

What comes next? 1000 Prefixes reached!

A snapshot of the global IPv6 routing table

Gert Döring, SpaceNet AG, Munich, Germany

Oct 30, 2008

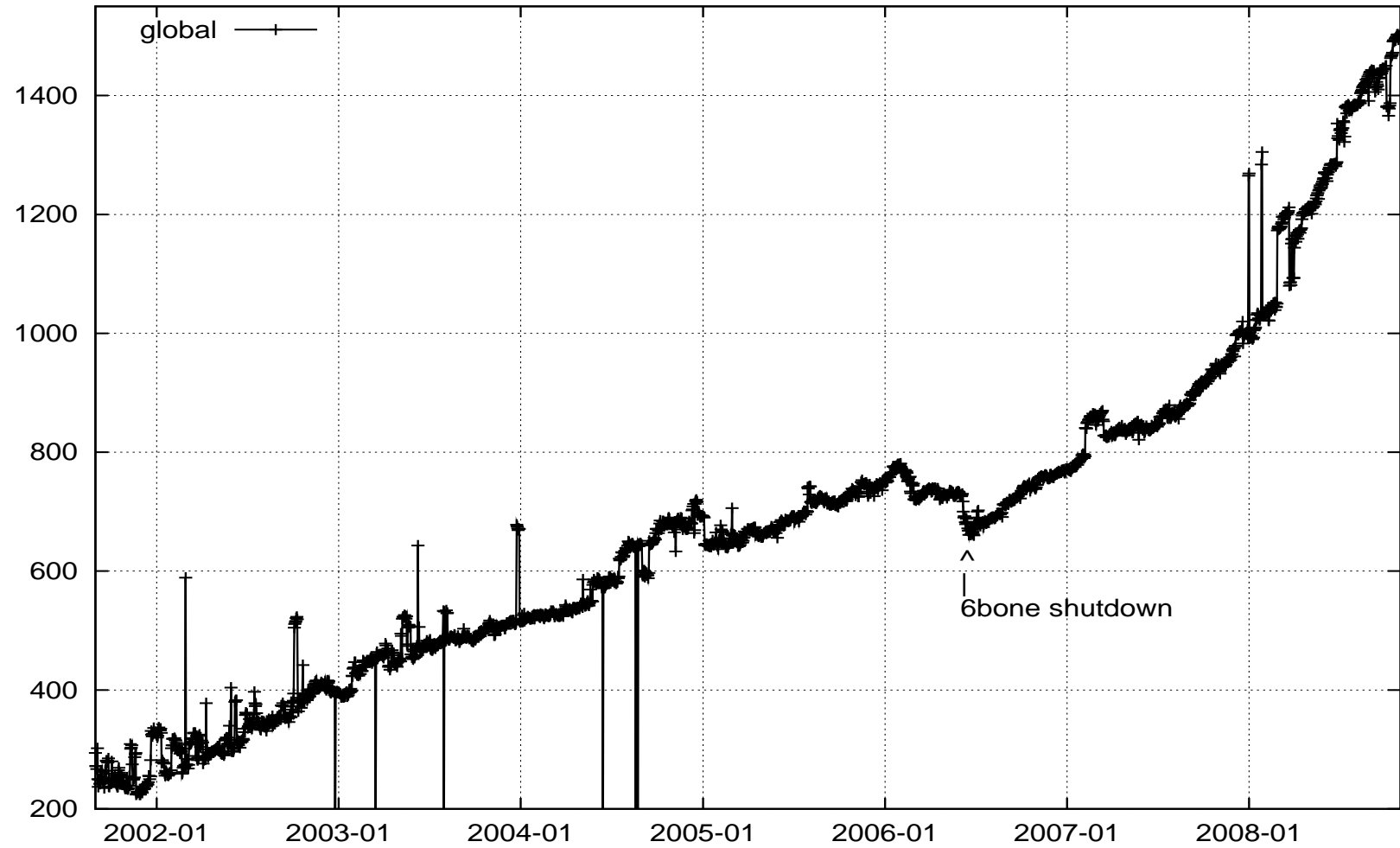
RIPE 57, Dubai

Overview

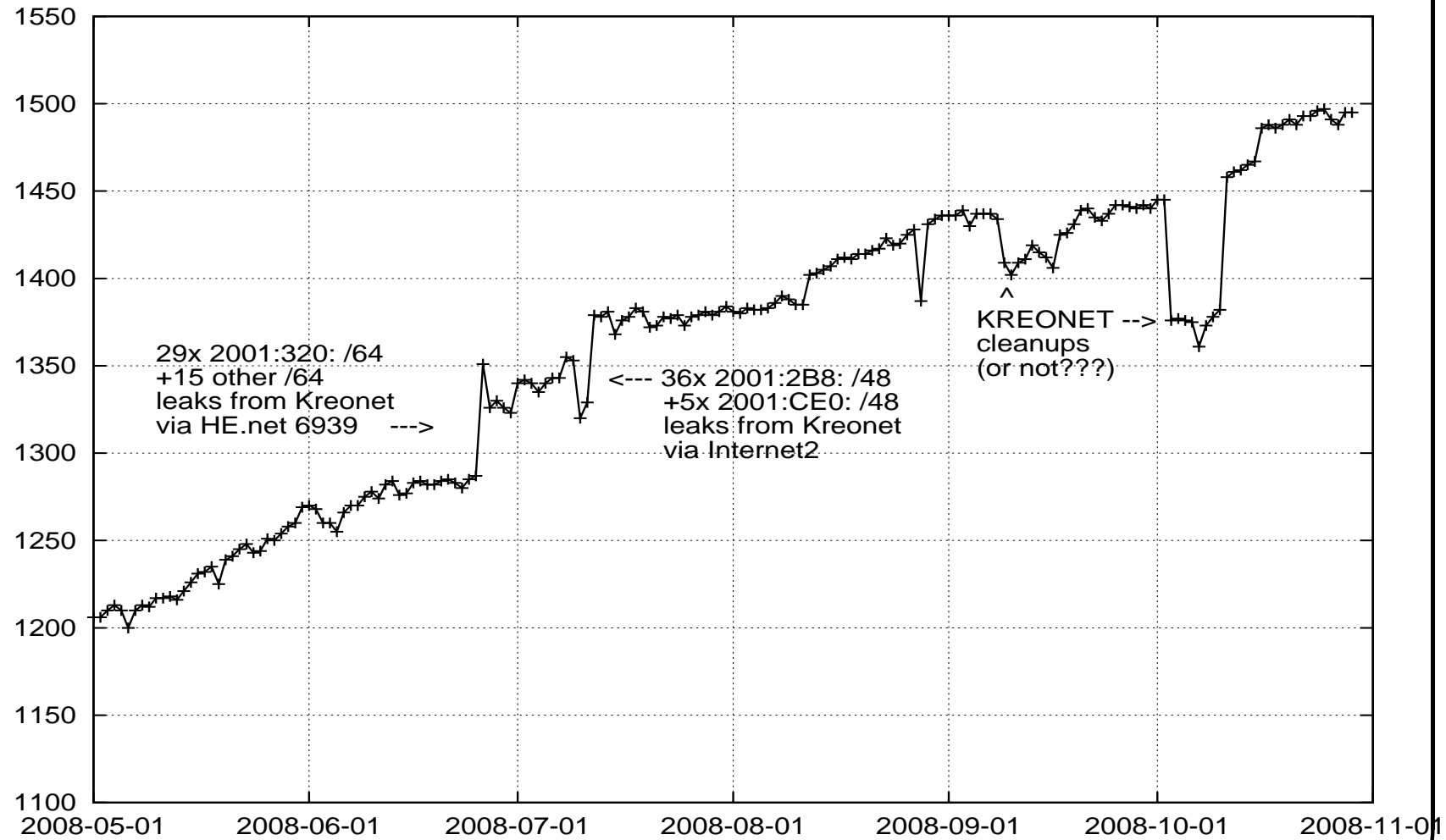
- pictures & trends
- numbers...
- things that should not be there...
- route6 current practices
- references

