

Impressions

An overview of the global IPv6 routing table

Gert Döring, SpaceNet AG, Munich, Germany

Sept 22nd, 2004

RIPE 49, Manchester, UK

Overview

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

Slides online at: <http://www.space.net/~gert/RIPE/R49-v6-table/>

Numbers - AS numbers

- as of 2004/09/20: 474 unique AS numbers visible (05/03: 421)
 - 292 origin-only ASes (no transit paths seen) (263)
 - 167 ASes originate & give transit (144)
 - 15 transit-only ASes (e.g. 1237, 1717, 4774, 6667, ...) (14)
- mixture of RIR (2001::) and 6Bone (3FFE::) space announced
 - 314 ASes originate 1 RIR prefix (258)
 - 45 ASes originate 1 6Bone prefix (50)
 - 43 ASes originate 1 6Bone + 1 RIR prefix (43)
 - 28 ASes originate 2 RIR prefixes (10 due to /32+/35)
 - 29 ASes with “more than that”, maximum is 55 prefixes
- 16 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*

2001:760::/32 3549 20965 137 i

2001:760::/35 680 6880 137 i

- experiments and/or leaks?

2001:2B8::/32 1752 17832 i

2001:2B8:2::/48 1853 6680 1103 11537 17579 1237 7623 17832 i

2001:2B8:80::/48 1853 6680 1103 11537 17579 1237 7623 17832 i

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

2001:610:140::/48 3549 1200 i

2001:7B8:200::/48 3549 1200 i

2001:7F8:1::/48 3549 1200 i (AMS-IX)

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

2001:650::/32 1273 i

2001:5000::/21 1273 i

Numbers - Prefixes

As of 2004/09/20: 645 prefixes in total (2004/05/03: 546)

/n	global	RIR space	6bone	6to4	(2004/05/03)
/16	1	0	0	1	(1 0 0 1)
/20-21	2	2	0	0	(0 0 0 0)
/ 24	38	0	38	0	(39 0 39 0)
/27	1	1	0	0	(1 1 0 0)
/ 28	40	1	39	0	(40 0 40 0)
/30	1	1	0	0	(0 0 0 0)
/ 32	402	373	29	0	(350 319 31 0)
/33	2	2	0	0	(2 2 0 0)
/ 35	42	41	1	0	(42 42 0 0)
/36-/39	0	0	0	0	(1 1 0 0)
/40	3	3	0	0	(5 4 1 0)
/41-/45	3	3	0	0	(5 5 0 0)
/48	109	92	17	0	(59 37 22 0)
/52-/60	0	0	0	0	(0 0 0 0)
/64	1	0	1	0	(1 0 1 0)
/65-/128	0	0	0	0	(0 0 0 0)

