0	3	0	(1 1 0)
/64	46	39	7	0	(13 5 8)
/126	1	1	0	0	(0 0 0)
/128	1	1	0	0	(1 1 0)

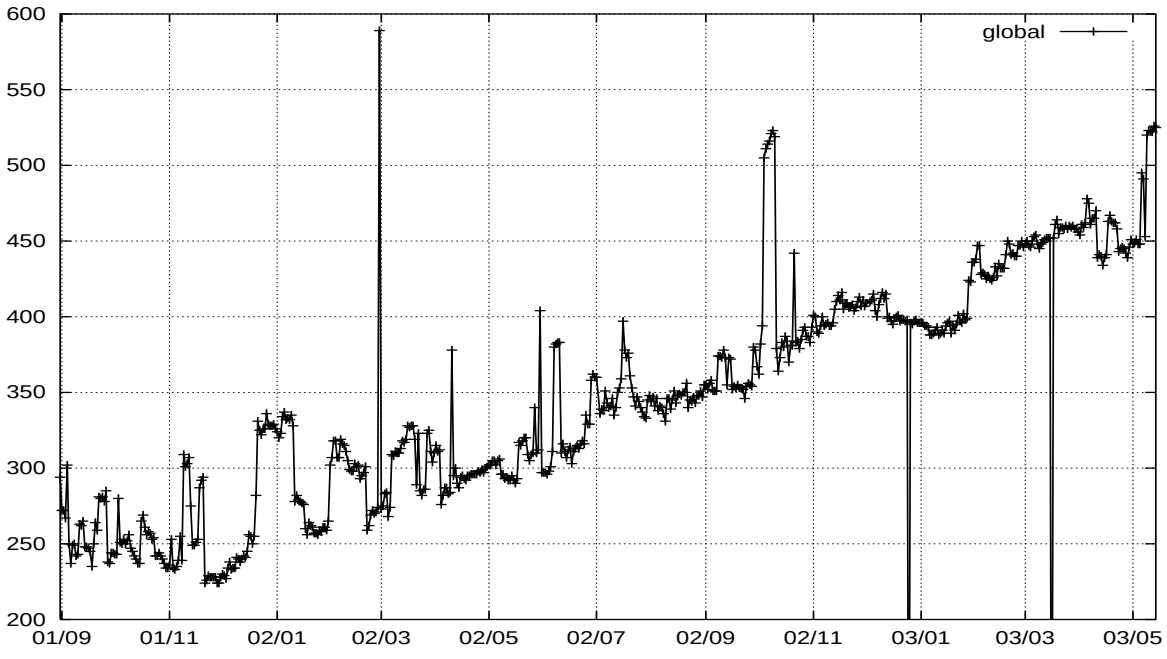
6to4 - 2002::/16

- 6to4 prefix 2002::/16 currently announced by (at least)