Slides online at: <http://www.space.net/~gert/RIPE/R57-v6-table/>

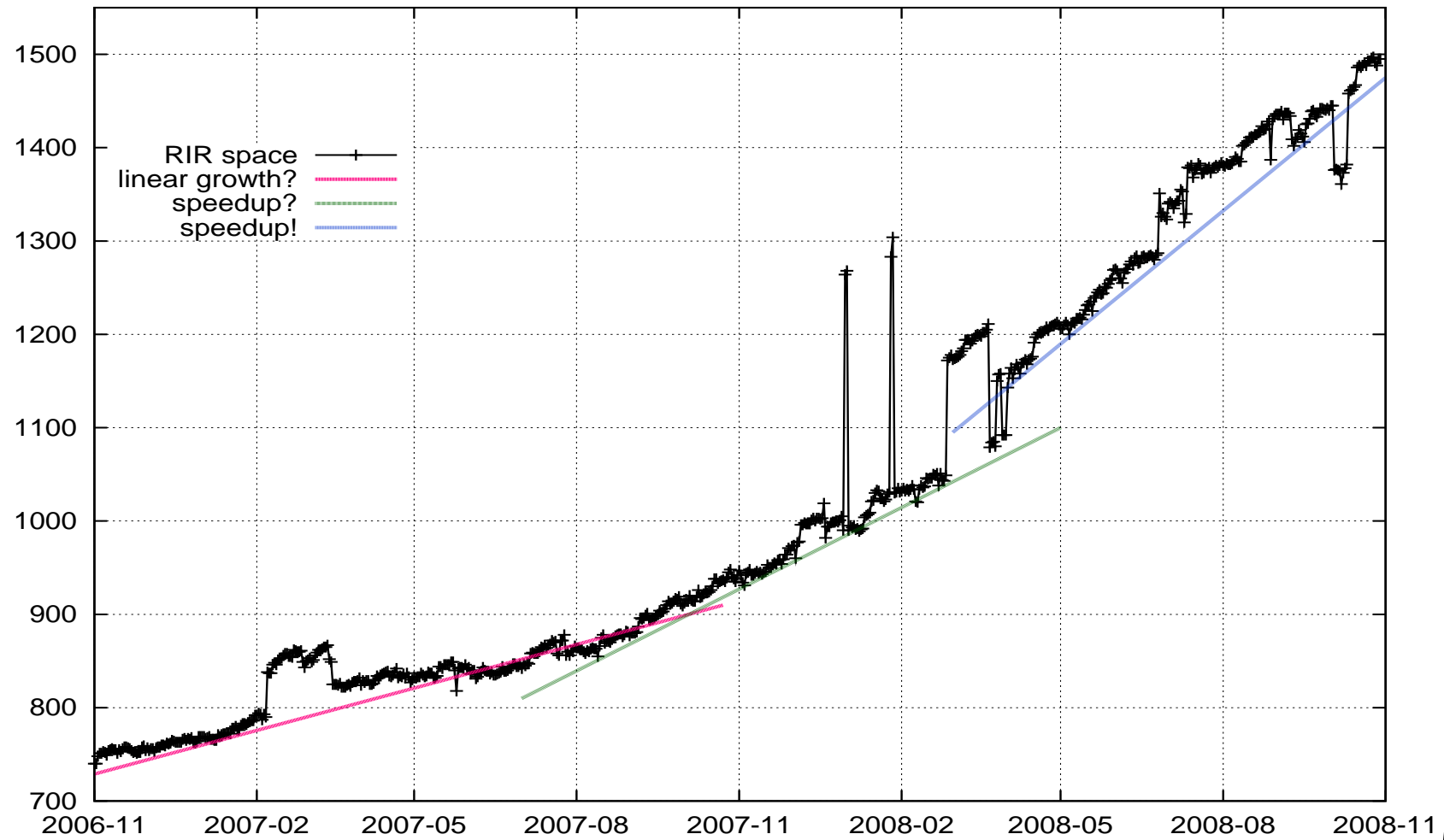
Graphics: Total Prefixes - 7 years



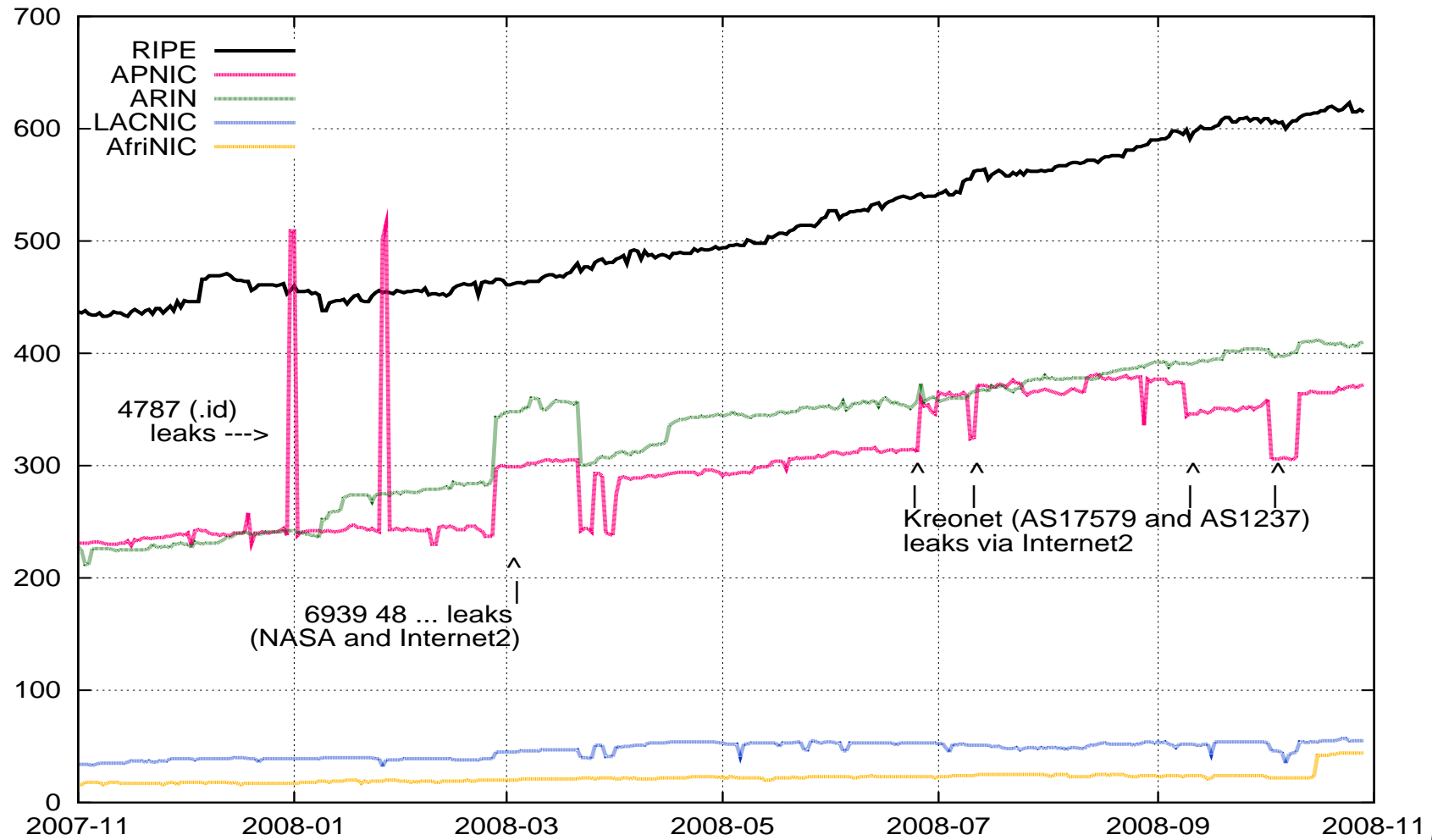
Graphics: zoom into last 6 months