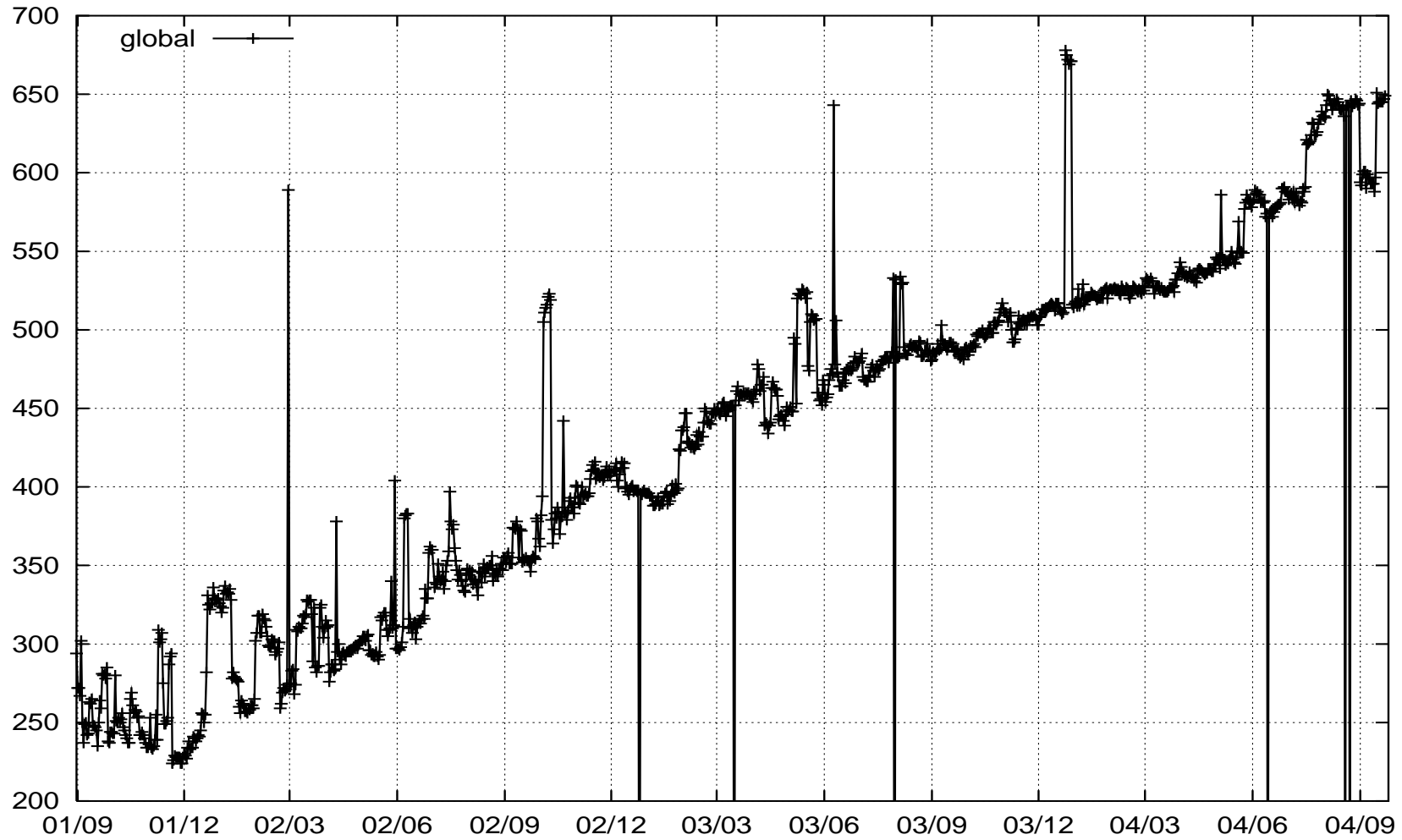
6to4 - 2002::/16

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

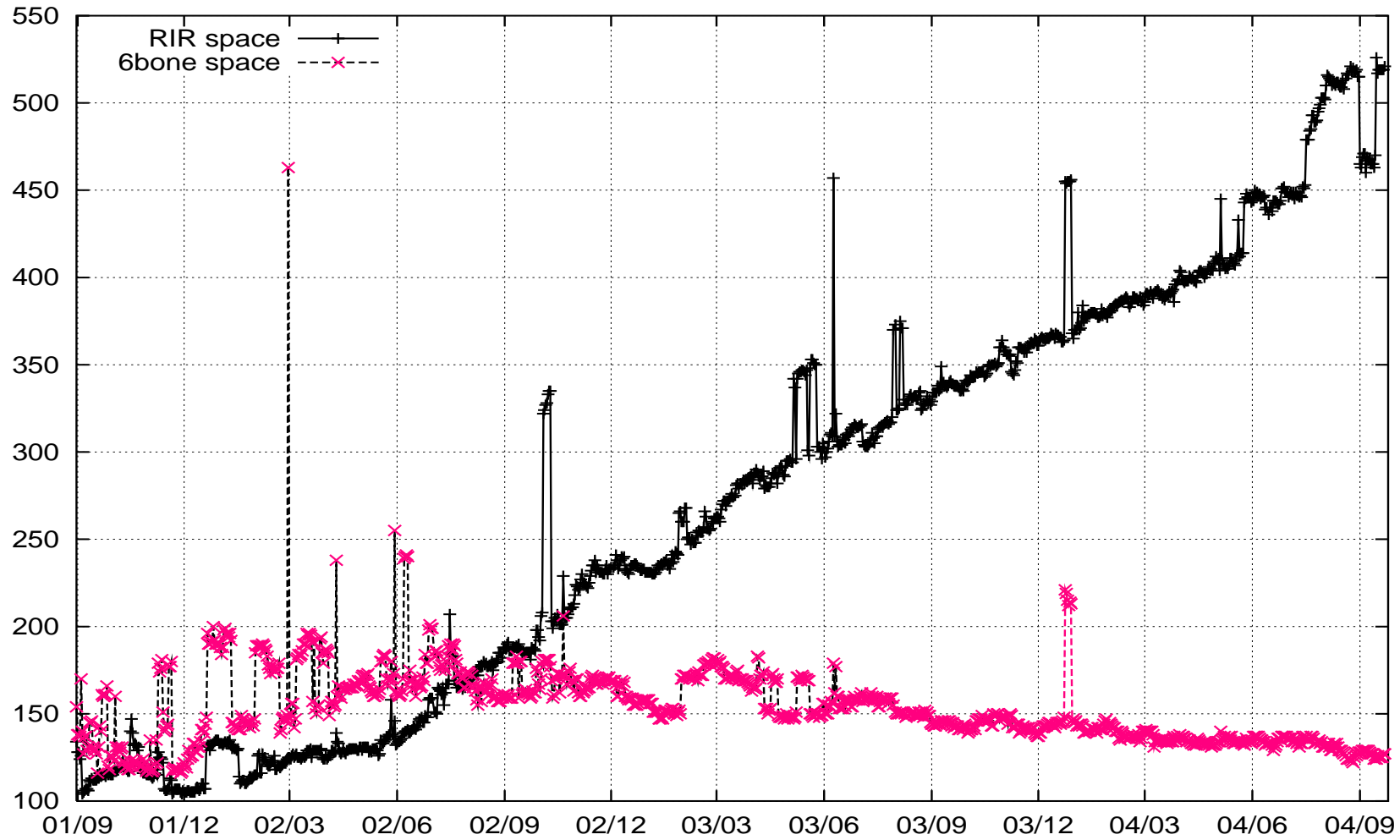
	Network	Next Hop	Path
*	2002::/16	2001:608:0:3::9	3320 1752 i
*		2001:948:0:F00F::1	2603 1741 i
*		3FFE:C00:8023:19::1	109 i
*		3FFE:8150:0:1::17	9044 559 i
*		2001:608:0:3::7	1930 i
*		3FFE:1108:40A:FFFF::1:2	3274 1741 i
*		2001:7F8:2:8001::2	1752 i
*		2001:7F8::3349:0:1	13129 9033 i
*		2001:7F8::CB9:0:1	3257 12859 ?
*		2001:450:1:2001::AA	3549 20965 559 i

- this is **fine**, anycast relay approach, see RFC3068
- this is just a snap-shot from AS5539, not all relays visible
- some research on non-publically visible 6to4 relays by David Malone (dwmalone@maths.tcd.ie): approximately 42 relays found. Good start, but more relays would be useful.
- some more-specific pfxs seen (prohibited by RFC3056 5.2.3!)

