

Impressions

An overview of the global IPv6 routing table

Gert Döring, SpaceNet AG, Munich, Germany

May 3rd, 2005

RIPE 50, Stockholm, SE

Overview

- numbers
- pictures & trends
- things that should not be there...
- conclusions & recommendations
- references

Slides online at: <http://www.space.net/~gert/RIPE/R50-v6-table/>

Numbers - AS numbers

- as of 2005/05/02: 517 unique AS numbers visible (04/09: 474)
 - 337 origin-only ASes (no transit paths seen) (292)
 - 167 ASes originate & give transit (167)
 - 13 transit-only ASes (e.g. 1237, 1717, 4774, 6667, ...) (15)
- mixture of RIR (2001::) and 6Bone (3FFE::) space announced
 - 366 ASes originate 1 RIR prefix (314)
 - 34 ASes originate 1 6Bone prefix (45)
 - 42 ASes originate 1 6Bone + 1 RIR prefix (43)
 - 29 ASes originate 2 RIR prefixes (5 due to /32+/35)
 - 33 ASes with “more than that”, maximum is 6 prefixes
- 8 ASes still announce their prefix as /32 and /35
- note: all paths observed from AS5539

ASes - why are people announcing 2 prefixes?

- 6bone to RIR migration: 1 6bone, 1 RIR prefix, *temporary*

2001:420::/35	109 i
3FFE:C00::/24	109 i

- /35 to /32 migration: 2 RIR prefixes, *temporary*

2001:258::/32	2914 2510 i
2001:258::/35	2914 2510 i

- experiments and/or leaks?

2001:4F8::/32	1273 3557 i
2001:4F8:8::/48	3549 3557 i
2001:4F8:FFFF::/48	3549 3557 i

- multi-uplink-/multi-homing-experiments? IXPs?

2001:770:80::/48	3257 2110 2128 i
2001:7F8:18::/48	3257 2110 2128 i

- mergers and acquisitions, business units, growth, ...

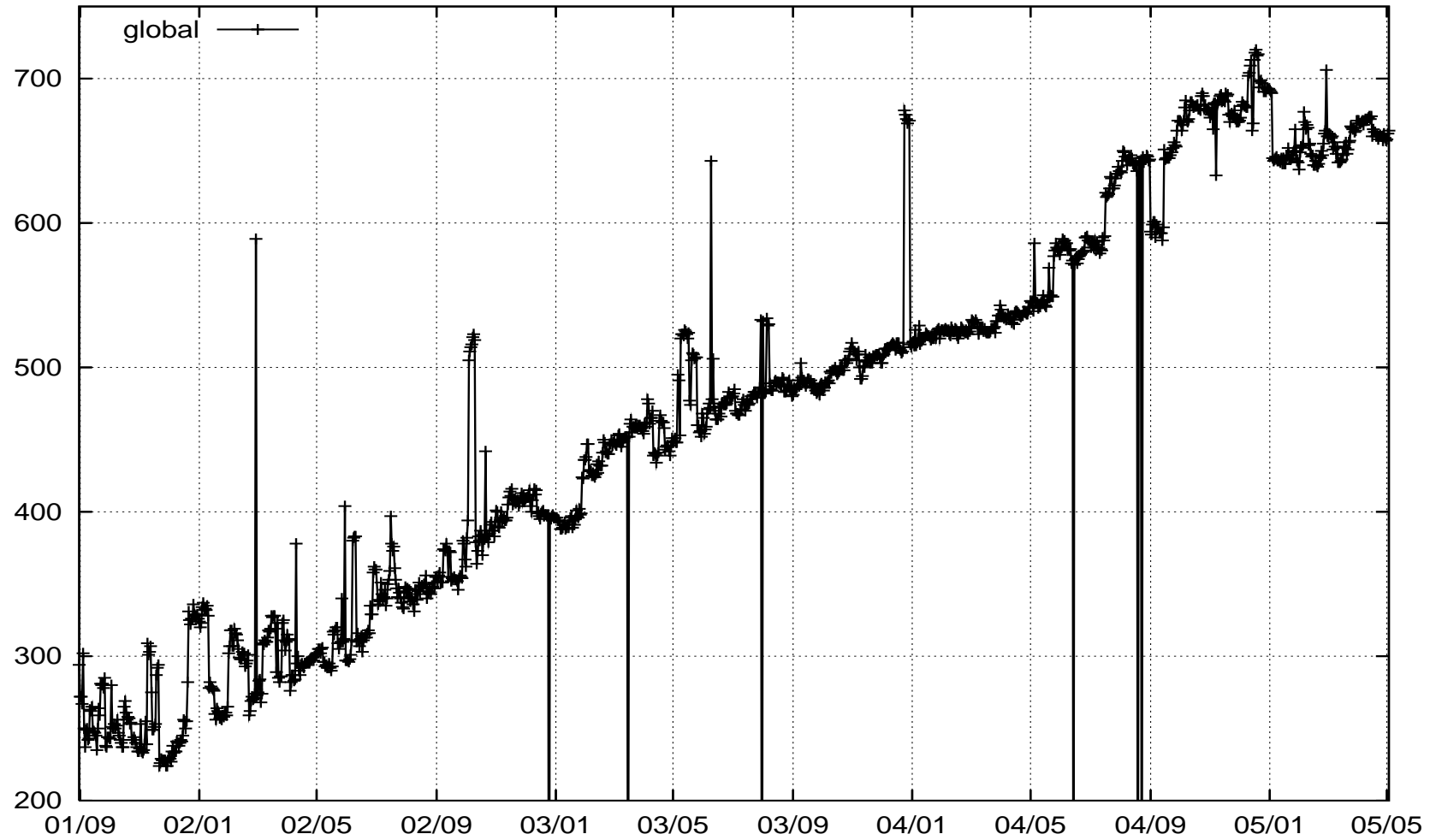
2001:360::/32	1221 i
2001:8000::/20	1221 i

