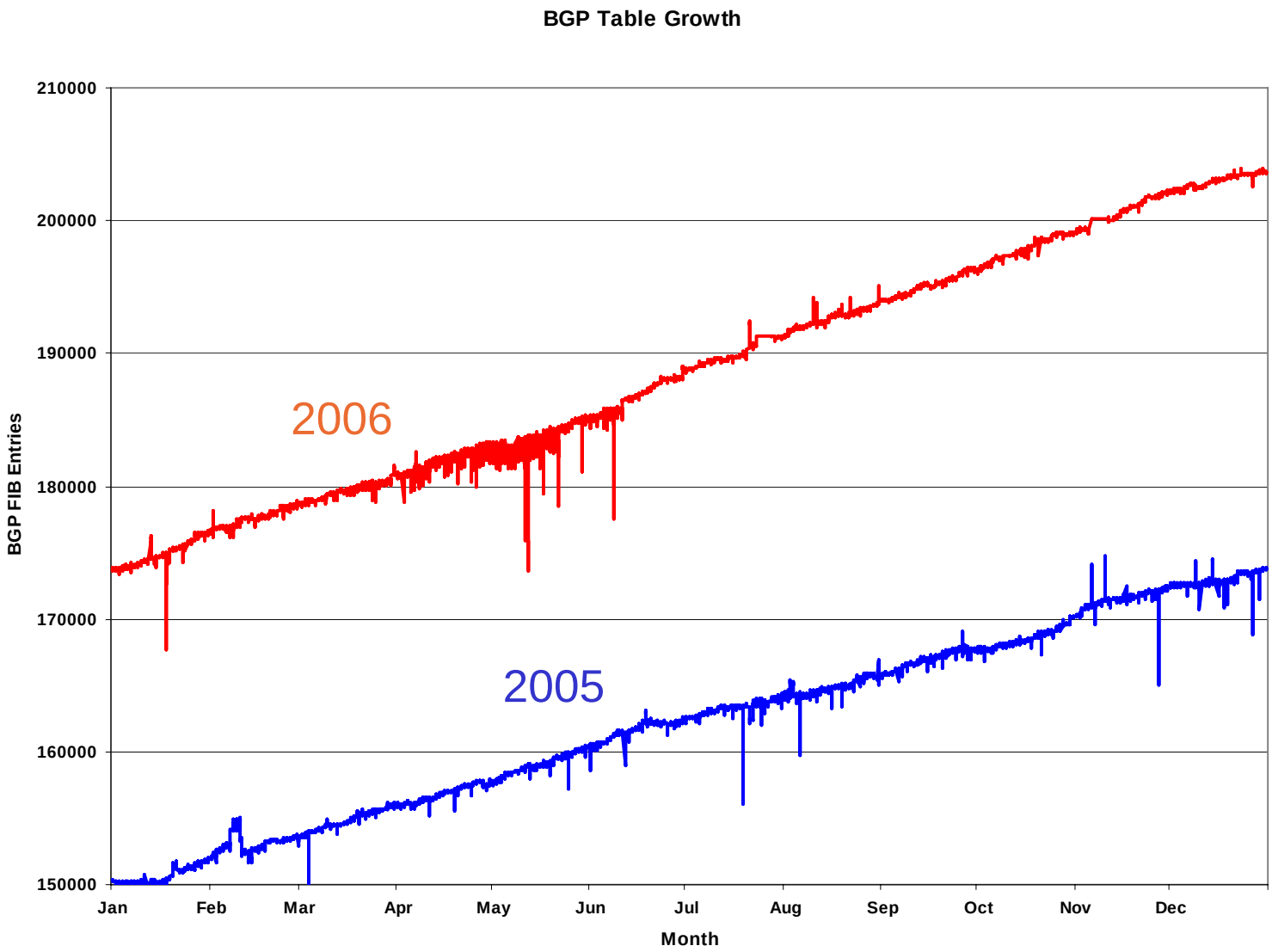
Routing the Internet in 2006

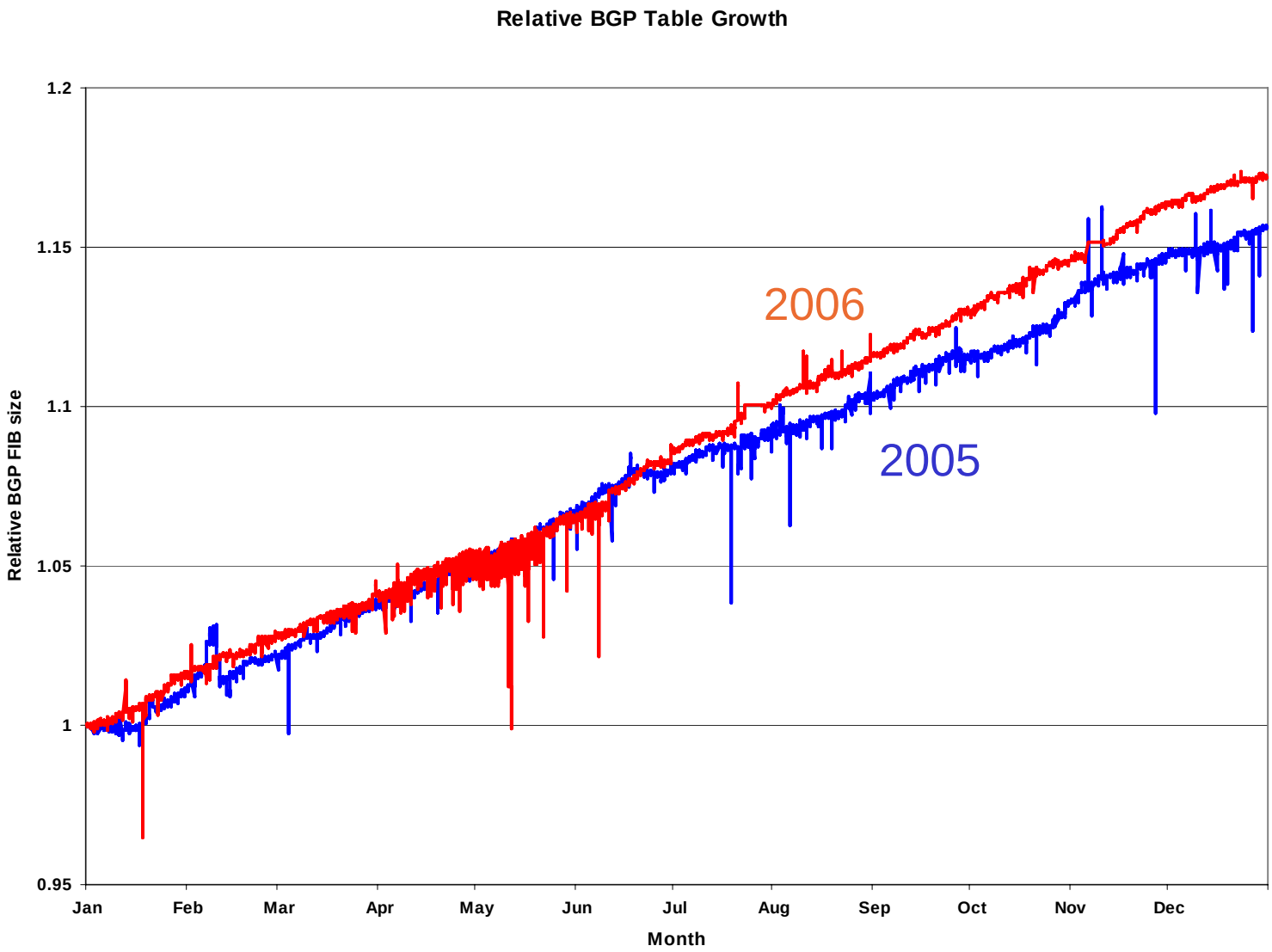
Geoff Huston
APNIC
March 2007