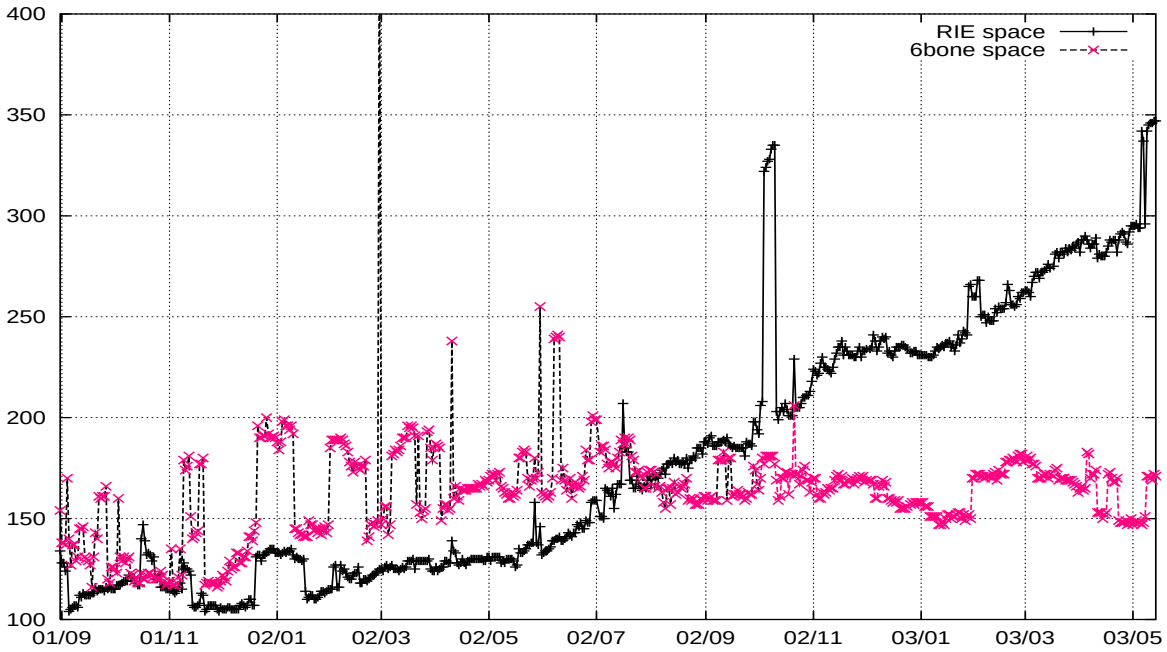
AS	Organization
109	Cisco
1741	FUnet.FI
1752	BT UK
1930	RCCN (portuguese NREN) (gone?)
2012	Eotvos Lorand University, Budapest
3246	Song Networks, Helsinki (only to Finnish peers)
5408	Greek Reasearch Network (recently: /48 only)
8379	Cybernet Germany (now PSInet)

- a number of networks have set up private relays (good!)
- some more-specific pfxs seen (prohibited by RFC3056 5.2.3!)
- anycast relay address 192.88.99.1/24, RFC3068