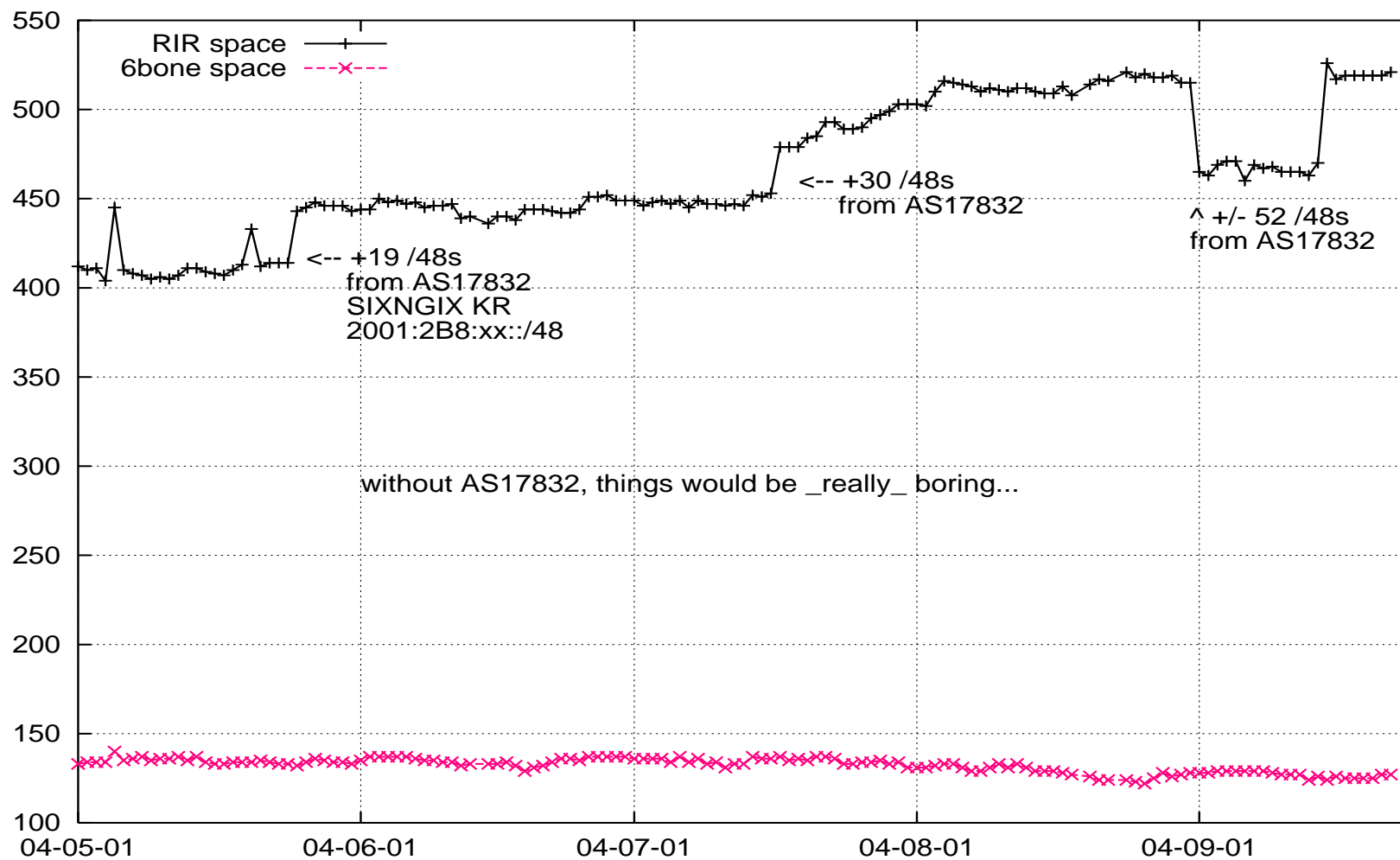
Graphics: Total Prefixes - 36 months



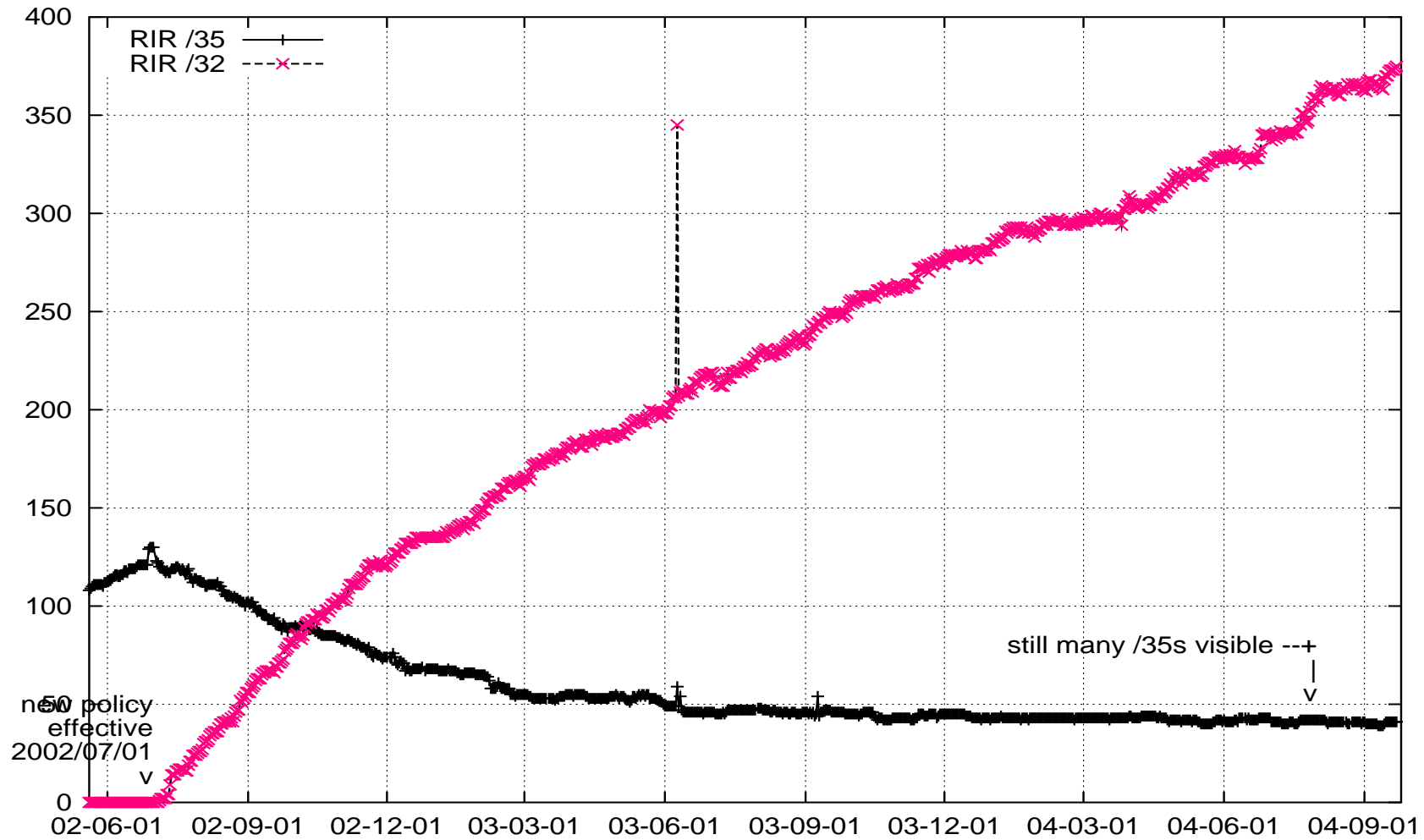
Graphics: RIR vs. 6Bone Prefixes - 36 months



Graphics: RIR vs. 6Bone Prefixes - 4 months



Graphics: RIR /35s vs. /32s



Why /35s?

- Non-upgraded /35 allocations (*only* /35 seen)?

APNIC: 7, ARIN: 4, RIPE: 4 prefixes - total: 15

- dual-announcements /32+/35 (same origin) - fear of ghosts?

APNIC: 7, ARIN: 6, RIPE: 3 prefixes - total: 16

- Traffic-Engineering with more-specific announcements??