Numbers - Prefixes

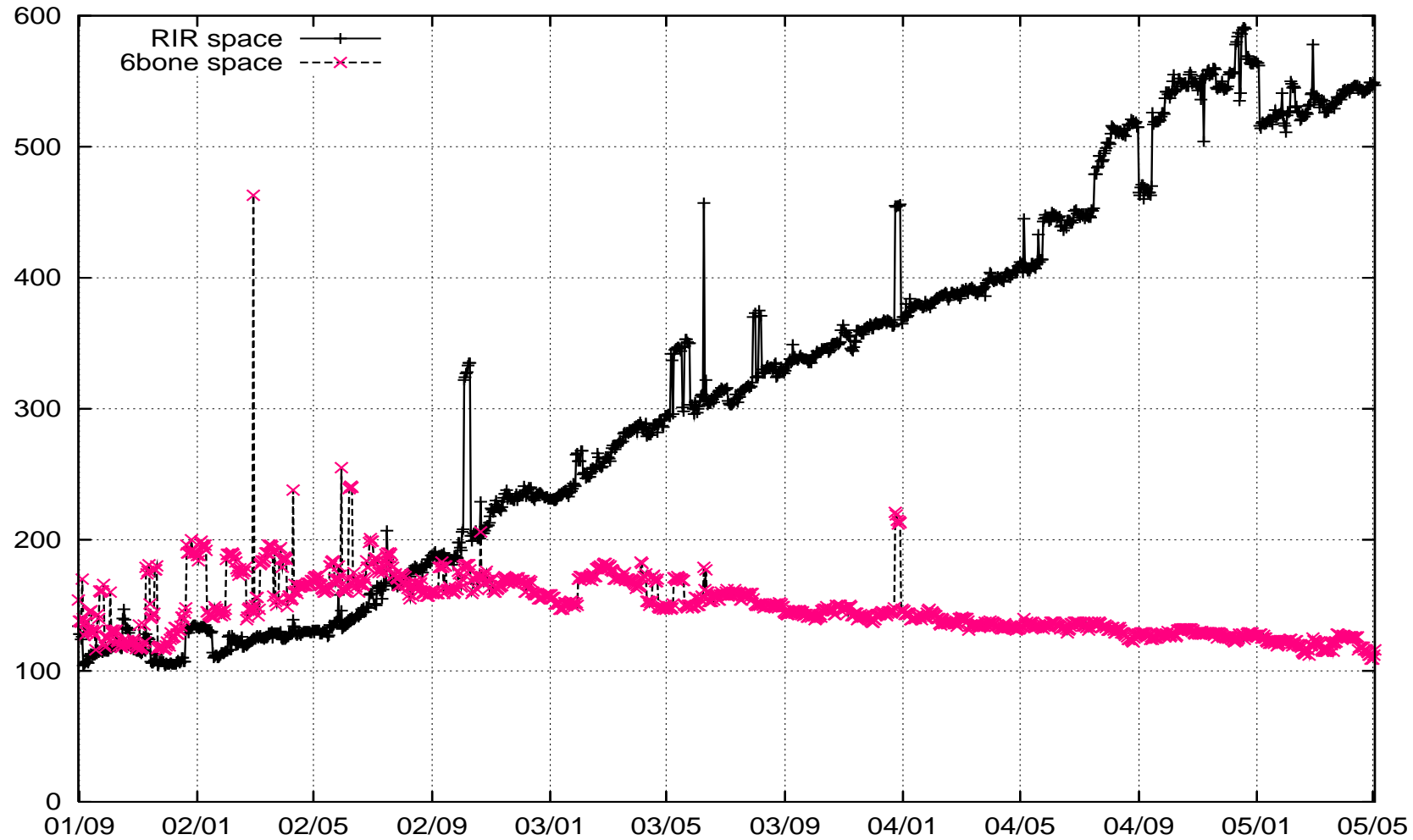
As of 2005/05/02: 662 prefixes in total (2004/09/20: 645)

/n	global	RIR space	6bone	6to4	(2004/09/20)
/16	1	0	0	1	(1 0 0 1)
/20-21	3	3	0	0	(2 2 0 0)
/24	36	0	36	0	(38 0 38 0)
/27	1	1	0	0	(1 1 0 0)
/28	33	1	32	0	(40 1 39 0)
/30	1	1	0	0	(1 1 0 0)
/32	459	431	28	0	(402 373 29 0)
/33-/34	3	3	0	0	(2 2 0 0)
/35	25	25	0	0	(42 41 1 0)
/36-/39	2	1	1	0	(0 0 0 0)
/40	9	8	1	0	(3 3 0 0)
/41-/45	3	3	0	0	(3 3 0 0)
/48	76	64	12	0	(109 92 17 0)
/52-/60	1	1	0	0	(0 0 0 0)
/64	8	6	2	0	(1 0 1 0)
/65-/128	1	1	0	0	(0 0 0 0)