IPv4 in 2006
Total Advertised BGP Prefixes



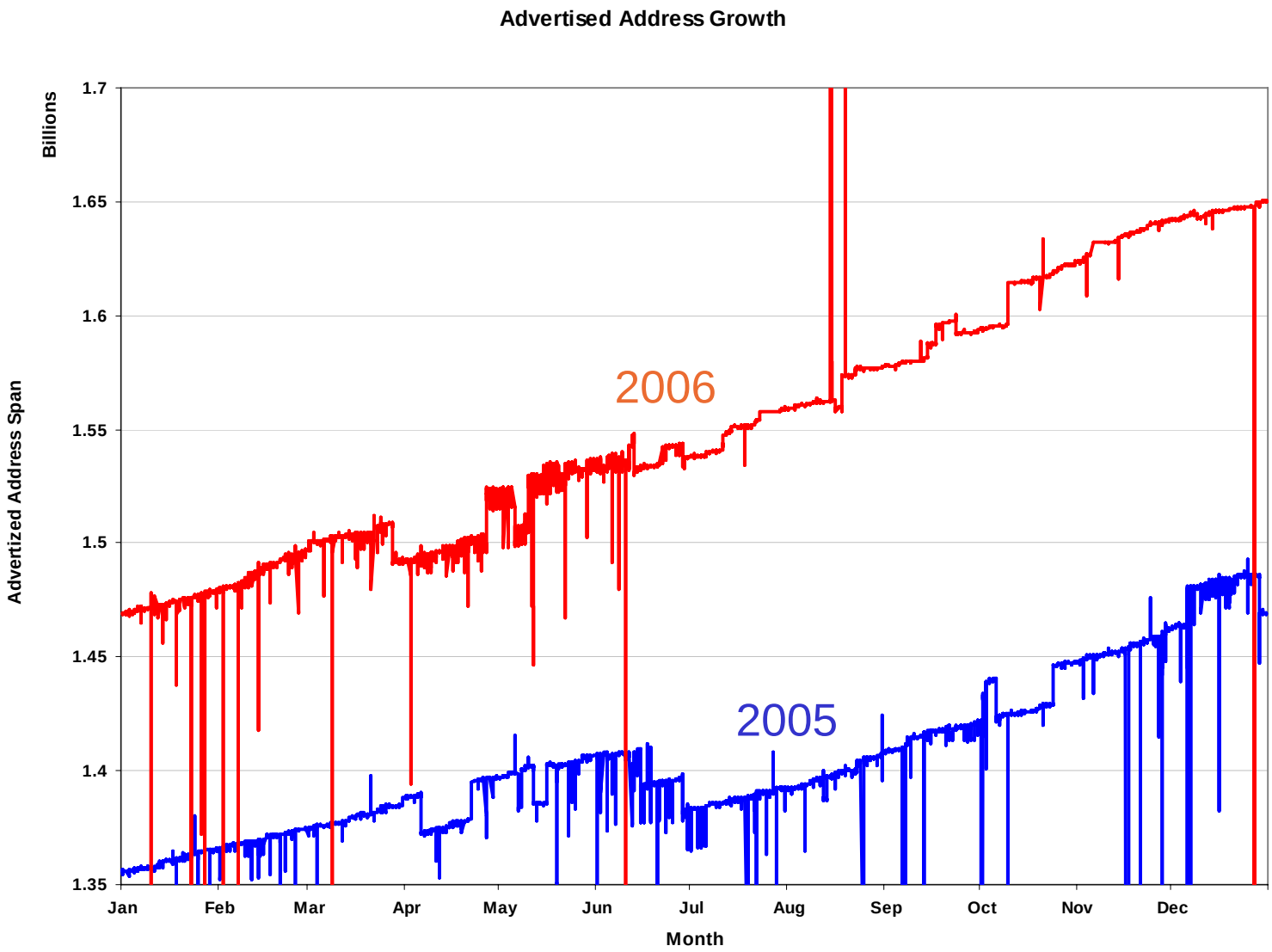
IPv4 in 2006

Relative Growth: 2005 to 2006



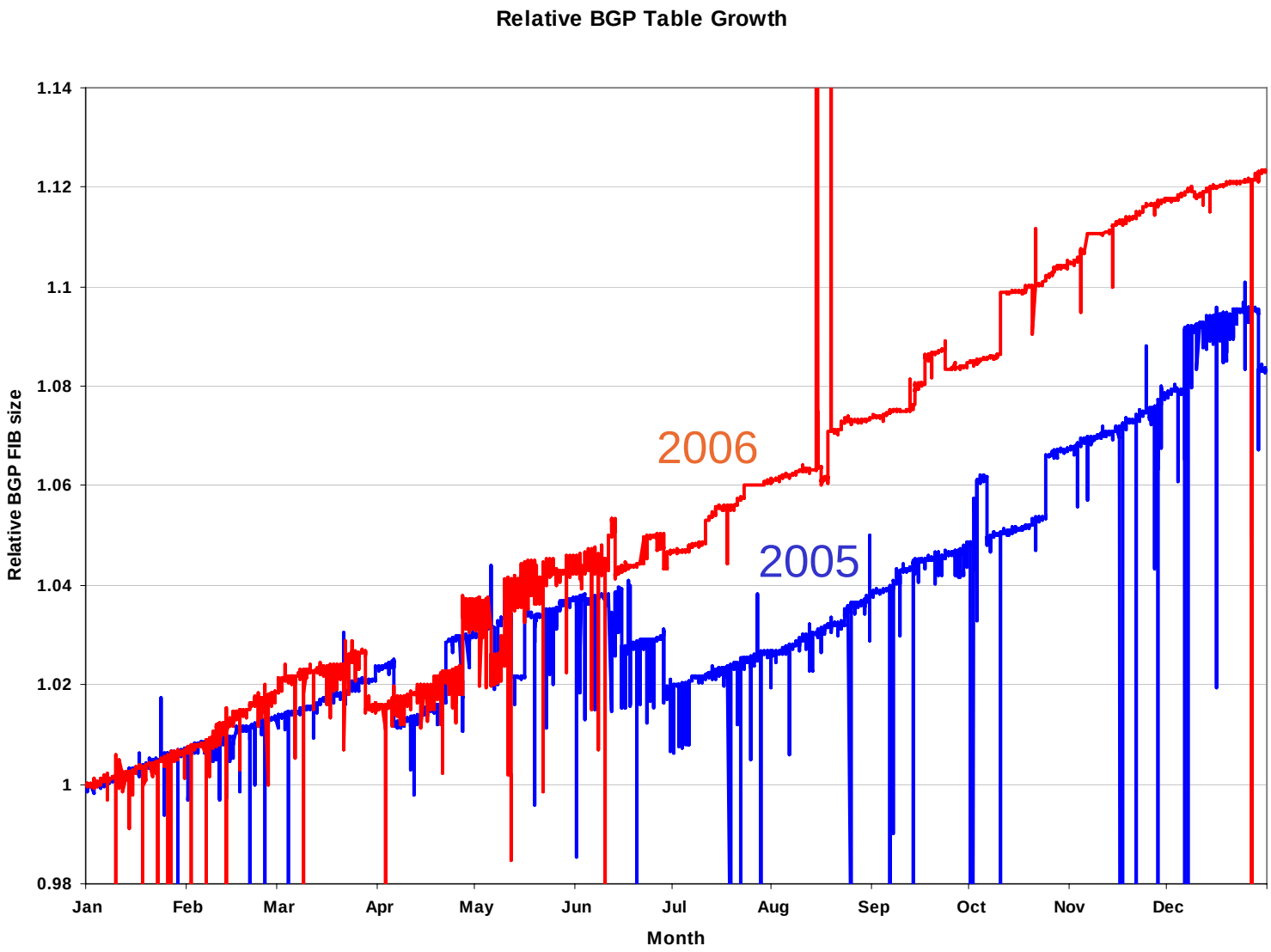