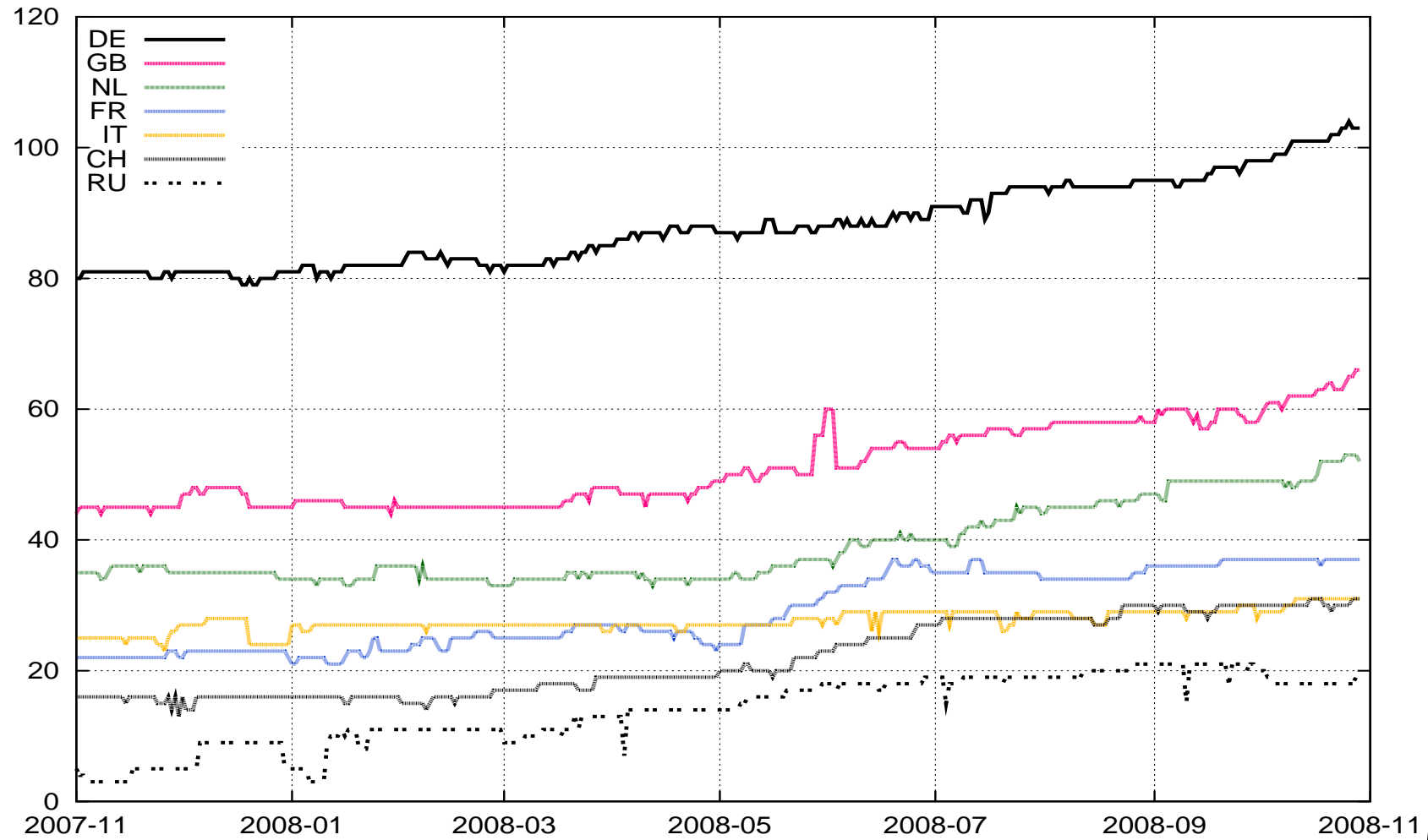
Graphics: trends? (24 months)



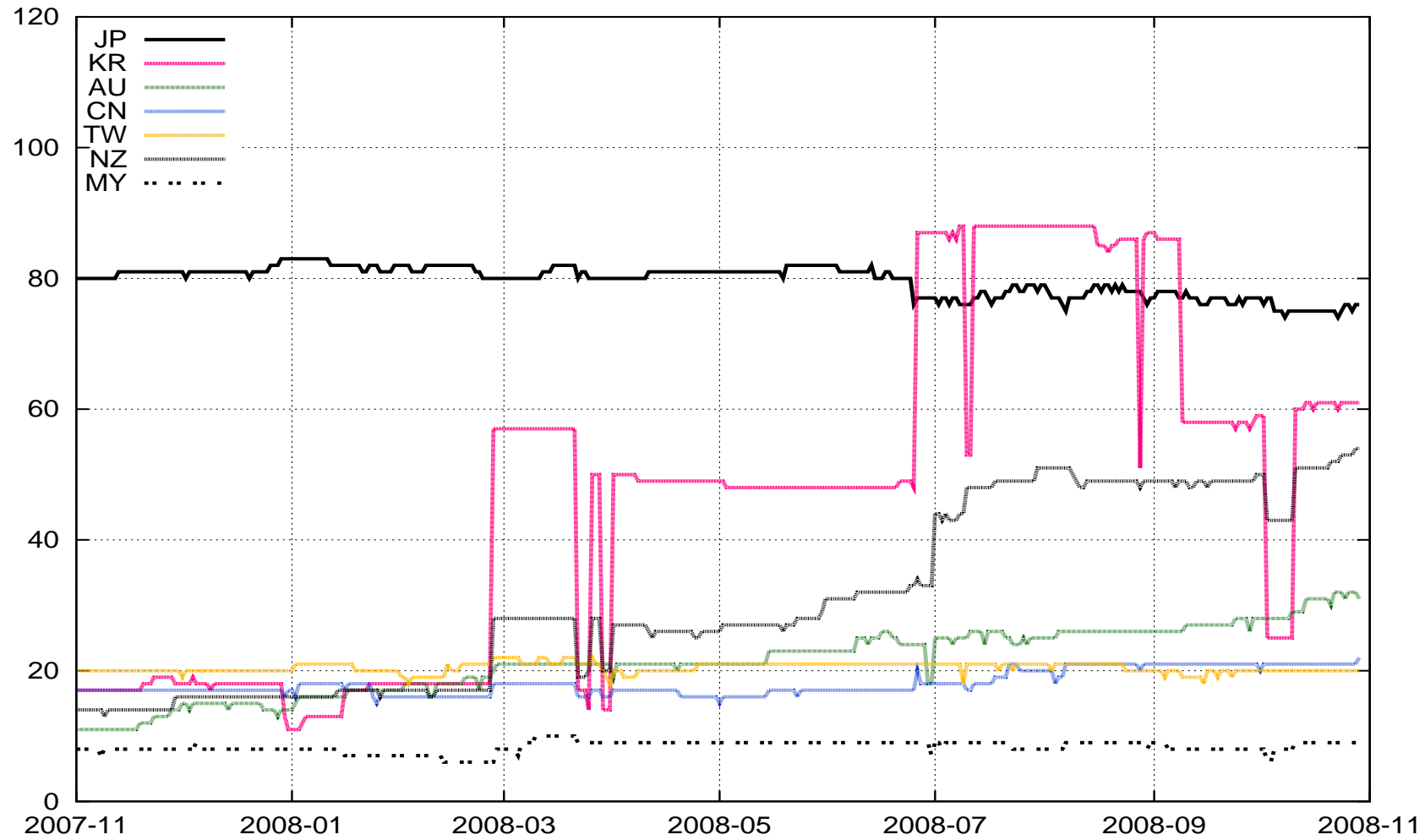
Graphics: prefixes by RIR region



Graphics: prefixes by country (RIPE)