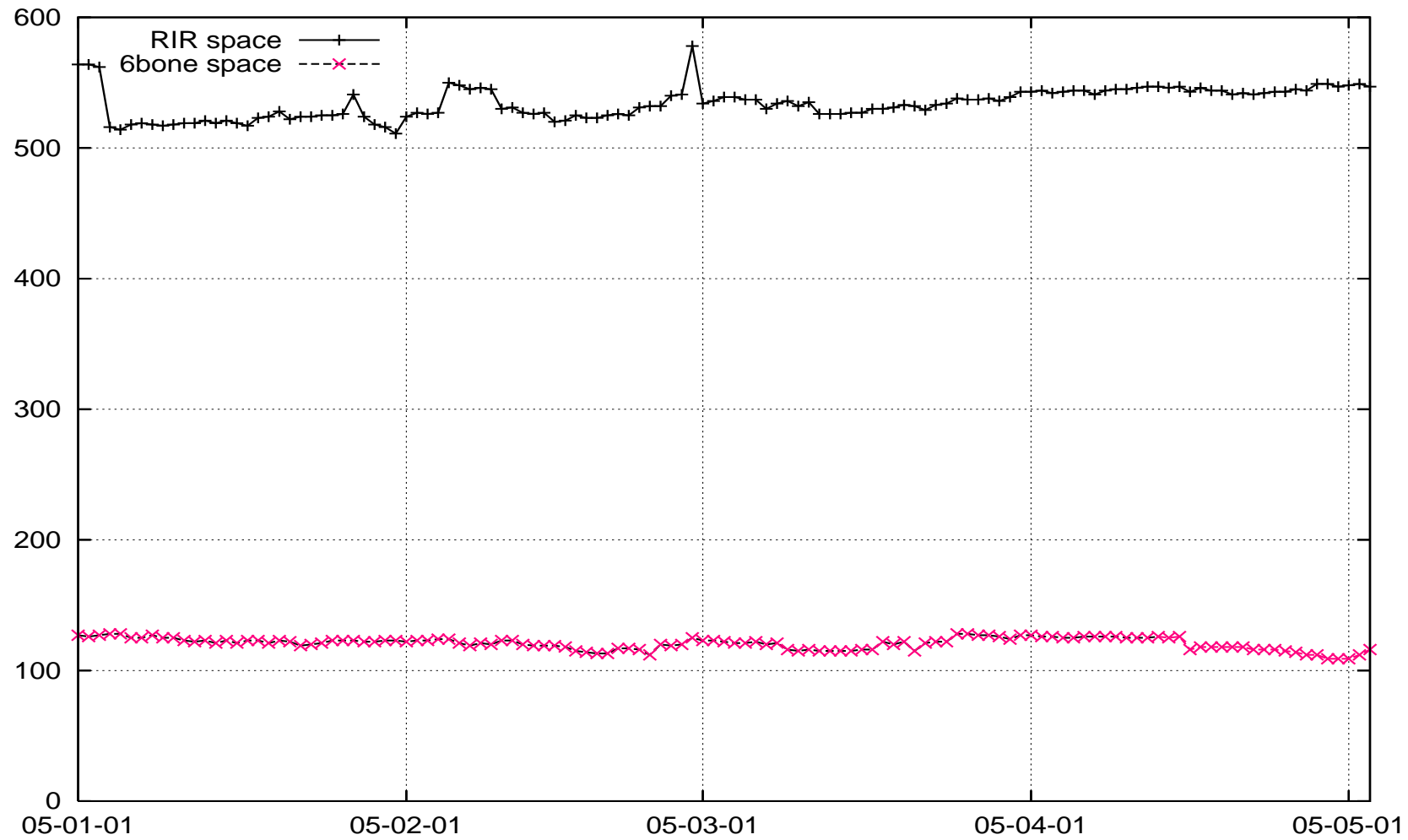
Graphics: Total Prefixes - 44 months



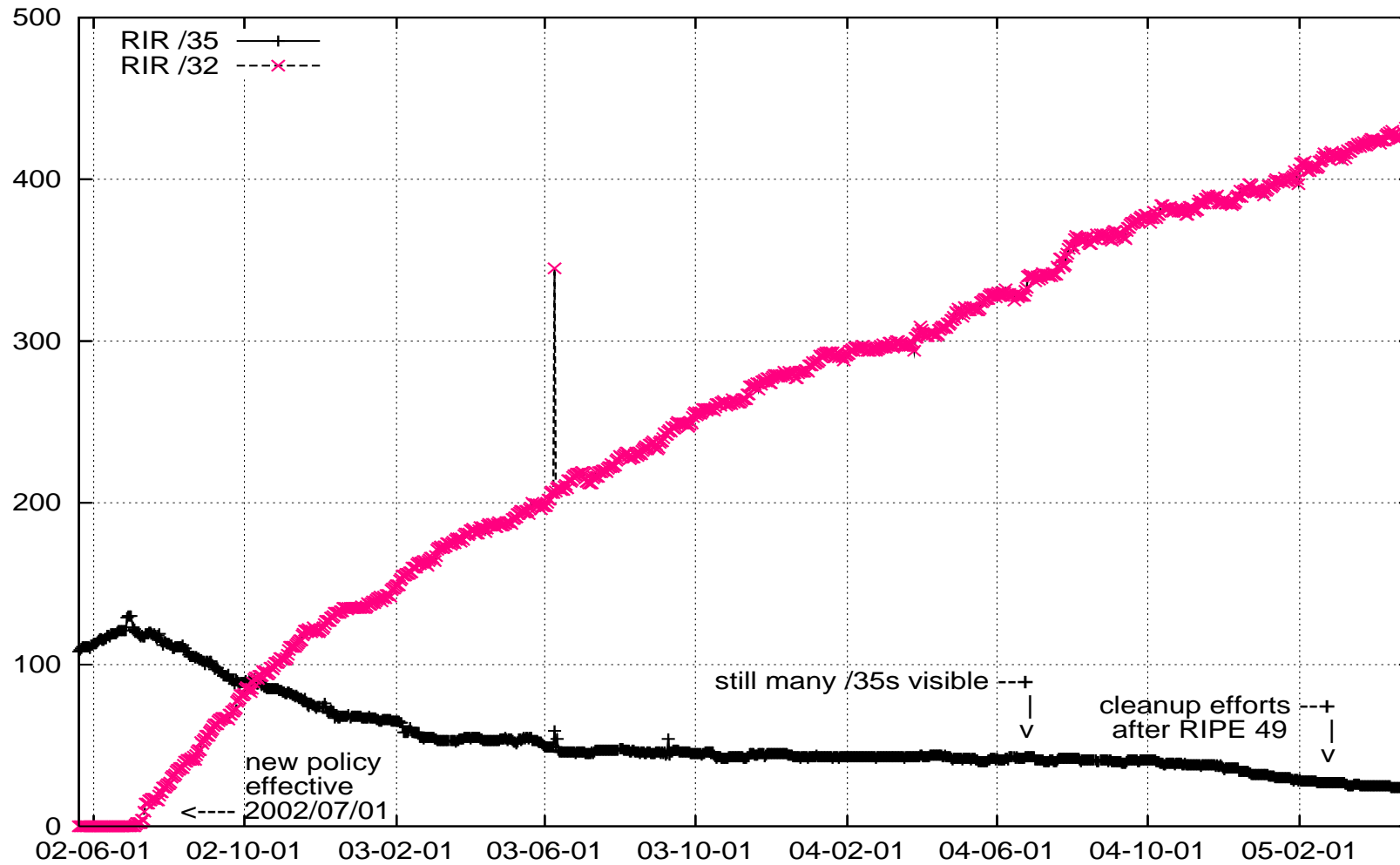
Graphics: RIR vs. 6Bone Prefixes - 44 months



Graphics: RIR vs. 6Bone Prefixes - 4 months



Graphics: RIR /35s vs. /32s



Why /35s?

- Non-upgraded /35 allocations (*only* /35 seen)?
APNIC: 5, ARIN: 4, RIPE: 2 prefixes - total: 11
- dual-announcements /32+/35 (same path) - fear of ghosts?
APNIC: 4, ARIN: 3, RIPE: 0 prefixes - total: 7
- Traffic-Engineering with more-specific announcements??

2001:808::/32	3549 20965 8501 9112	i
2001:808::/35	3320 9112	i
- Multihomed customers? Shared prefixes?

2001:BF0::/32	20646 12732	i
2001:BF0:C000::/35	6939 2497 4589 29670	i
2001:490::/32	3549 14277	i (Nokia)
2001:490::/35	3549 12702 1248	i (Nokia Finland)
2001:490:C000::/35	3549 18084 18666	i (Nokia Dallas)

Numbers: RIRs, Allocations, ...