IPv4 in 2006

Total Advertised IPv4 Address Span

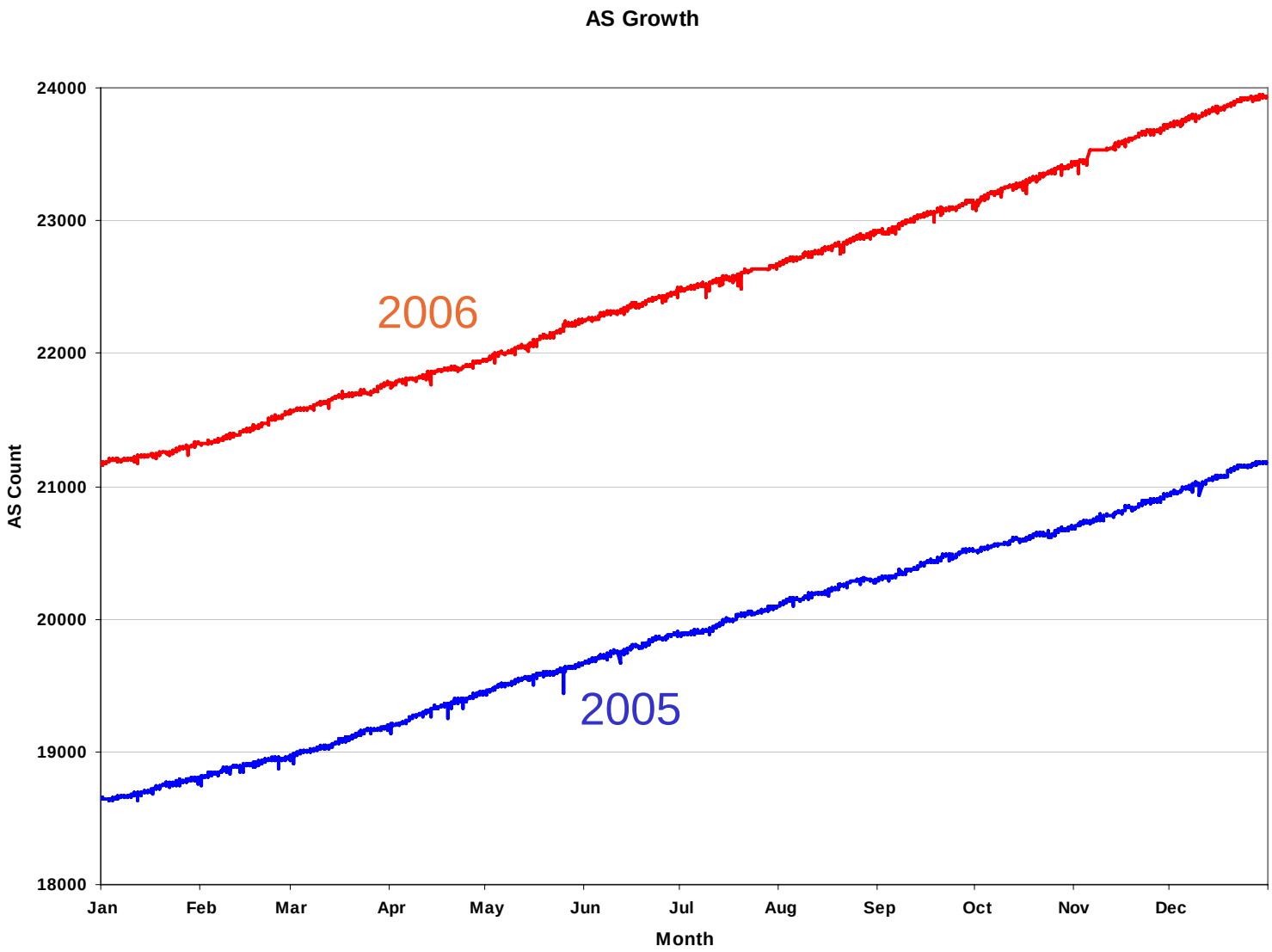


IPv4 in 2006

Relative Growth: 2005 to 2006