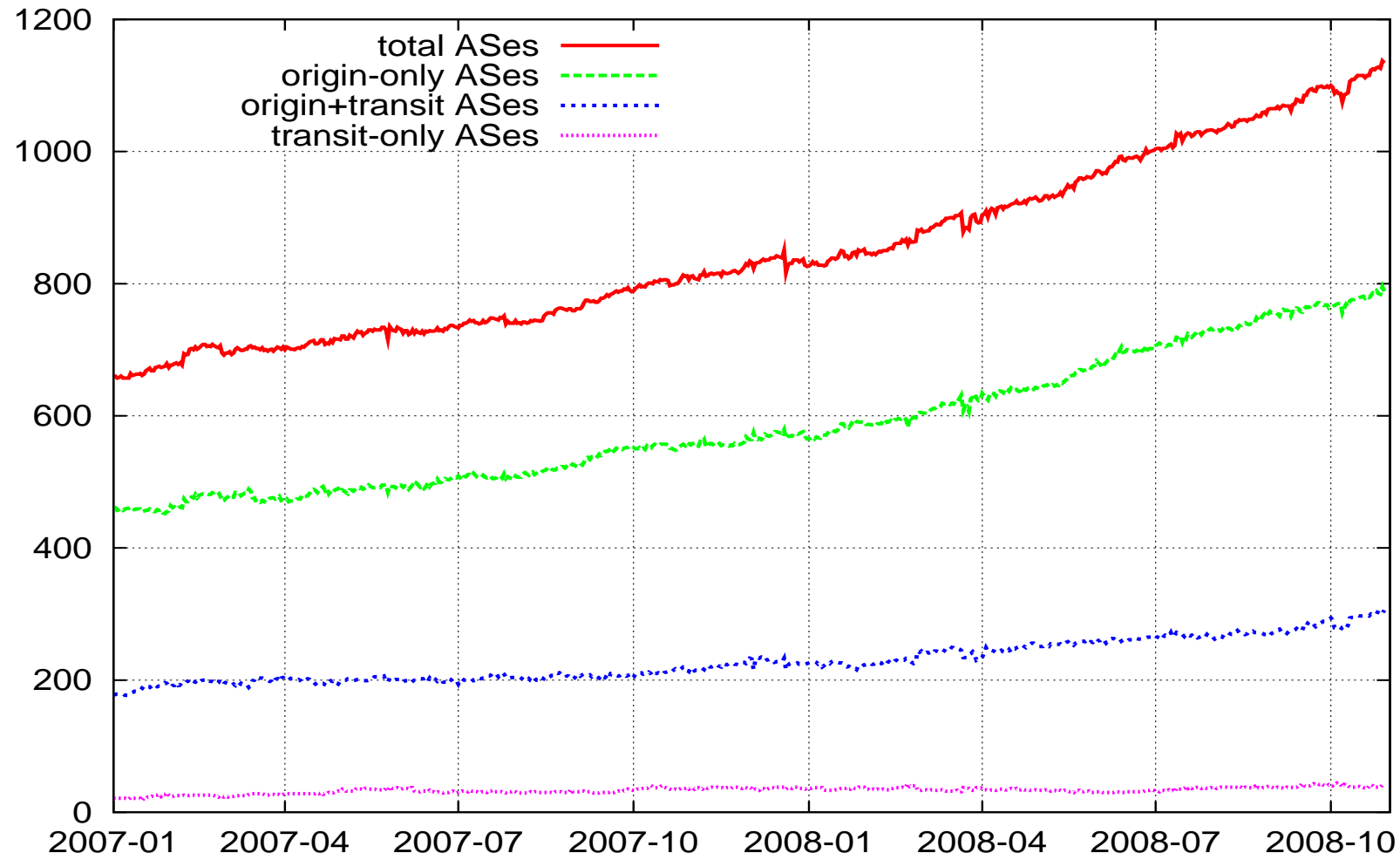
Graphics: prefixes by country (APNIC)



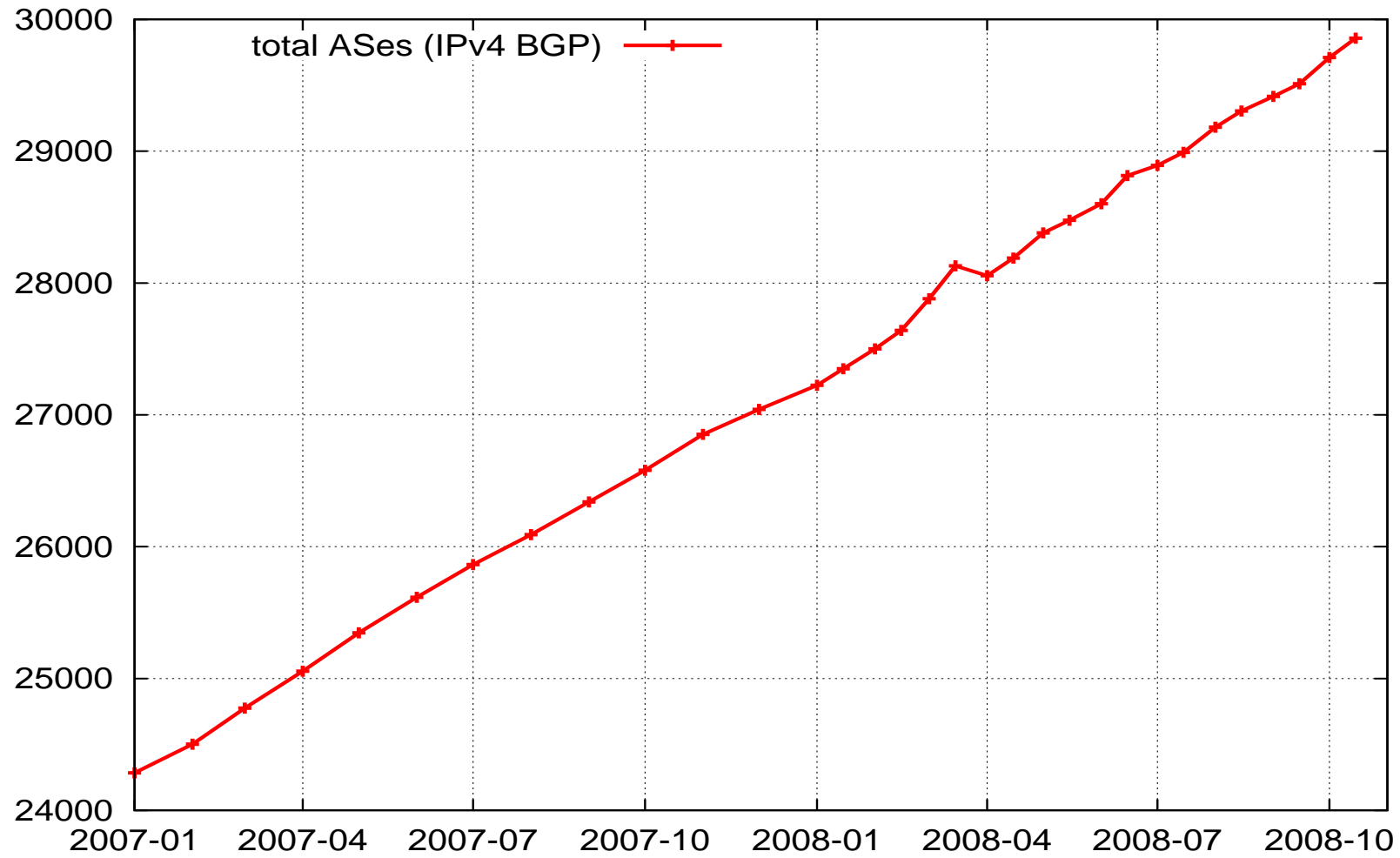
Numbers - AS numbers

- as of 2008-10-29: 1137 unique AS#s visible (2008-05-04: 933)
 - 795 origin-only ASes (no transit paths seen) (646)
 - 301 ASes originate & give transit (253)
 - 41 transit-only ASes (e.g. 57, 1125, 1659, 6667, ...) (34)
- different number of prefixes announced
 - 936 ASes originate 1 prefix (775)
 - 97 ASes originate 2 prefixes (2 due to /32+/35)
 - 32 ASes originate 3 prefixes
 - 12 ASes originate 4 prefixes
 - 18 ASes with “more than that”, max. is 30 & 35 prefixes
- 2 ASes still announce their prefix as /32 and /35
- note: all paths observed from AS5539