- On 2005/05/02, 827 LIR blocks (2001::/16) allocated by RIRs:

RIR	alloc.	members	perc.	on 2004/09/20
ARIN	148	~ 2294	6.5%	118 (+25%)
APNIC	197	~ 1791	11.0%	170 (+16%)
RIPE	465	~ 4050	11.5%	392 (+18%)
LACNIC	17	~ 231	7.4%	10 (+70%)
AfriNIC	2+6 (*)	-	-	-

- actual *percentage* with IPv6 similar for RIPE and APNIC
- 442 (R49: 392) allocations visible in routing table (*only 53%!)*

Numbers: RIRs: notable allocations (1)

- more “very large” allocations seen:
 - 2001:4600::/24 to Telenor Norway (2004/10/06)
 - 2001:4a00::/27 to PIPEX UK (2004/12/01)
 - 2001:4ba0::/31 to Smartcall Norway (2005/02/18)
 - 2001:8000::/20 to Telstra Australia (AP) (2004/12/02)
 - 2001:A000::/21 to NTT West Japan (AP) (2004/12/01)
 - 2003::/19 to Deutsche Telekom (2005/01/13)
 - 2A00::/22 to Arcor DE (2005/04/20)
- ⇒ **check your BGP filters!!**
- ⇒ what sort of filters are *useful*? Some are more harmful...!