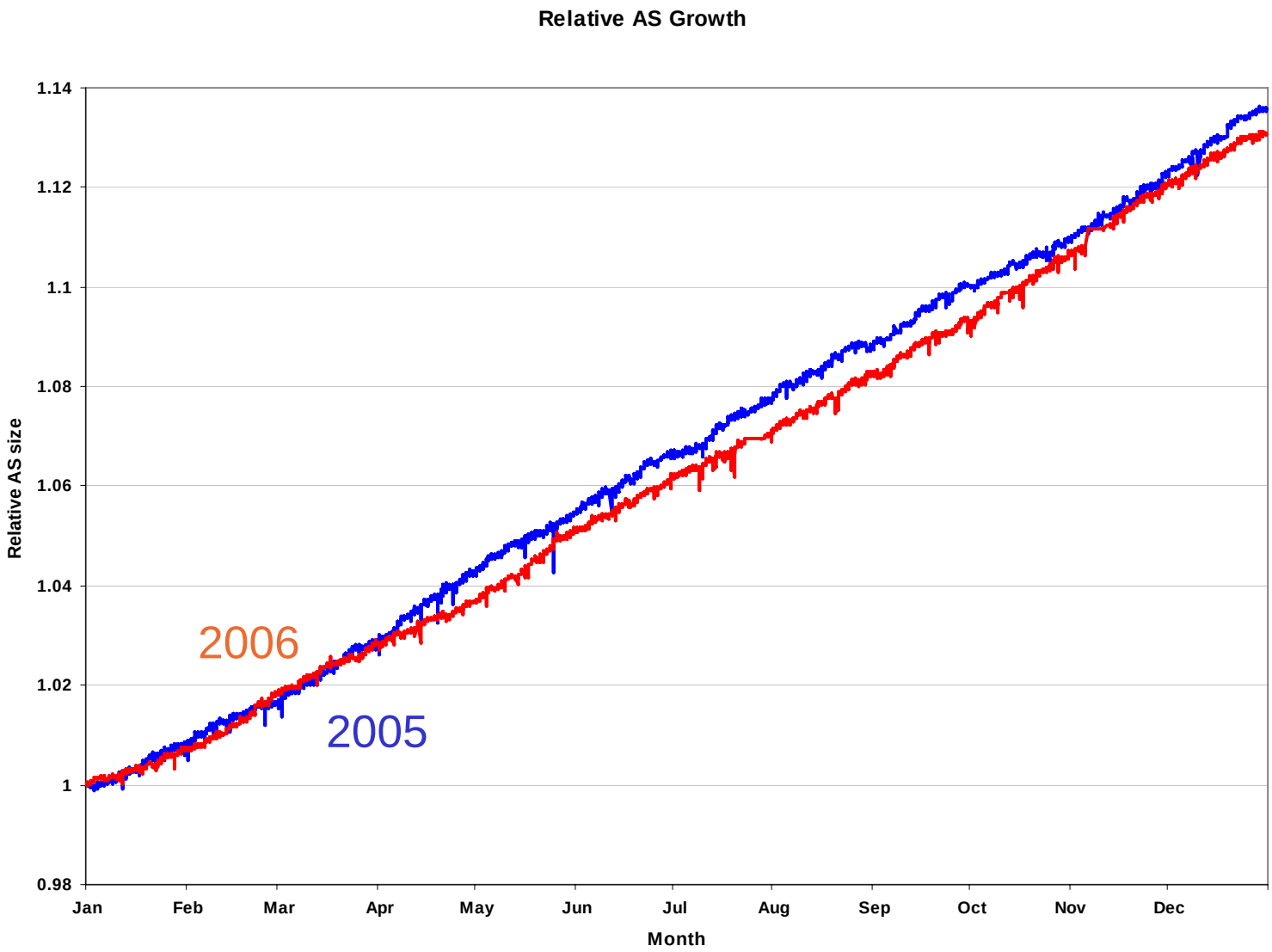


IPv4 in 2005
Total Advertised AS Numbers



IPv4 in 2005

Relative Growth: 2005 to 2006