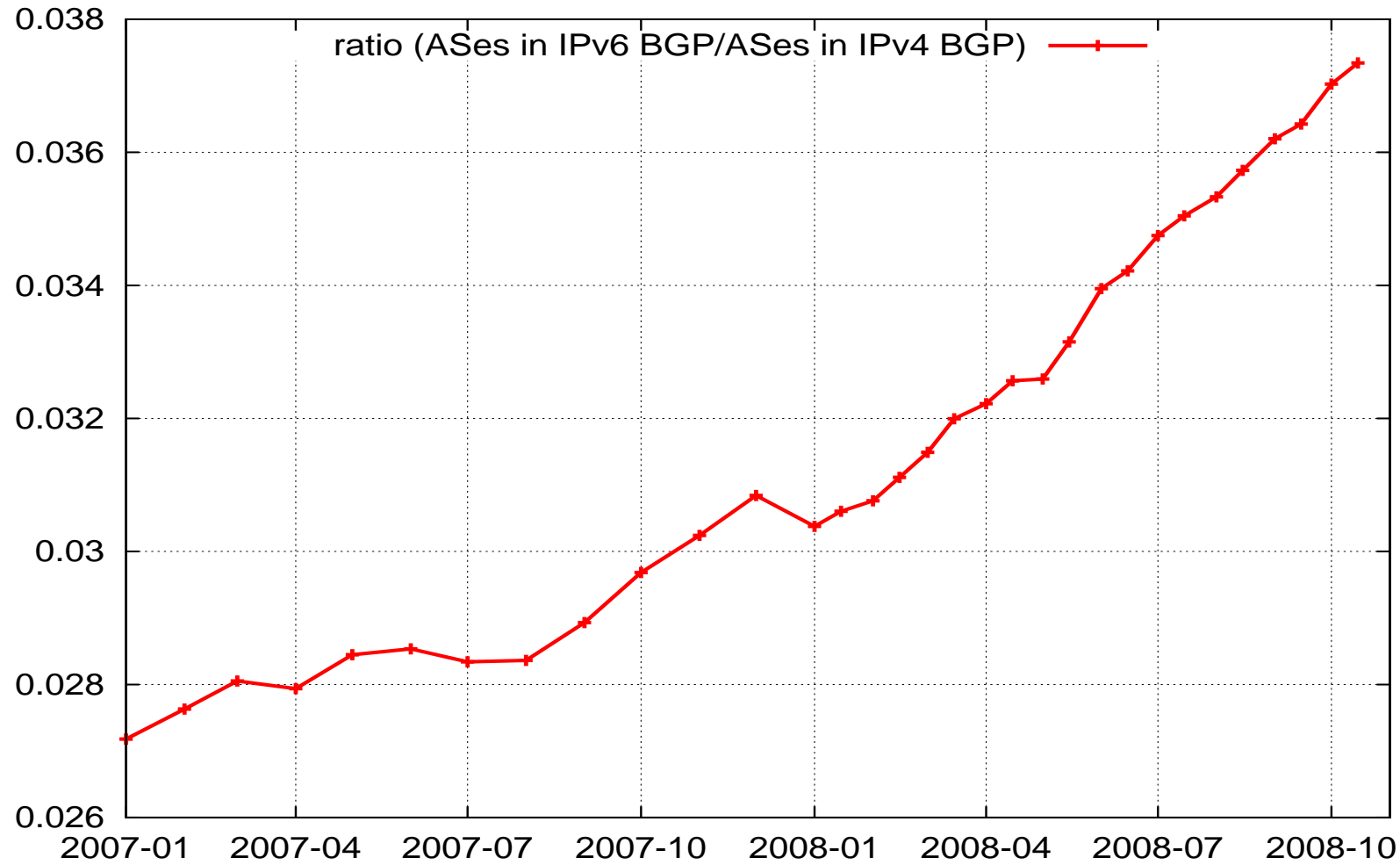
Graphics: AS Numbers (v6 BGP)



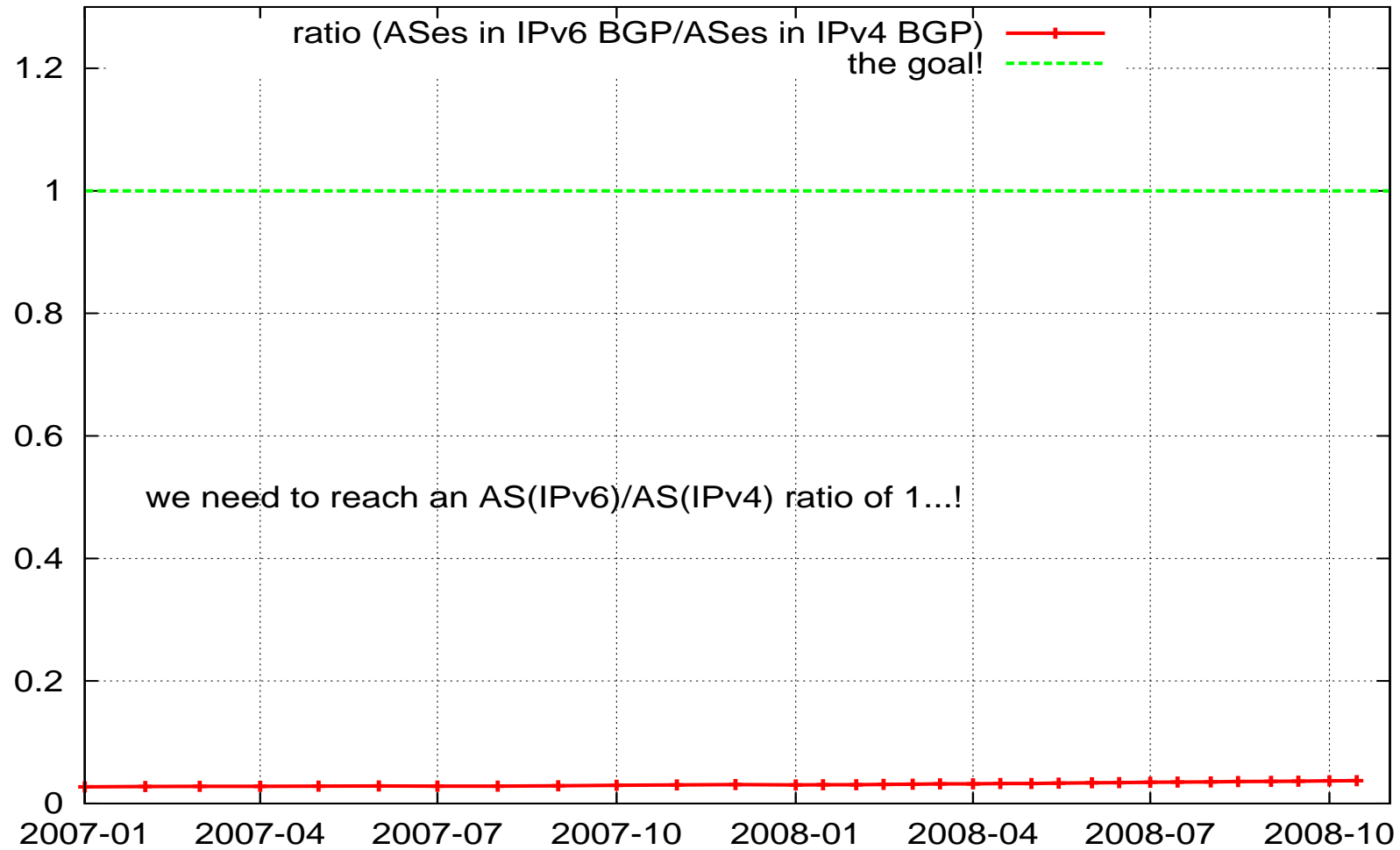
Graphics: AS Numbers (v4 BGP)