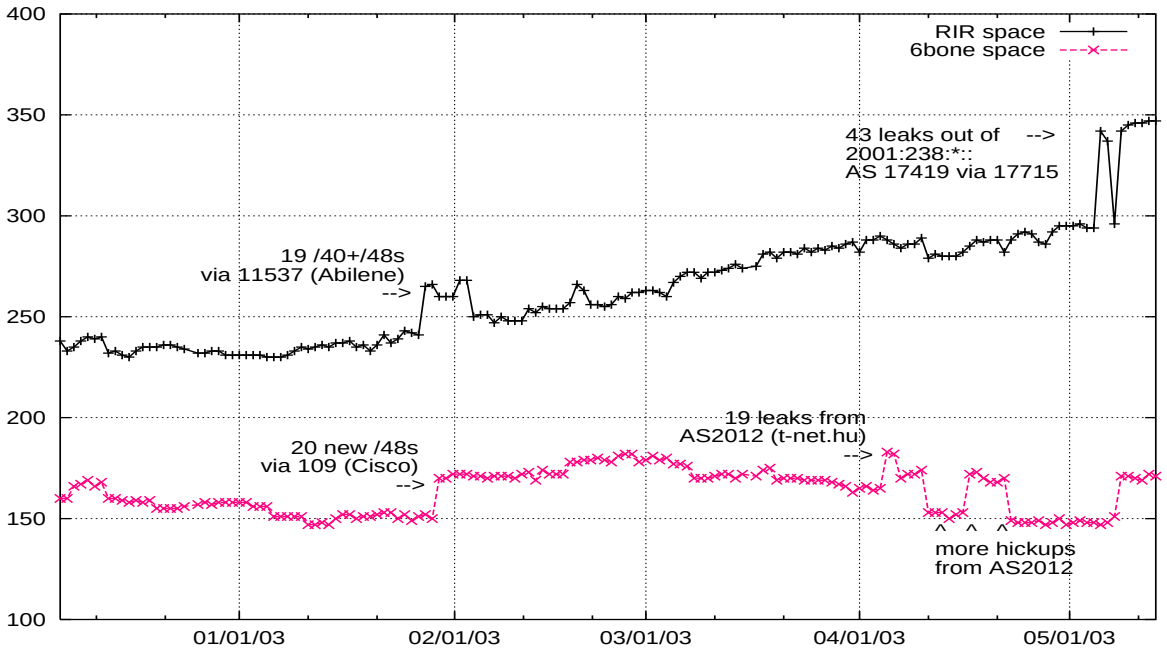
Graphics: Total Prefixes - 20 months



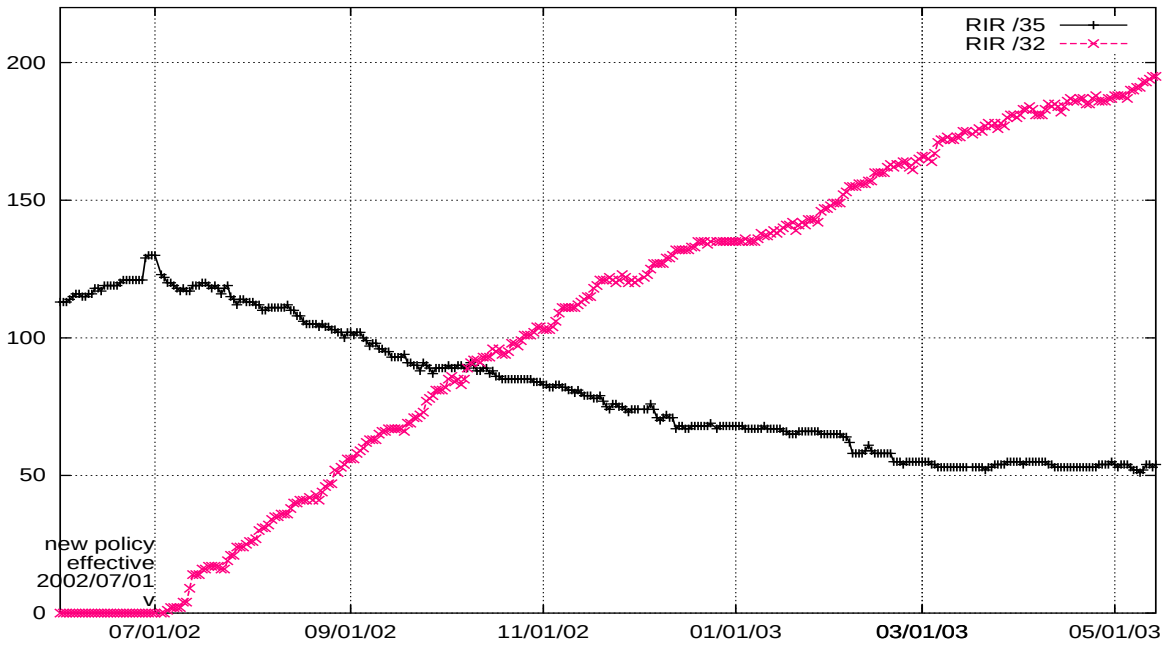
Graphics: RIR vs. 6Bone Prefixes - 20 months