IPv4 – Vital Statistics for 2006

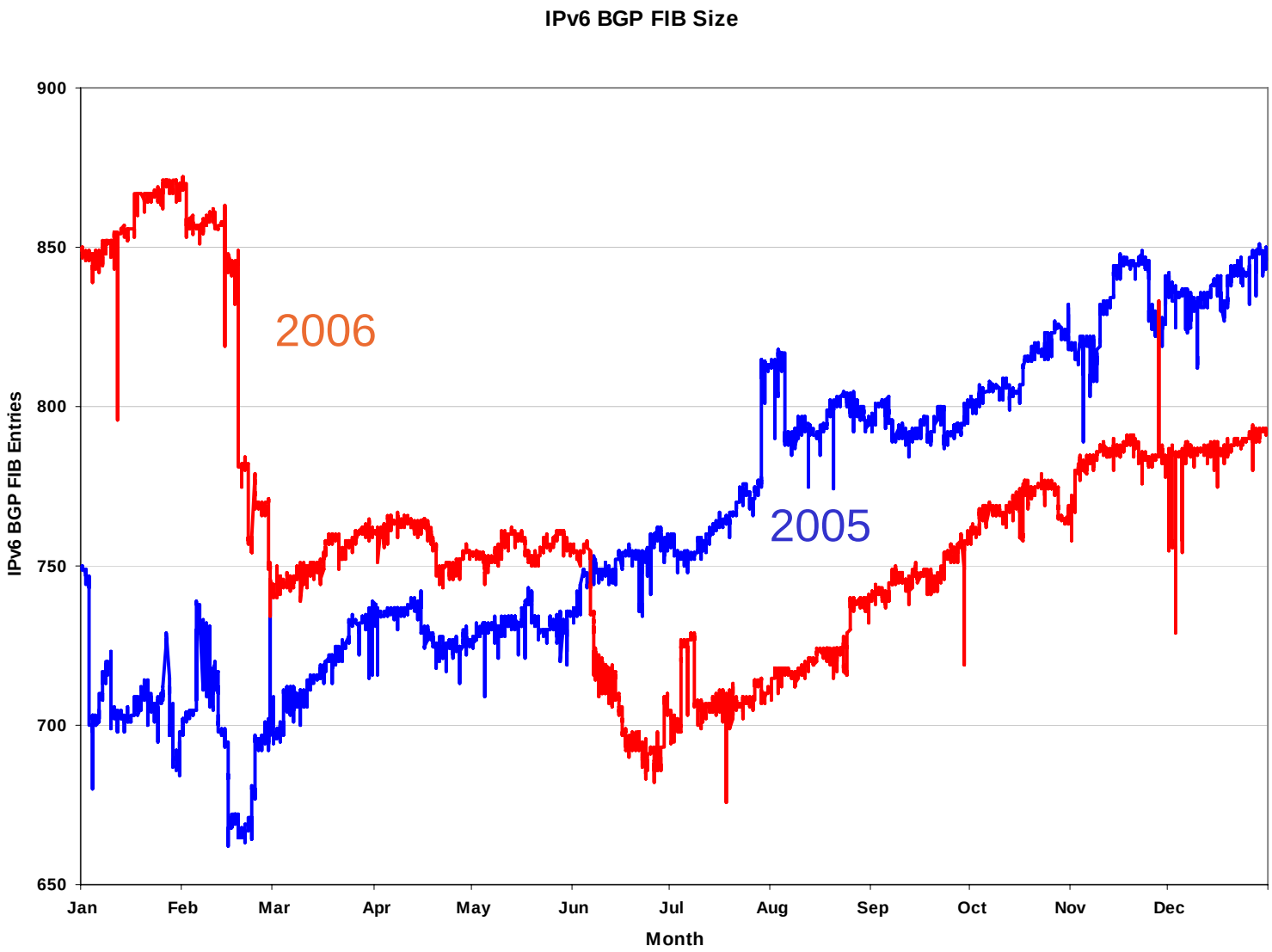
Prefixes	173,800 –	203,800	+17%
Roots	85,800 –	100,800	+17%
Specifics	88,000 –	103,000	+17%
Addresses	87.6 –	98.4 (/8)	+12%
ASNs	21,200 –	24,000	+13%

Average advertisement size is getting smaller (8,450 – 8,100)
Average address origination per AS is getting smaller (69,600 – 69,150)
Average AS Path length steady (3.4)
AS transit interconnection degree rising (2.56 – 2.60)