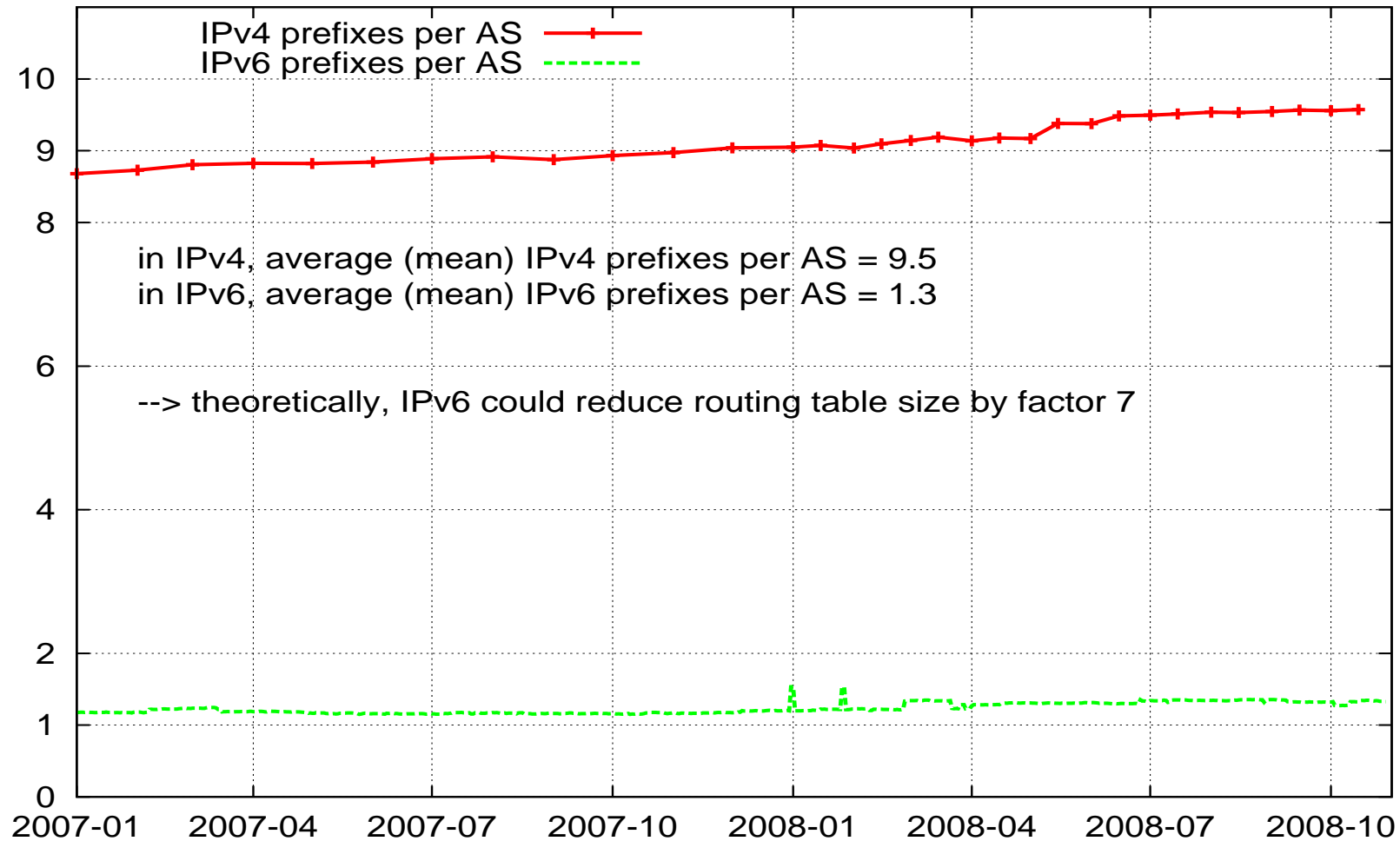
Graphics: AS Number Ratio (v6 BGP/v4 BGP)



Graphics: AS Number Ratio (v6 BGP/v4 BGP)



Graphics: Prefixes per AS (v4+v6)



ASes - why are people announcing 2+ prefixes

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

2001:420::/35	109 i
2001:420::/32	109 i

- ISP/LIR address space plus IXP prefixes

2001:5000::/21	1273 i	(C&W LIR space)
2001:7F8:2B::/48	1273 i	(IXP: INXS HAM)
2001:7F8:2C::/48	1273 i	(IXP: INXS MUC)

- mergers and acquisitions, business units, customer pfxs, ...

2001:218::/32	2914 i	NTT JP
2001:418::/32	2914 i	NTT America
2001:500:13::/48	2914 i	ARIN PI
2001:728::/32	2914 i	Verio Europe
2406:A000::/32	2914 i	Sempernet V6 Australia
2610:F8:8000::/35	2914 i	Command Information Inc.

- networks with multiple sites and multiple PI prefixes

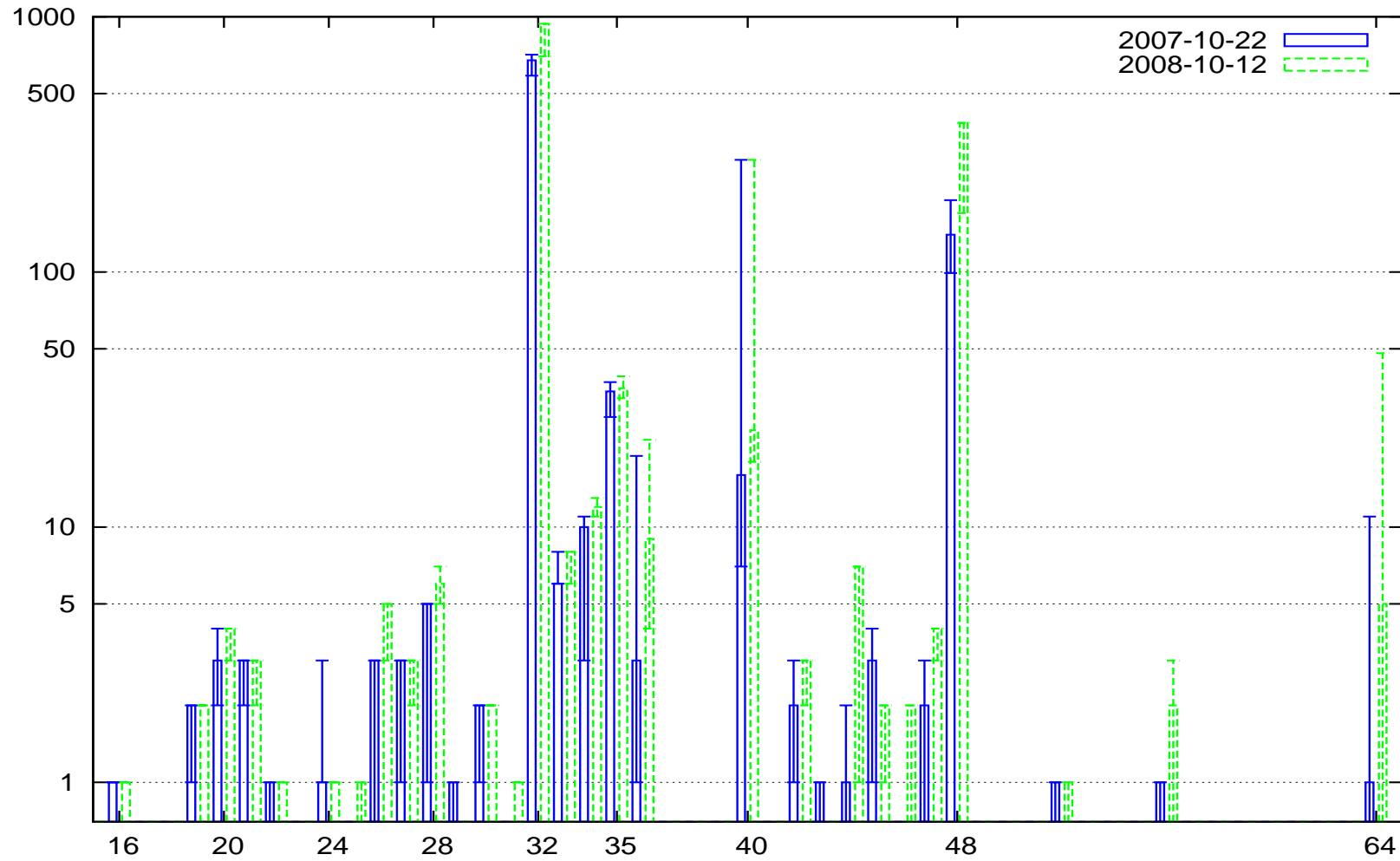
2001:500:16::/48	3257 6453 12041 i	Afilias
2001:500:17::/48	12041 i	Afilias
2001:500:18::/48	12041 i	Afilias
2001:500:19::/48	2914 12041 i	Afilias
2001:500:1A::/48	3257 6453 12041 i	Afilias
2001:500:1B::/48	12041 i	Afilias

Numbers - Prefixes

As of 2008-10-28: 1501 prefixes in total (2008-05-04: 1214)

/n	global	RIPE	APNIC	ARIN	Lacn.	Afri.	oth
/16	1	0	0	0	0	0	1
/19	2	2	0	0	0	0	0
/20..23	8	4	4	0	0	0	0
/24..27	10	5	4	1	0	0	0
/28..31	9	2	6	0	1	0	0
/32	956	513	198	184	37	18	6
/33..34	20	7	5	8	0	0	0
/35	35	8	18	9	0	0	0
/36	9	3	0	5	1	0	0
/40	26	8	5	9	2	2	0
/42	4	3	0	1	0	0	0
/44..47	15	3	1	11	0	0	0
/48	397	57	123	178	13	24	2
/49..63	2	0	1	1	0	0	0
/64..128	6	0	4	0	1	0	1