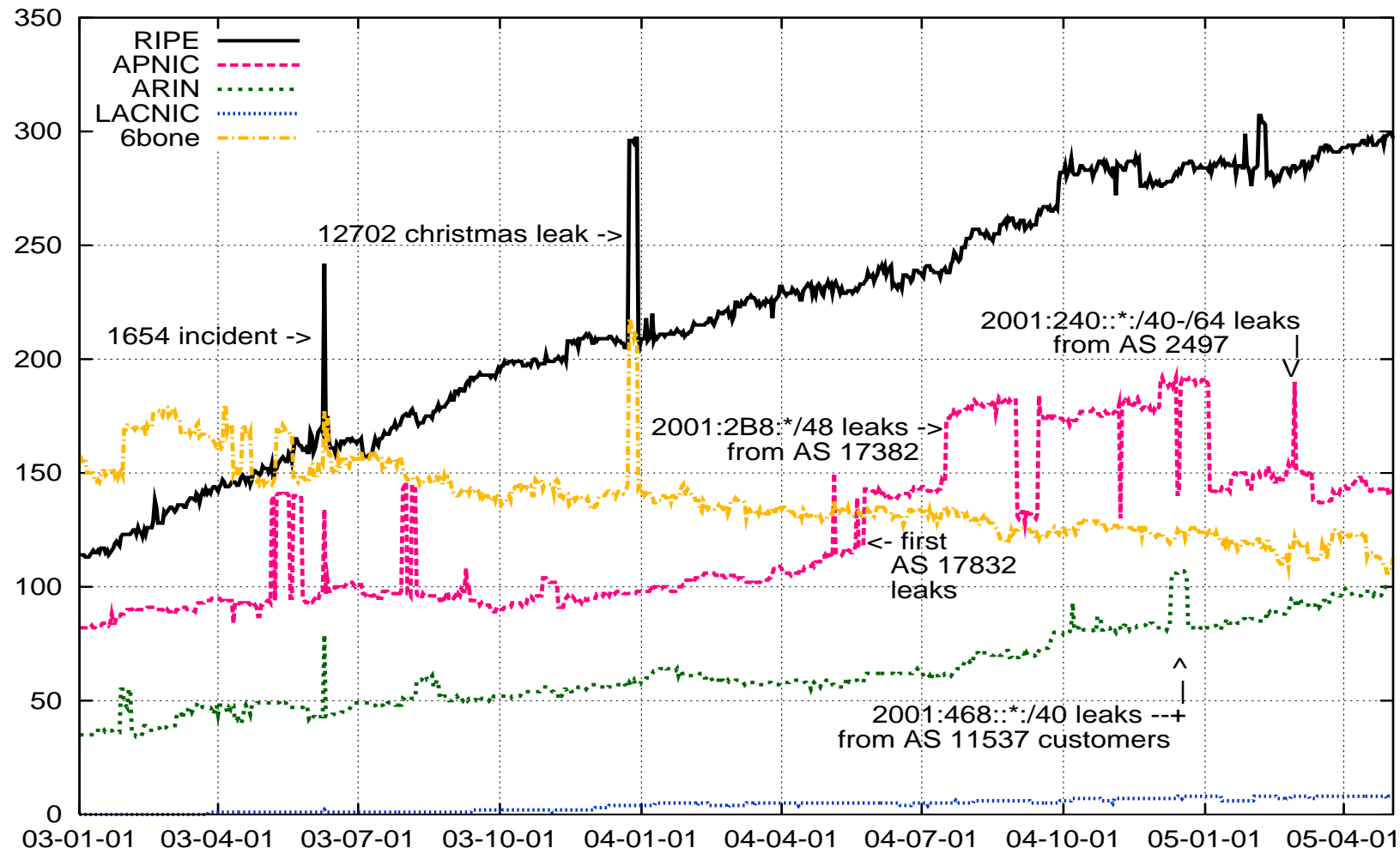
Numbers: RIRs: notable allocations (2)

- Allocations ICANN \Rightarrow RIRs since RIPE 49

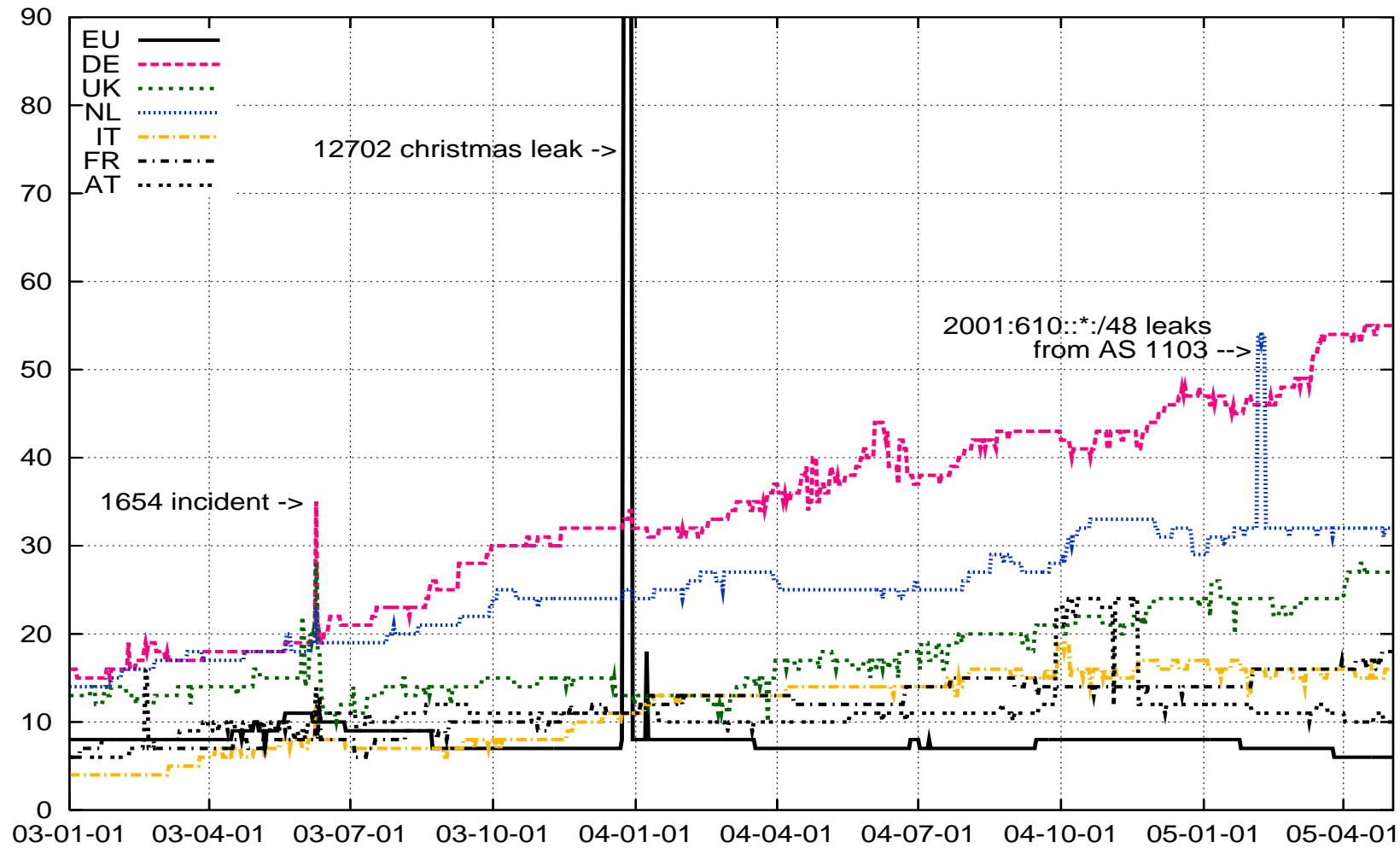
Prefix	RIR	Date	Comment
2001:4A00::/23	RIPE NCC	2004/10	
2001:4C00::/23	RIPE NCC	2004/12	
2001:5000::/20	RIPE NCC	2004/09	EU C&W
2001:8000::/19	APNIC	2004/11	AU/Telstra
2001:A000::/20	APNIC	2004/11	JP/NTT West
2003:0000::/18	RIPE NCC	2005/01	DE/DTAG
2600:0000::/22	ARIN	2005/04	US DoD
2604:0000::/22	ARIN	2005/04	US DoD
2608:0000::/22	ARIN	2005/04	US DoD
260C:0000::/22	ARIN	2005/04	US DoD
2A00:0000::/21	RIPE NCC	2005/04	DE/Arcor