2001:700::/32 3257 224 i

2001:700:E000::/35 3274 790 3549 6830 3320 1752 786 6680 2603 224 i

- Multihomed customers? Shared prefixes? (8)

2001:798::/32 3549 20965 i (GEANT)

2001:798::/35 680 6680 i (6NET)

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 2004/09/20, 690 LIR blocks (2001::/16) allocated by RIRs:

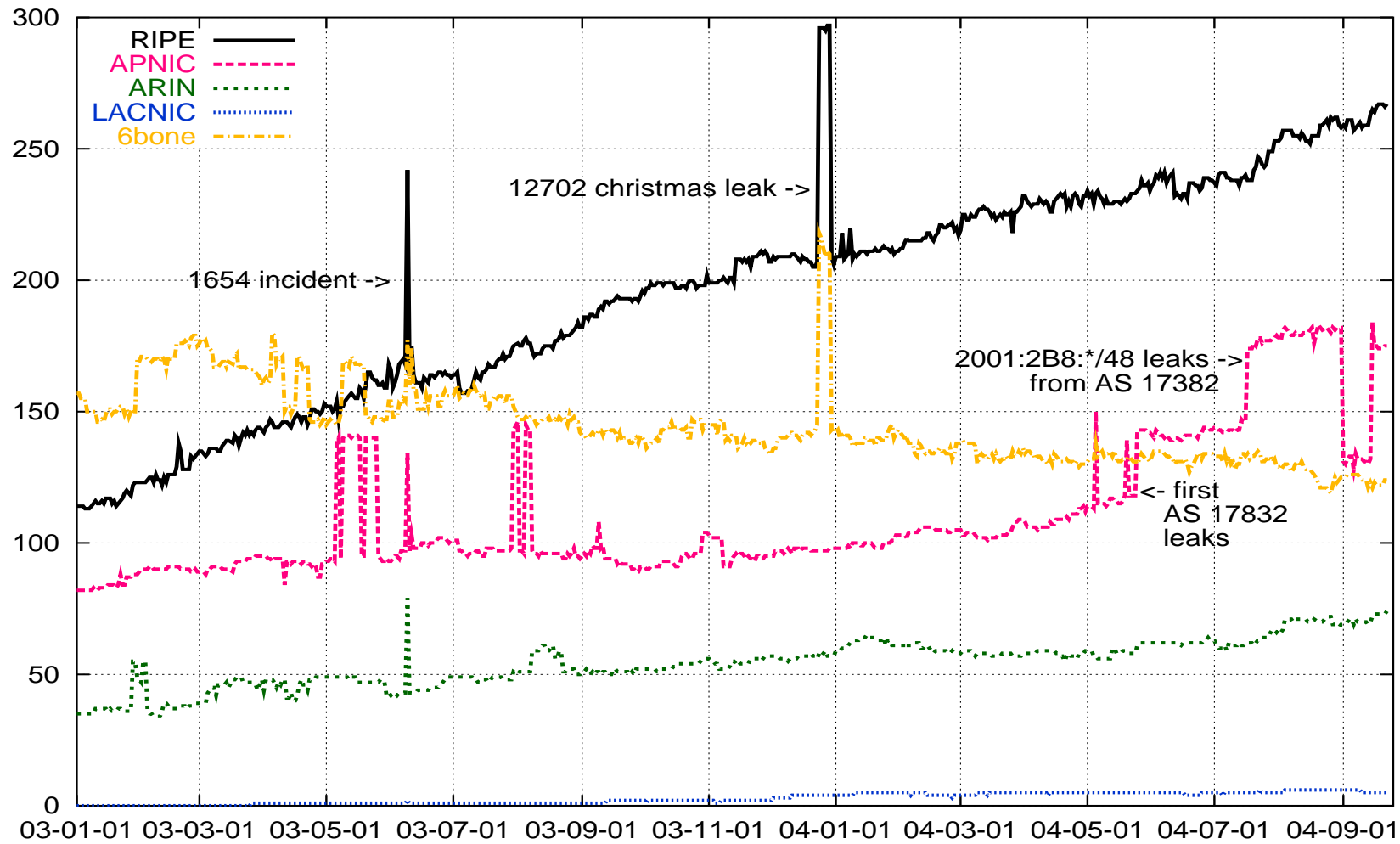
RIR	alloc.	members	perc.	on 2004/05/03
ARIN	118	2099	5.7%	107 (+10%)
APNIC	170	~ 1500	11.3%	152 (+12%)
RIPE	392	~ 3800	10.3%	330 (+18%)
LACNIC	10	~ 231	4.3%	6 (+66%)