Graphics: RIR vs. 6Bone Prefixes - 3 months



Graphics: RIR /35s vs. /32s



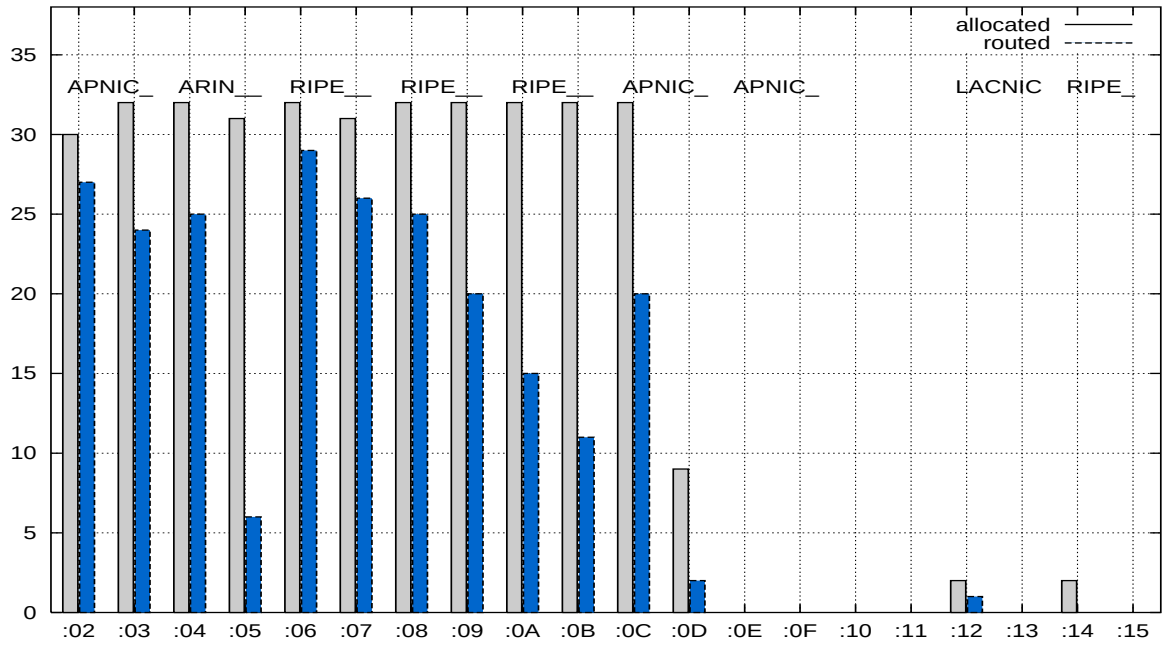
Numbers: RIRs, ASes and AS paths

- 361 LIR blocks out of 2001::/16 allocated by RIRs:
  ARIN 63(45), APNIC 103(96), RIPE 193(151), LACNIC 2(0)
  as of 2003/05/13 (2003/01/27)
- plus some IXP and other microallocations
- 235(193) allocations visible
- allocations take up 248(208) routes:
  53(66) /35s, 195(142) /32s, 18(15) allocations visible as /32 *and* /35
- total unique ASes in the IPv6 BGP table: 335(305)
- new IPv6 block for APNIC since 2003/02/04: **2001:0E00::/23**
- new IPv6 block for RIPE since 2003/02/26: **2001:1400::/23**
- new IPv6 block for ARIN since 2003/04: **2001:1800::/23**

Numbers: new /32s and /35s since 2003/01/27

RIPE		APNIC	ARIN
2001:890::/32	2001:B10:C000::/35	2001:C20::/32	2001:530::/32
2001:8E8::/32	2001:B18::/32	2001:C70::/32	2001:570::/48 ?
2001:9C0::/32	2001:B30::/32	2001:CD0::/32	2001:588::/32
2001:9E0::/32	2001:B38::/32 !	2001:CE0::/32	2001:598::/23 !
2001:A30::/32 !	2001:B40::/32	2001:CE8::/32	2001:5A8::/32
2001:A40::/32	2001:B60::/32	2001:CF8::/32	2001:5B8::/32
2001:A50::/32 !	2001:B70::/32	2001:D20::/32	2001:5C0::/32
2001:A88:E000::/35	2001:BA0::/32	2001:DC0::/35	
2001:AA0::/32	2001:BB0::/32	2001:DC0:2000::/35	
2001:AB8::/32	2001:BB8::/32		
2001:AC0::/32	2001:BC0::/32		
2001:AD0::/32	2001:BF8::/32		LACNIC
2001:AE0::/32			2001:1200::/32
2001:AE8::/32			
2001:AF8::/32		! = invisible	
		@ 2003/05/13	