Graphics - Prefixes



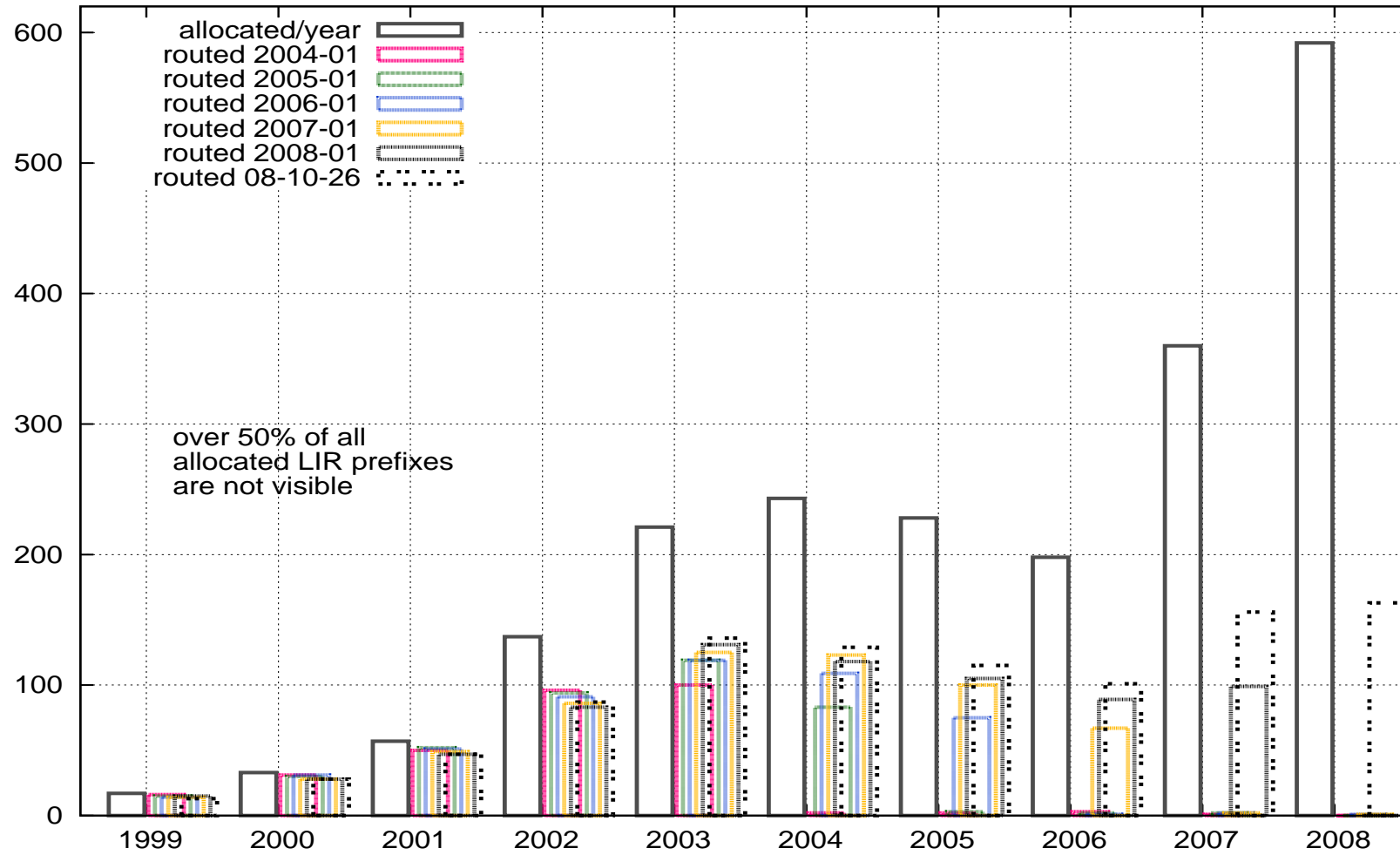
Numbers: RIRs, Allocations, ...

- On 2008-10-12, 2122 LIR blocks (2000:: $/4$) allocated by RIRs:

RIR	alloc.	members	perc.	on 2008-05-04
ARIN	487	~ 3200	15.2%	408 (+19%)
APNIC	430	~ 2863	15.0%	366 (+17%)
RIPE	1052	~ 5920	17.7%	878 (+20%)
LACNIC	101	~ 1006	10.1%	95 (+6%)
AfriNIC	52	~ 347	15.0%	43 (+21%)

- note: not counting $/48$ microallocs and $/35 \Rightarrow /32$ extentions
- actual *percentage* with IPv6 similar among regions
- 949 (R56: 799) allocations visible in routing table (*only 45%!*)

Graphics: Allocated vs. Routed



Allocated vs. Routed - by region & class

RIR	type	alloc.	visible	perc.	subnets
ARIN	LIR	459	183	40%	88
	IXP	22	0	0%	0
	Critical Inf.	52	23	42%	32
	Internal Inf.	3	0	0%	0
	PI	169	34	20%	48 (*)
APNIC	LIR	426	207	49%	129
	IXP	21	2	10%	0
	PI	38	6	16%	18 (*)
RIPE	LIR	1052	509	48%	101
	IXP	72	13	18%	0
	Anycast DNS	8	5	63%	0
LACNIC	LIR	100	33	33%	19
	Crit.Inf.+PI	8	2	25%	0
AfriNIC	LIR	50	17	34%	2
	PI	8	3	38%	0

Allocated vs. Routed - reasons?

- “early adopters” already losing interest in IPv6?
- “prepare for the future” allocations?
- “for internal use” allocations? (some, yes)
- distribution of non-announced prefixes does not show any specific characteristic, like “academia” vs. “commercial networks” etc.
- some delay between prefix allocation and announcement is to be expected (expect some more statistics in this space...)
 - but this cannot explain effects seen on 2003+2004 allocations – about 40% don’t show up after over 3 years...

Eastern Asian Research

```

Network          Path
* > 2001:2B8::/32 3257 2497 4725 6939 17832 i
* 2001:2B8:90::/48 109 5511 10764 11537 17579 1237 i
* 2001:2B8:94::/48 109 5511 10764 11537 17579 1237 i
* 2001:2B8:9A::/48 109 5511 10764 11537 17579 1237 i
* 2001:2B8:9C::/48 109 5511 10764 11537 17579 1237 i
...
* > 2001:3C8::/32 3257 2497 4725 2500 7660 24287 24490 24475 4621 i
* 2001:3C8:100D::/48 109 5511 10764 11537 4621 4621 4621 4621 4621 ?
* 2001:3C8:1202::/48 109 5511 10764 11537 4621 4621 4621 4621 4621 ?
* 2001:3C8:1303::/48 109 5511 10764 11537 4621 4621 4621 4621 4621 ?
* 2001:3C8:9007::/48 109 5511 10764 11537 4621 4621 4621 4621 4621 ?
...

```

- Internet2 (AS 11537) now leaks via NCSA and OpenTransit
- Korean RENs 1237 and 17579 nowadays *do* have proper transit!
- Thai REN 4621 has other transit, but still leaks via I2 :(
- long-standing problem: research networks with poor connectivity, due to political issues or plain disinterest

Egyptian Research

Network	Path
* 2001:4300::/32	6939 6175 i
*>	3257 6175 i
*> 2001:4300:2001::/48	109 5511 10764 11537 33789 24863 i
*> 2001:4300:2002::/48	109 5511 10764 11537 33789 24863 i
*> 2001:4300:2003::/48	109 5511 10764 11537 33789 24863 i
*> 2001:4300:2004::/48	109 5511 10764 11537 33789 24863 i
...	
*> 2001:4300:2009::/48	109 5511 10764 11537 33789 24863 i
*> 2001:4300:2010::/48	109 5511 10764 11537 33789 24863 i
...	
*> 2001:4300:2019::/48	109 5511 10764 11537 33789 24863 i
*> 2001:4300:2020::/48	109 5511 10764 11537 33789 24863 i

- 10764 = NCSA, 11537 = I2
33789 = Egyptian MCIT Internet2
24863 = LINKdotNET AS number
- can you recognize a pattern?
- (this *might* be legitimate, but there is no IRR data to check)

local-pref is evil (as is NREN policy)