- actual *percentage* with IPv6 similar for RIPE and APNIC
- B, F, H, I, K, M Root-Servers have IPv6 addresses (some are visible on www.root-servers.org, some are already in BGP)
- 392(344) allocations visible in routing table (as per GRH)

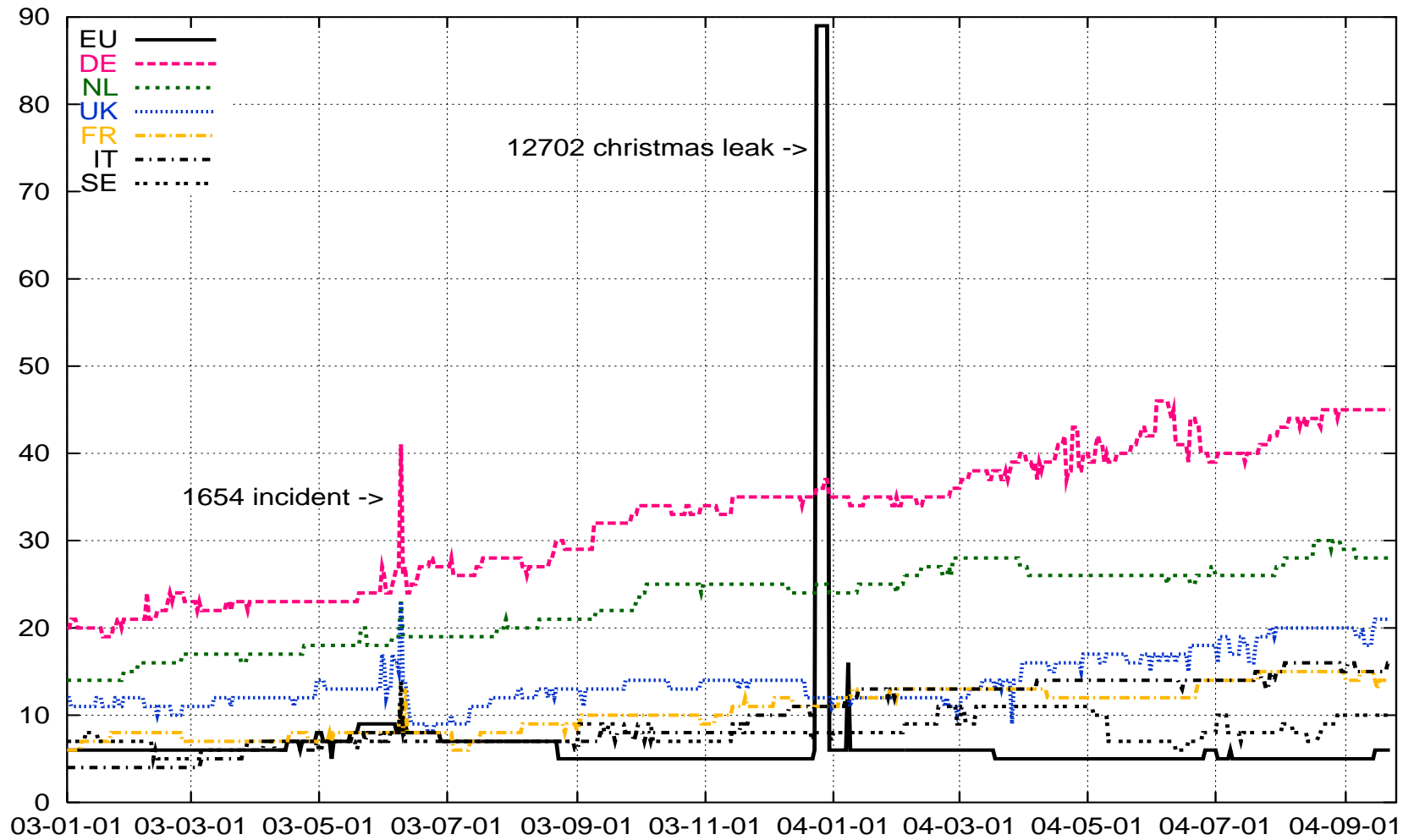
Numbers: RIRs: notable allocations

- more “very large” allocations seen:
 - 2001:1C00::/23 to NL-BENELUX (2004/05/10)
 - 2001:2000::/20 to EU-TELIANET (2004/05/10)
 - 2001:0F60::/28 to VECTANTNET-JP (2004/08/23)
 - 2001:0628::/32 extended to /30 (AT-ACONET)
 - 2001:5000::/21 to Cable&Wireless Global (2004/09/10)
- new 2001:40xx::/23 blocks from ICANN ⇒
 - 2001:4000::/23 to RIPE NCC (2004/07)
 - 2001:4200::/23 to ARIN (2004/07)
 - 2001:4400::/23 to APNIC (2004/07)
 - 2001:4600::/23 to RIPE NCC (2004/08)
 - 2001:4800::/23 to ARIN (2004/08)