Graphics: Allocated vs. Routed



Interesting Observations (1) - Ghost Busting

```
   Network       Path
* 3FFE:1D00::/24 1930 2200 3425 293 5609 5623 i
*>i              3257 5623 i
*                6939 6939 5623 i
*                4554 5609 5623 i

* 3FFE:1F00::/24 3561 8379 1849 12702 3549 4697 3320 293
                   6175 7580 10566 15180 1251 1916 11537 22 5609 6830 1755 i
*                3274 790 3549 4697 3320 293
                   6175 7580 10566 15180 1251 1916 11537 22 5609 6830 1755 i
* i              3257 5511 4697 3320 293
                   6175 7580 10566 15180 1251 1916 11537 22 5609 6830 1755 i
*                1930 2200 9270 17832 3748 4725 4697 3320 293
                   6175 7580 10566 15180 1251 1916 11537 22 5609 6830 1755 i
*                3320 6939 14277 4725 1752 5408
                   6175 7580 10566 15180 1251 1916 11537 22 5609 6830 1755 i
*                6939 14277 4725 1752 5408
                   6175 7580 10566 15180 1251 1916 11537 22 5609 6830 1755 i
*                4554 278 6939 4725 1752 5408
                   6175 7580 10566 15180 1251 1916 11537 22 5609 6830 1755 i
```

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 (2) - Count-to-Infinity

```
Start:
*  2001:650::/35    2001:470:1FFF:2::          6939 3257 3561 ?
*                   3FFE:1108:40A:FFFF::1:2    3274 790 790 3561 ?
*                   2001:608:0:3::7            1930 3561 ?
*>                  2001:650:0:2::26           3561 ?

After withdrawal (snapshots every minute or so):
  3320 13110 25336 25489 1752 1930 3561 ?
  8560 12731 109 109 10479 10566 6175 6830 1752 3561 ?
  109 10479 10566 6175 6830 1752 3561 ?
...
  109 1251 1916 11537 6939 14277 33 15589 25336 25489 1752 1930 3561 ?
...
  3320 5430 13285 786 11537 22 10566 2042 4774 2497 4697 3748 4725 14277 +
       33 15589 25336 25489 1752 1930 3561 ?
...
  680 2200 1916 11537 145 7580 10566 13944 22 5609 3320 6939 4716 2500 4697 +
       3748 4725 14277 33 15589 25336 25489 1752 1930 3561 ?
...
  8472 8903 16091 513 559 6680 1103 11537 786 1849 20834 2549 3320 293 5609 +
       22 10566 2042 4774 2497 5994 8002 6939 4716 2500 4697 3748 4725 14277 +
       33 15589 25336 25489 1752 1930 3561
```