- <http://www.iana.org/assignments/ipv6-unicast-address-assignments>

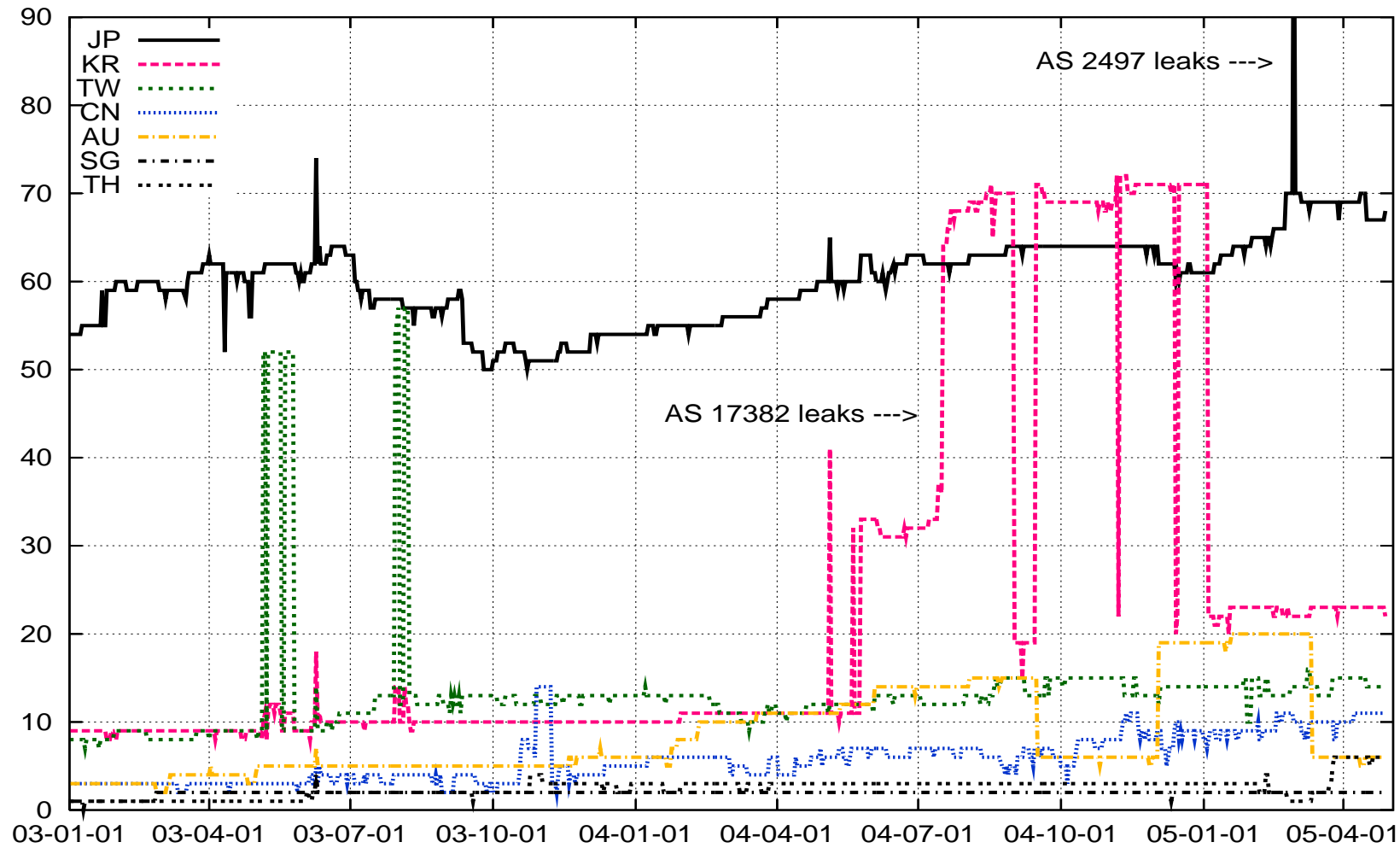
Graphics: prefixes by RIR region



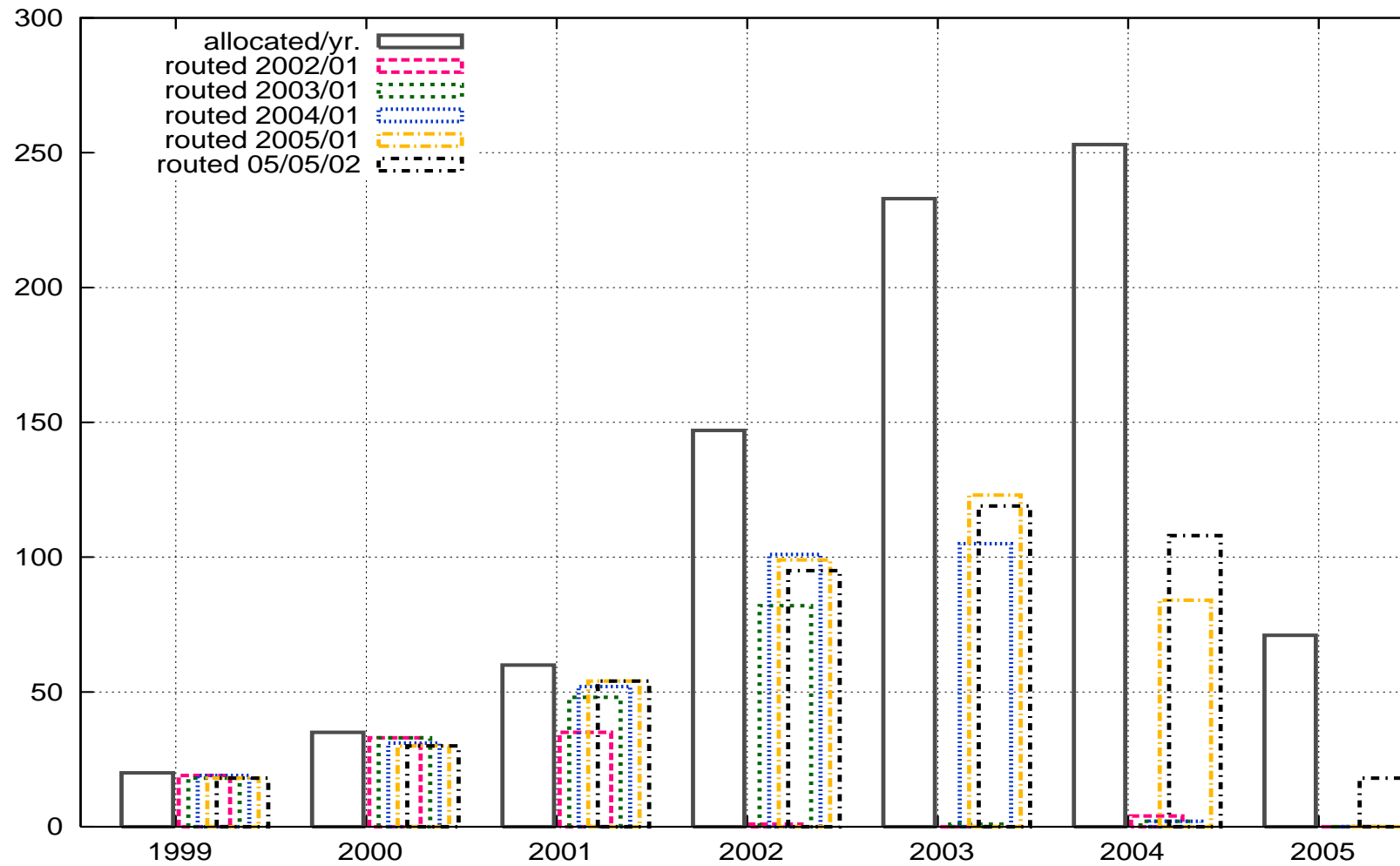
Graphics: prefixes by country (RIPE)



Graphics: prefixes by country (APNIC)



Graphics: Allocated vs. Routed



Interesting Observations (1) - Ghost Busting

```

Network      Path
* > 2001:450::/32  3549 i
*                6939 3549 i
*                3320 5430 3549 i

* > 2001:450::/35  3274 790 6667 6830 30071 3557 2914      1916 3549 i
*                1221 109 109 1251          1916 3549 i
*                1930 1930 1930 1930 1930 10566 109 1251 1916 3549 i
*                3320 5609 6939 109 1251    1916 3549 i
*                6939 109 1251             1916 3549 i
*                109 1251                 1916 3549 i
*                1752 5511 2500 4697 2914   1916 3549 i

```

- Ghosts = BGP withdrawal bug, caused by *buggy software*.
Long paths can stay *mostly unchanged* in the table for weeks.
- don't confuse with BGP count-to-infinity (= paths *change*).
- Ghost above seem to be caused by *current* IOS 12.2S at 3549!?
- Cisco ISP team acknowledges problem, but not yet fixed :-(
- maybe related to CSCeg30291?