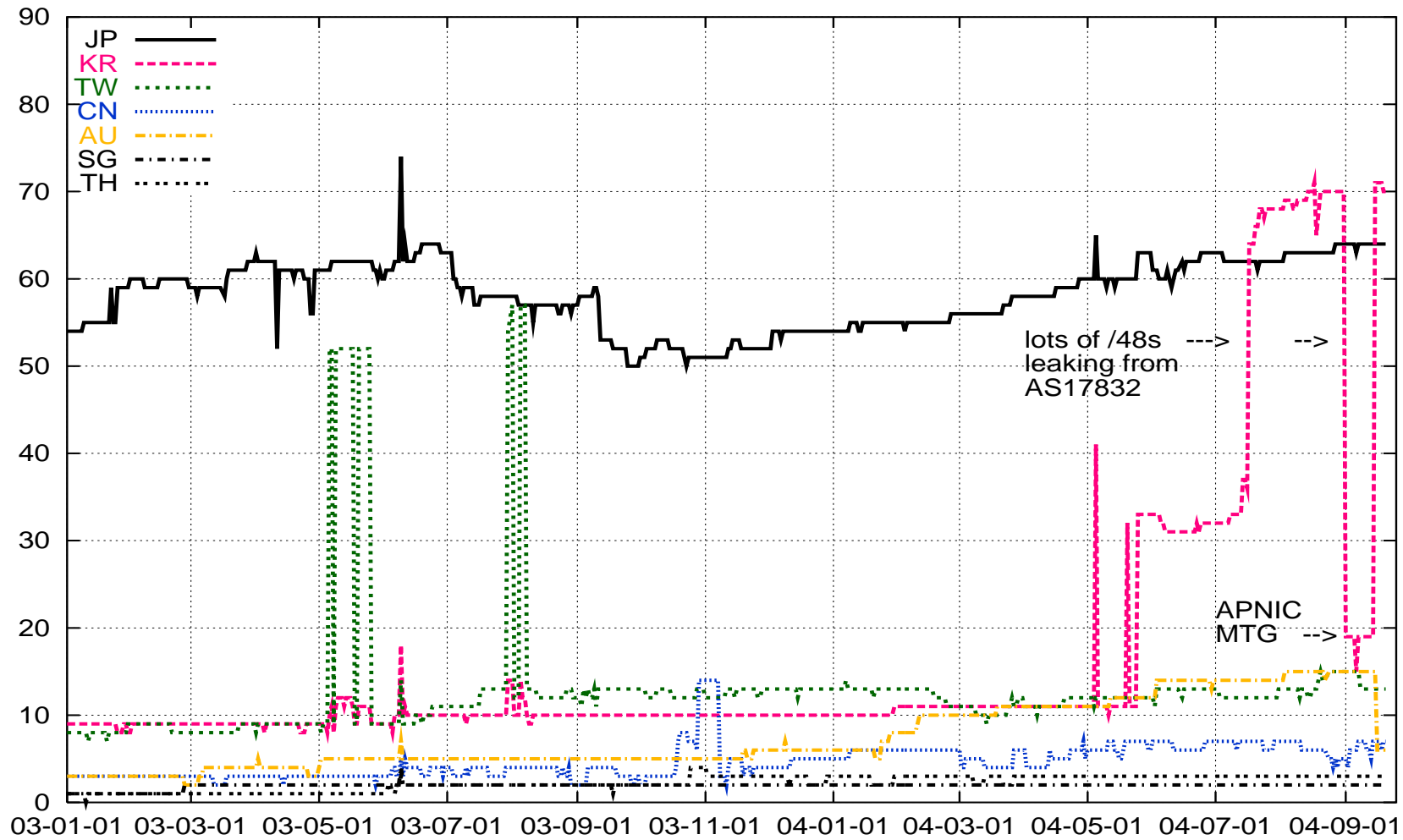
Graphics: prefixes by RIR region



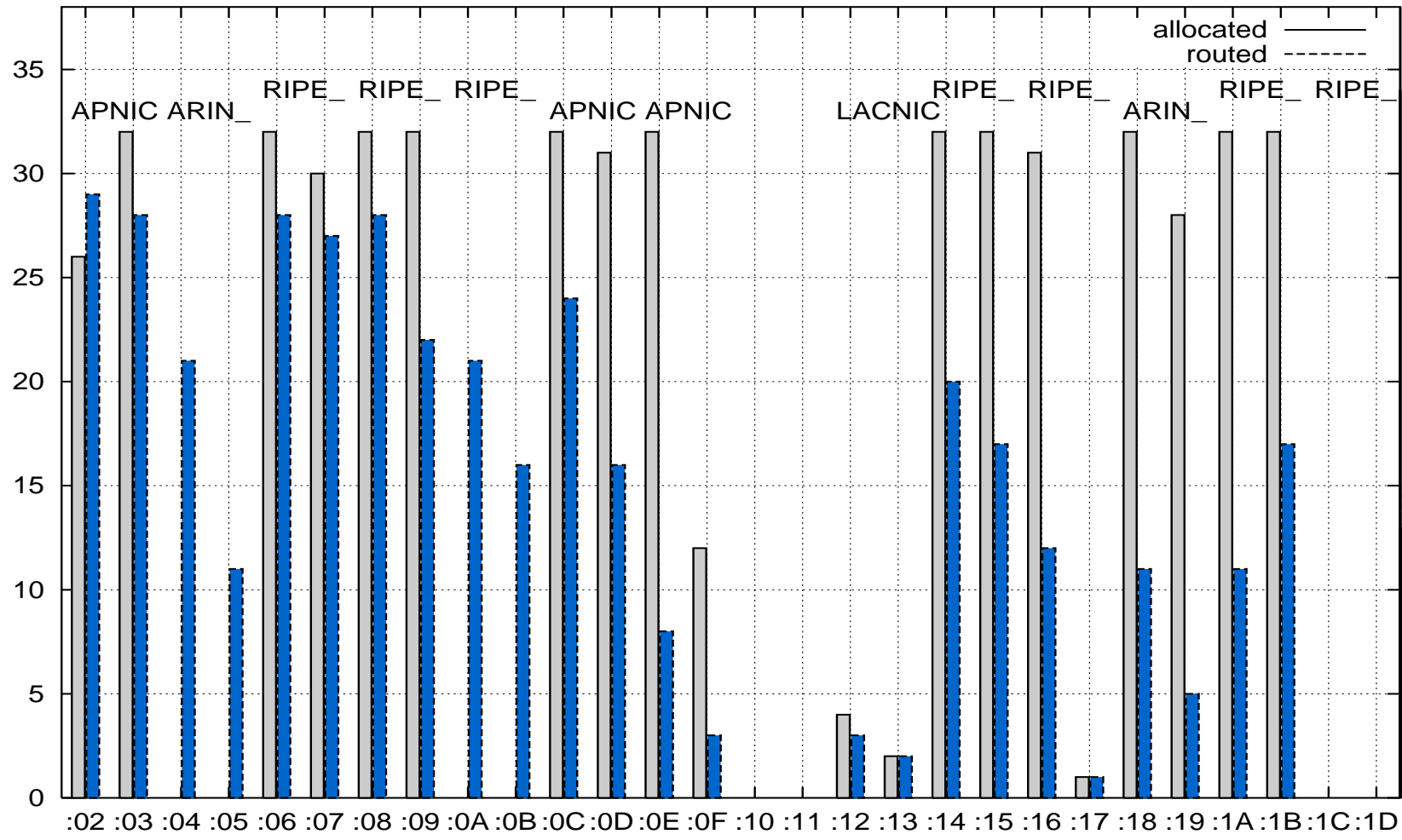