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

Interesting Observations (3) - ugly tunnels

```
traceroute to moebius2.space.net (195.30.1.100), 30 hops max, 38 byte packets
 1  213.150.184.129 (213.150.184.129)  4.404 ms  2.820 ms  2.793 ms
 2  172.16.0.1 (172.16.0.1)  7.230 ms  4.513 ms  4.052 ms
 3  193.95.43.99 (193.95.43.99)  4.290 ms  4.840 ms  4.535 ms
 4  193.95.1.38 (193.95.1.38)  5.073 ms  5.237 ms  5.922 ms
 5  193.95.50.202 (193.95.50.202)  6.776 ms  6.549 ms  5.535 ms
 6  pal1-tunisie-telecom-9-tn.seabone.net (195.22.205.137)  11.543 ms 11.621 ms
 7  decix-fra2-racc1.fra.seabone.net (195.22.211.205)  39.658 ms  40.669 ms
 8  decix.Space.net (80.81.192.105)  39.890 ms  39.544 ms  39.796 ms
 9  Cisco-F-V-Fa6-0-0.Space.Net (193.149.44.254)  41.516 ms  41.866 ms 42.858 ms
10  Cisco-M-XII-POS4-0.Space.Net (195.30.3.33)  44.635 ms  42.800 ms 43.279 ms
11  Cisco-M-XI-Vlan1.Space.Net (195.30.0.116)  49.897 ms
12  moebius2.Space.Net (195.30.1.100)  47.008 ms  47.636 ms  47.945 ms

traceroute6 to moebius2.space.net (2001:608:0:1::100) from
2001:970:6800:0:202:2dff:fe03:e156, 30 hops max, 16 byte packets
 1  2001:970:6800::1 (2001:970:6800::1)  3.115 ms *  2.926 ms
 2  2001:970:0:2::20 (2001:970:0:2::20)  9.085 ms *  18.411 ms
 3  tu-viagenie.ipv6.noris.de (2001:780::b)  157.448 ms  159.842 ms 155.208 ms
 4  3ffe:b00:c18::11 (3ffe:b00:c18::11)  307.153 ms  336.514 ms  315.435 ms
 5  Cisco-M-VI-Tu18.Space.Net (2001:608:0:3::6)  375.616 ms  358.7 ms 358.521 ms
 6  Cisco-M-VIII.space.net (2001:608::119)  362.091 ms  365.437 ms 361.65 ms
 7  moebius2 (2001:608:0:1::100)  359.427 ms  368.408 ms  359.534 ms
```

⇒ IPv6 goes via tunnel to US and back to EU ⇒ *slow*

Interesting Observations (4) - more-specific leaks

```
   Network            Next Hop              Path
*  2001:238::/32      2001:478:FFFF::1      4555 6939 3425 17419 i
* i                   2001:7F8::CB9:0:1     3257 17419 i
*                     2001:470:1FFF:2::     6939 3425 17419 i
*                     2001:7F8:2:8001::2    1752 2914 3549 17419 i
*>i                   2001:450:1:2001::AA   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
```

This is a recent leak of more-specifics, likely not intentional.

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

Interesting Observations (5) - inconsistant origin

```
   Network   Next Hop                Path
*  2001:200::/35
*            2001:608:0:3::D         5430 3549 2500 i
*            2001:470:1FFF:2::       6939 4716 2500 i
*>           2001:7F8:2:8001::2      1752 ?
*            2001:608:0:3::9         3320 1752 ?
*  2001:200::/32
*            2001:608:0:3::D         5430 3549 2500 i
*            2001:470:1FFF:2::       6939 4716 2500 i
*>           2001:7F8:2:8001::2      1752 ?
*            2001:608:0:3::9         3320 1752 ?
```

This was a typo at BT (1725), effectively blackholing WIDE (2500) from Europe for about a week. It was resolved quickly after notifying BT (thanks to Hank N. and Stuart P.) - demonstrates need for proper filtering.

There are a few more examples - *show bgp ipv6 inconsistant-as*.

Interesting Observations (6) - Invalid AS numbers

```
   Network   Next Hop                Path
* i3FFE:400A::/32
            2001:650:0:200::4       8472 8903 65200 16091 i
*           3FFE:1108:40A:FFFF::1:2 3274 790 790 6830 559 513 16091 i
*           2001:608:0:3::D         5430 13285 6939 6939 513 513 16091 i

*>i3FFE:4005:A::/48
            2001:650:F807::DE9:1    3561 5511 6830 8733 12702 12702 65272 i

*> 3FFE:2C05::/32
            2001:608:0:3::7         1930 1752 64734 i
```

Only very short-term leaks of private/reserved AS numbers observed.

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

Good News

- no martian (bogon) networks seen since 2002/10/21
- private/unallocated AS numbers mostly gone
- people build native commercial IPv6 backbones across Europe (Tiscali, Easynet)
- people massively cleaning up their international tunnels (Cisco)
- people are doing bogon filtering (/64s and such)
- 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)

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/R45-v6-table/

Questions?

gert@space.net