Interesting Observations (2) - Accidental Hijack

Network	Next Hop	Path
* 2001:218::/32	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
*> 2001:228::/35	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
*> 2001:230::/35	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
*> 2001:270::/32	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
*> 2001:2b0::/32	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
*> 2001:400::/32	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
*> 2001:400::/35	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
*> 2001:408::/32	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
*> 2001:410::/32	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
* 2001:418::/32	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
...		
* 2001:608::/32	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
* 2001:610::/32	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
*> 2001:618::/32	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
*> 2001:620::/32	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
...		

- 29657 *really* likes to do this - happens every few months
- likely cause: IGP \Leftrightarrow BGP redistribution

Interesting Observations (3) - more specifics

Network	Next Hop	Path
* 2001:478::/32	2001:7F8:2:8001::2	1752 5408 6175 i
*	2001:470:1FFF:2::	6939 109 6175 i
*>i	2001:7F8::CB9:0:1	3257 6175 i
*	3FFE:C00:8023:19::1	109 6175 i
* 2001:478:65::/48	2001:470:1FFF:2::	6939 109 4555 i
*	3FFE:1108:40A:FFFF::1:2	3274 790 3549 6939 109 4555 i
*>	3FFE:C00:8023:19::1	109 4555 i
*> 2001:478:9200:51::/64	::FFFF:203.14.5.7	1221 3856 42 i
*> 2001:478:9200:52::/64	::FFFF:203.14.5.7	1221 3856 42 i
*> 2001:478:9200:53::/64	::FFFF:203.14.5.7	1221 3856 42 i
*> 2001:478:9200:FFFF::2/127	::FFFF:203.14.5.7	1221 3856 42 i
*> 2001:478:9200:FFFF::4/127	::FFFF:203.14.5.7	1221 3856 42 i
*> 2001:478:9200:FFFF::6/127	::FFFF:203.14.5.7	1221 3856 42 i
*> 2001:478:9200:FFFF::8/127	::FFFF:203.14.5.7	1221 3856 42 i

- /64s and /127s appeared for 3 days only - looks non-intentional
- the /48 looks like “more-specific BGP multihoming”

Interesting Observations (4) - Martian Networks

Network	Next Hop	Path
*> 2000:2:2::/64	2001:470:1FFF:2::	6939 4538 4134 ?
*> 2000:3:3::/64	2001:470:1FFF:2::	6939 4538 4134 ?
*>i3FF3:2C04::/32	2001:7F8:2C:1000:0:A500:8472:1	8472 8903 i

- alle announcements visible only a short time (days...weeks)
- only few ASes paths actually seen → *good news*
- downstream BGP announcement filtering still lacking, obviously

Observations (5) - funny AS paths

```
Feb 27 19:54:16: %BGP-6-ASPATH: Long AS path 20965 1299 3320 15589 15589 5397
    {33,109,145,293,559,816,1103,1273,1275,1752,1853,1930,2042,2200,
    2497,2500,2914,3257,3265,3333,3352,3425,3549,4691,4697,4716,4725,
    5511,5539,5609,5623,6175,6435,6453,6762,6830,6939,7580,7660,8447,
    8472,8763,9264,10566,12779,12793,12859,13944,14277,15897,17715,
    17965,24136,24895,25358,29377,29686,31103,32266}
    received from 2001:798:201B:10AA::1: More than configured MAXAS-LIMIT
Apr 19 17:36:10 %BGP-6-ASPATH: Long AS path 20965 1299 3320 15589 15589 5397
    {33,109,145,278,293,513,559,1257,1275,1752,1853,2042,2497,2500,
    2607,2914,3257,3265,3292,3352,3425,3549,3748,3786,4691,4697,4716,4725,
    ...
    received from 2001:798:201B:10AA::1: More than configured MAXAS-LIMIT
```

- research project from AS5397, Roma Tre University (idea seems to be to do BGP AS charting via special announcements).
- is this a good idea?
- danger of triggering BGP implementation bugs?!

new tool: GRH prefix comparison

- check prefix distribution through your peers / upstream's filters
- <http://www.sixxs.net/tools/grh/compare/>
- was built because C&W's new prefix (2001:5000::/21) was filtered in many places ("nobody needs a prefix that big")
- \Rightarrow compare 2001:650::/32 to 2001:5000::/21

ASN	2001:650::/32's ASPath	2001:5000::/21's ASPath	GRH peer
2914	*6175 *6830 8472 1273	*6680 *786 *1752 8472 1273	Verio
3257	*6175 6830 8472 1273	*3549 6830 8472 1273	Tiscali
3265	3549 6830 8472 1273	3549 6830 8472 1273	XS4ALL

- \Rightarrow *if you deploy BGP filters (which is a good thing!), make sure that you stay up-to-date on recent developments!*

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
- ...and now we need *content* on the v6 network

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

- Ghost Route Hunter: <http://www.sixxs.net/tools/grh/>
- 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>
- IPv6 sample prefix filter page
<http://www.space.net/~gert/RIPE/ipv6-filters.html>
- Slides are available at:
<http://www.space.net/~gert/RIPE/R50-v6-table/>

Questions?

`gert@space.net`