Graphics: prefixes by country (RIPE)



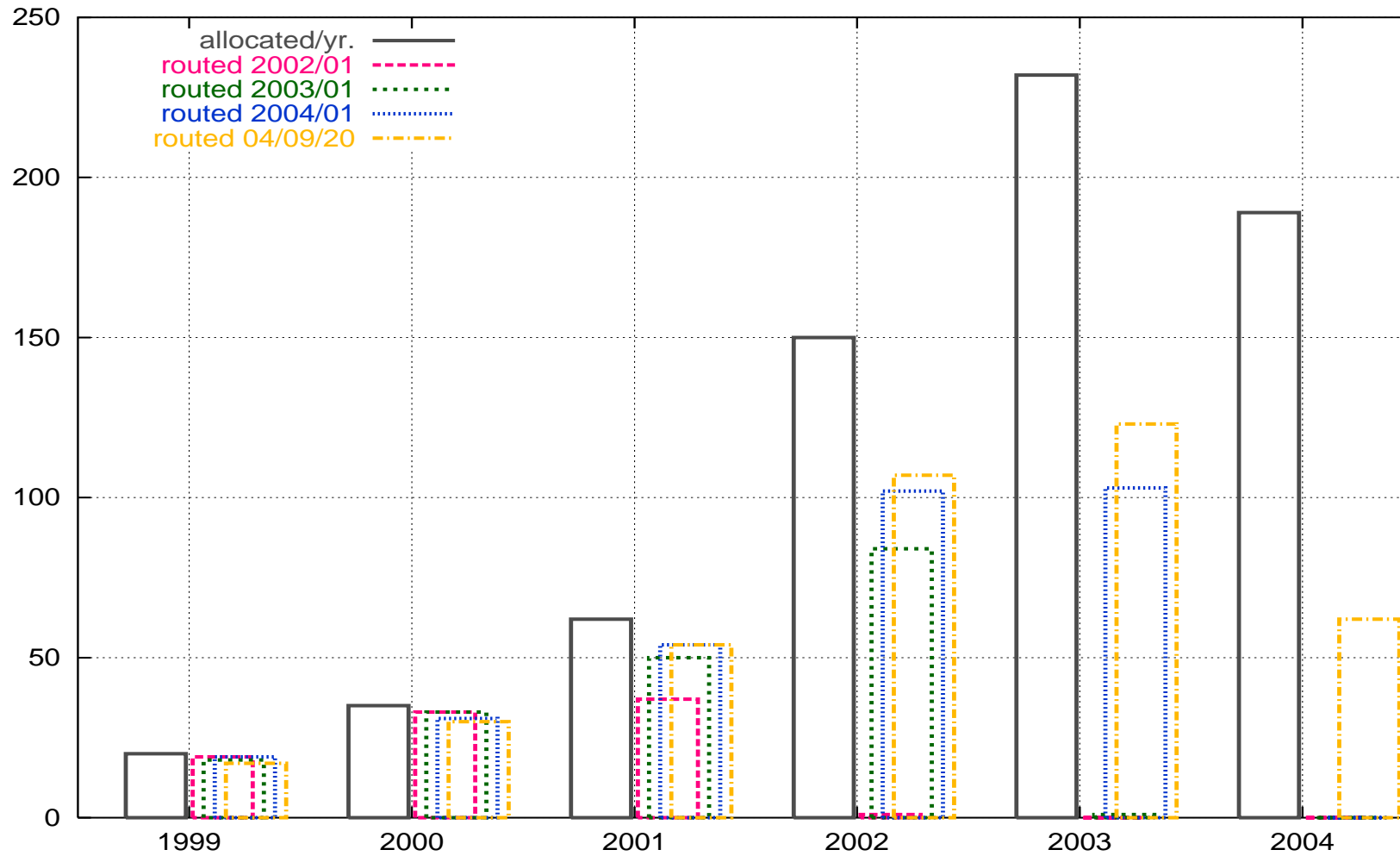
Graphics: prefixes by country (APNIC)