- I2 leaks stuff. Geant2 local-prefs I2. DFN local-prefs Geant2.
- ... traffic goes from DE to DE via US&HK:

```
1 vl-23.csr1-2wr.lrz-muenchen.de (2001:4ca0:0:f000::1) 0.374 ms
2 xr-gar1-te1-3-108.x-win.dfn.de (2001:638:c:a003::1) 0.441 ms
3 2001:638:c:c043::2 (2001:638:c:c043::2) 8.484 ms
4 dfn.rt1.fra.de.geant2.net (2001:798:14:10aa::1) 7.879 ms
5 abilene-wash-gw.rt1.fra.de.geant2.net (2001:798:14:10aa::12) 100.643 ms
6 so-0-2-0.0.rtr.chic.net.internet2.edu (2001:468:ff:209::2) 117.230 ms
...
9 so-0-0-0.0.rtr.seat.net.internet2.edu (2001:468:ff:716::1) 181.476 ms
10 kreonet-1-lo-jmb-706.sttlwa.pacificwave.net (2001:504:b:10::6) 169.102 ms
11 2001:320:1b00:1::1 (2001:320:1b00:1::1) 283.341 ms
12 hurricaneelectric-RGE.hkix.net (2001:7fa:0:1::ca28:a19e) 327.876 ms
13 v1026.core1.sjc1.he.net (2001:470:0:c3::1) 331.108 ms
14 10gigabitethernet2-1.core1.sjc2.he.net (2001:470:0:55::2) 327.817 ms
...
18 10gigabitethernet1-1.core1.fra1.he.net (2001:470:0:47::2) 327.742 ms
19 1g-bge0.tserv6.fra1.ipv6.he.net (2001:470:0:69::2) 328.126 ms
```
- This Needs To Stop. Move to the 21st century!
- unconditionally local-pref'ing certain ASes is a *very bad* idea

OSI networking is the future(?)

- since 2008-02-28, we observe this beauty...

Network	Path
339:752F:14::/48	6939 4716 6175 1880 i
	1221 4777 38610 30071 6175 1880 i
	109 6175 1880 i
	3257 6175 1880 i

- based on RFC 1888, section 2 and 4...

If the first byte of an IPv6 address is hexadecimal 0x03 (binary 00000011), then the remaining 15 bytes SHALL contain a truncated NSAPA [...]

- RFC 1888 has been declared historic by RFC 4048
- this is *NOT* an accident, but seems to be a conscious protest demonstration against the non-availability of IPv6 PI space
- ⇒ better downstream BGP filters needed...
- (...and we *are* working on a RIPE policy for IPv6 PI...)

Cisco brings back the Ghosts!

- recent IOS versions have *new* ghost bugs, v4 & v6

```
Router-A> show ip bgp 195.216.245.0/24
BGP routing table entry for 195.216.245.0/24, version 63676028
Paths: (1 available, best #1, table Default-IP-Routing-Table)
```

```
1273 33775 47348
```

```
193.149.B.B (metric 133888) from 193.149.B.B
```

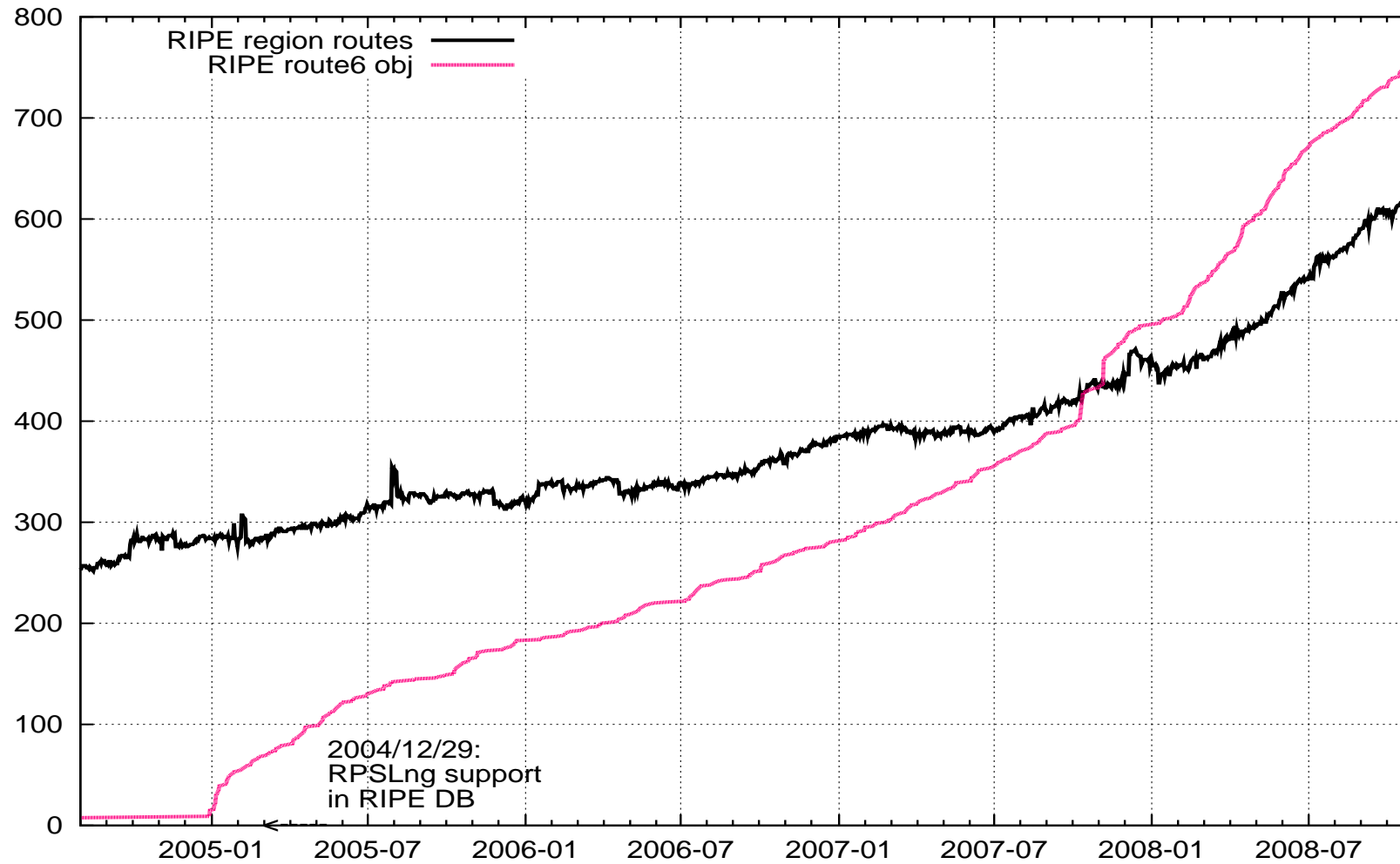
```
Origin IGP, metric 0, localpref 100, valid, internal, best
```

```
Community: 1273:12826 5539:200
```

```
Router-B> show ip bgp 195.216.245.0/24
% Network not in table
```

- old ghost bug required BGP filter change during updates
- this bug is independent of any config changes, “just happens”
- this is CSCsu59917 and CSCsu03167, in 12.2(33)SXH3 + 12.4T
- *if you run affected IOS, upgrade!*

Graphics: route6 objects vs. routes seen



route6 correlation (RIPE region)

- on 2008-05-04:
 - 492 BGP routes from RIPE region
 - 630 route6: objects in RIPE DB
- correlation?
 - multiple origin route6's (11x 2002::/16, 5x 2001::/32, ...)
 - \Rightarrow 612 route6 objects for *unique* prefixes
 - 43 route6 objects for prefixes from *other* RIRs...
- so...

route6 correlation (2)

- ... and this is what I found:

RIPE prefix, route6 ok	358	:-)
RIPE prefix, route6 missing	125	!!!
RIPE prefix, route6 origin mismatch	8	
RIPE prefix, BGP inconsistant AS	6	
route6 objects without BGP route	201	???
other region, route6 ok	40	
other region, route6 missing	679	
other region, route6 origin mismatch	3	
other region, BGP inconsistant AS	9	

- \Rightarrow close-up view shows “more work needed”
- in other RIR regions, situation is worse (no IRR DBs yet, etc.)

route6 object example

- it's as easy as this...

```
route6:      2001:608::/32
descr:      DE-SPACE-2001-0608
descr:      SpaceNET AG, Munich
origin:     AS5539
notify:     noc@space.net
mnt-by:     SPACENET-N
changed:    gert@space.net 20041230
source:     RIPE
```

- strongly recommended, helps upstream/peer ASes build decent BGP filters, based on IRR data

References

- Ghost Route Hunter: <http://www.sixxs.net/tools/grh/>
- List of IPv6 blocks allocated by the RIRs:
<http://www.ripe.net/rs/ipv6/stats/index.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/R57-v6-table/>

Questions?

gert@space.net