Graphics: Allocated vs. Routed (old)



Graphics: Allocated vs. Routed (new)



Interesting Observations (1) - Ghost Busting

```

Network      Path
*> 2001:200::/32  3549 2500 i
*              1752 4725 2500 i
*              1930 20965 11537 2500 i

* 3FFE:3500::/24 1221 109 109 10566 9044
                    5424 10318 6175 13944 3344 25396 25396 15703 12731 8319 i
*              3320 1275
                    5424 10318 6175 13944 3344 25396 25396 15703 12731 8319 i
*              1930 1930 1930 1930 1930 10566 9044
                    5424 10318 6175 13944 3344 25396 25396 15703 12731 8319 i
*              6939 6939
                    5424 10318 6175 13944 3344 25396 25396 15703 12731 8319 i
*              3274 790 3549 6939 6939
                    5424 10318 6175 13944 3344 25396 25396 15703 12731 8319 i
*              1752 12853 1275
                    5424 10318 6175 13944 3344 25396 25396 15703 12731 8319 i
*              109 10566 9044
                    5424 10318 6175 13944 3344 25396 25396 15703 12731 8319 i
*>              9044
                    5424 10318 6175 13944 3344 25396 25396 15703 12731 8319 i

```

Ghosts = BGP withdrawal bug, caused by old and buggy software.
 Long paths can stay *mostly unchanged* in the table for weeks.
 Don't confuse with BGP count-to-infinity (= paths change quickly).

Interesting Observations (2) - Accidental Hijack

Network	Next Hop	Path
* 2001:6c8::/32	3ffe:8150::1	9044 10566 29657 ?
*> 2001:6c8::/32 >	3ffe:401c:0:3:20c:ceff:fe05:da0e	29657 ?
* 2001:6c8::/32	2001:468:ff:121d::2	11537 3425 293 6175 3292 i
* 2001:6c8::/32	2001:770:8::	1213 11537 3425 293 6175 3292 i
* 2001:6c8::/32	2001:1548:1:10::4	12565 5609 4555 6830 3292 i
* 2001:6c8::/32	2001:15a8:1:1::6	29449 3320 6830 3292 i
* 2001:6c8::/32	2001:15f8:1::1	25384 1752 8472 6830 3292 i
* 2001:6c8::/32	3ffe:4017::1	24776 25358 3549 6830 3292 i

- correct origin AS is 3292
- caused by static route and static → BGP redistribution
- fixed in less than 20 hrs - thanks!

Interesting Observations (3) - Martian Leaks

Network	Next Hop	Path
*> ::/0	2001:7F8:1::A500:1103:1	1103 i
...		
*> 1000::/8	2001:610:25:5062::62	1103 i
* 1000::/8	2001:668:0:1:34:49:6900:40	3257 1103 i
* 1000::/8	2001:728:0:1000::f000	33 2914 6680 1103 i
* 1000::/8	2001:608:0:fff::6	399 5539 3257 1103 i
* 1000::/8	2001:780:0:2::6	12337 3257 1103 i
...		
* 1000:2550:100::/40	2001:728:0:1000::f000	33 2914 6680 1103 i
* 1000:2550:200::/40	2001:1418:1:400::1	12779 3549 6939 3257 1103 i
* 1000:2550:300::/40	2001:15a8:1:1::6	29449 6939 3257 1103 i
* 1000:2550:400::/40	2001:608:0:fff::6	399 5539 3257 1103 i

- likely to be caused by some Cisco/6PE weirdness
- noticed and fixed in very short time (thanks!)
- only documented martian leak since 2002/10/21 :-)
- but: displays potential for improving BGP prefix filters...

Interesting Observations (4) - Invalid AS numbers

Network	Next Hop	Path
*> 3FFE:1CE1::/48	2001:478:FFFF::1	4555 64555 65000 i (08/09)
*> 3FFE:1CE1::/48	2001:478:FFFF::1	4555 64555 3 i (08/12)

- private AS numbers should not be announced world wide
- long-standing offenders 64702 and 64600 gone since 05/13 :-)
- only one AS paths remaining with private ASes - good news!
- example above looks like accidental leak of customer pfx

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!*

News (?)

- 6bone (3FFE:...) going away, end date: 2006/06/06
- people start discontinuing *.ip6.int reverse delegation (proposal to switch off ip6.int globally on 2004/09/09 not accepted)
- private/unallocated AS numbers + ghosts seem under control
- still quite a number of “unsolicited full transit” links
- but: 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
- ...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/R49-v6-table/>

Questions?

`gert@space.net`