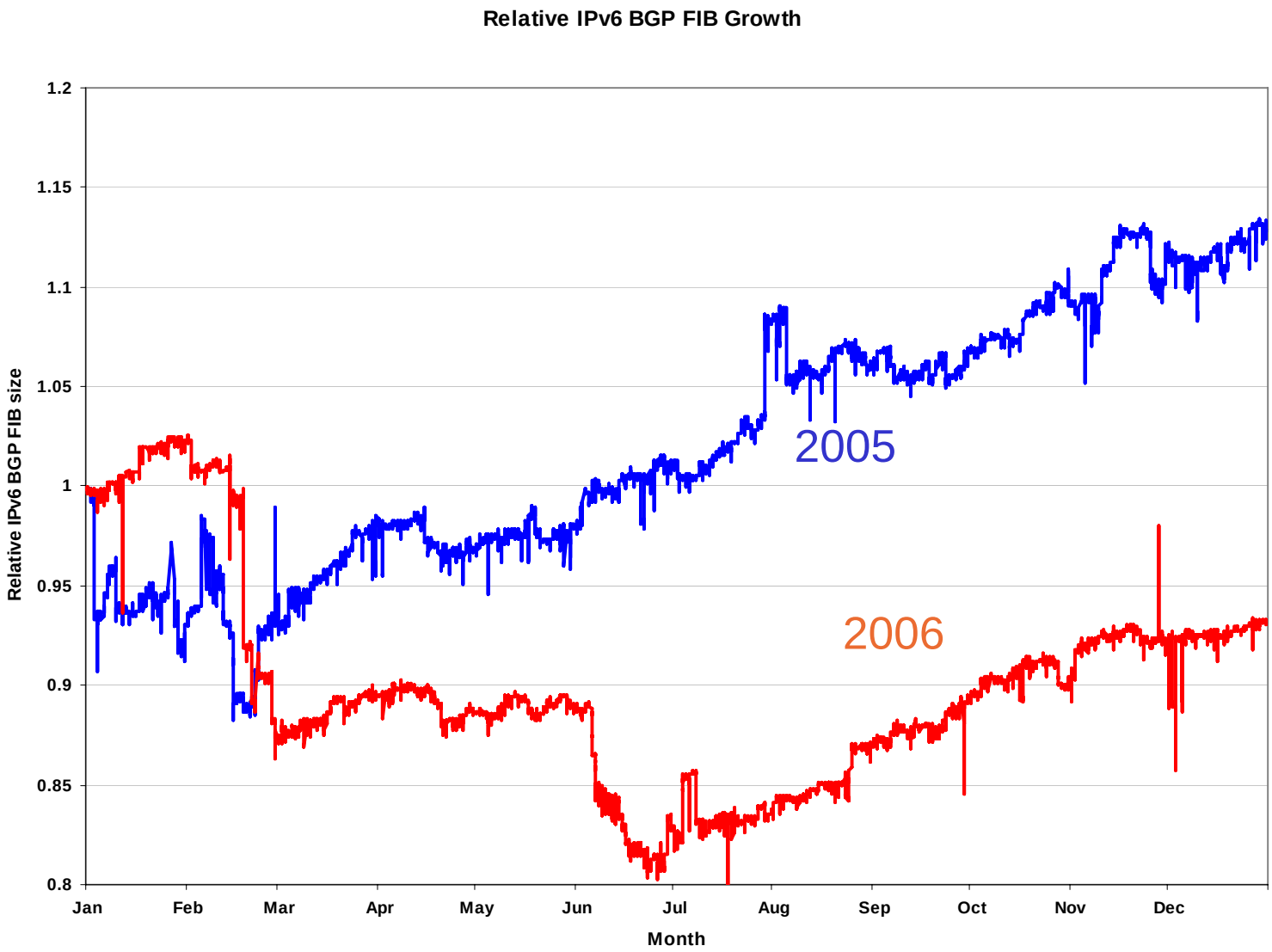
The IPv4 network continues to get denser, with finer levels of advertisement granularity.

More interconnections, more specific advertisements

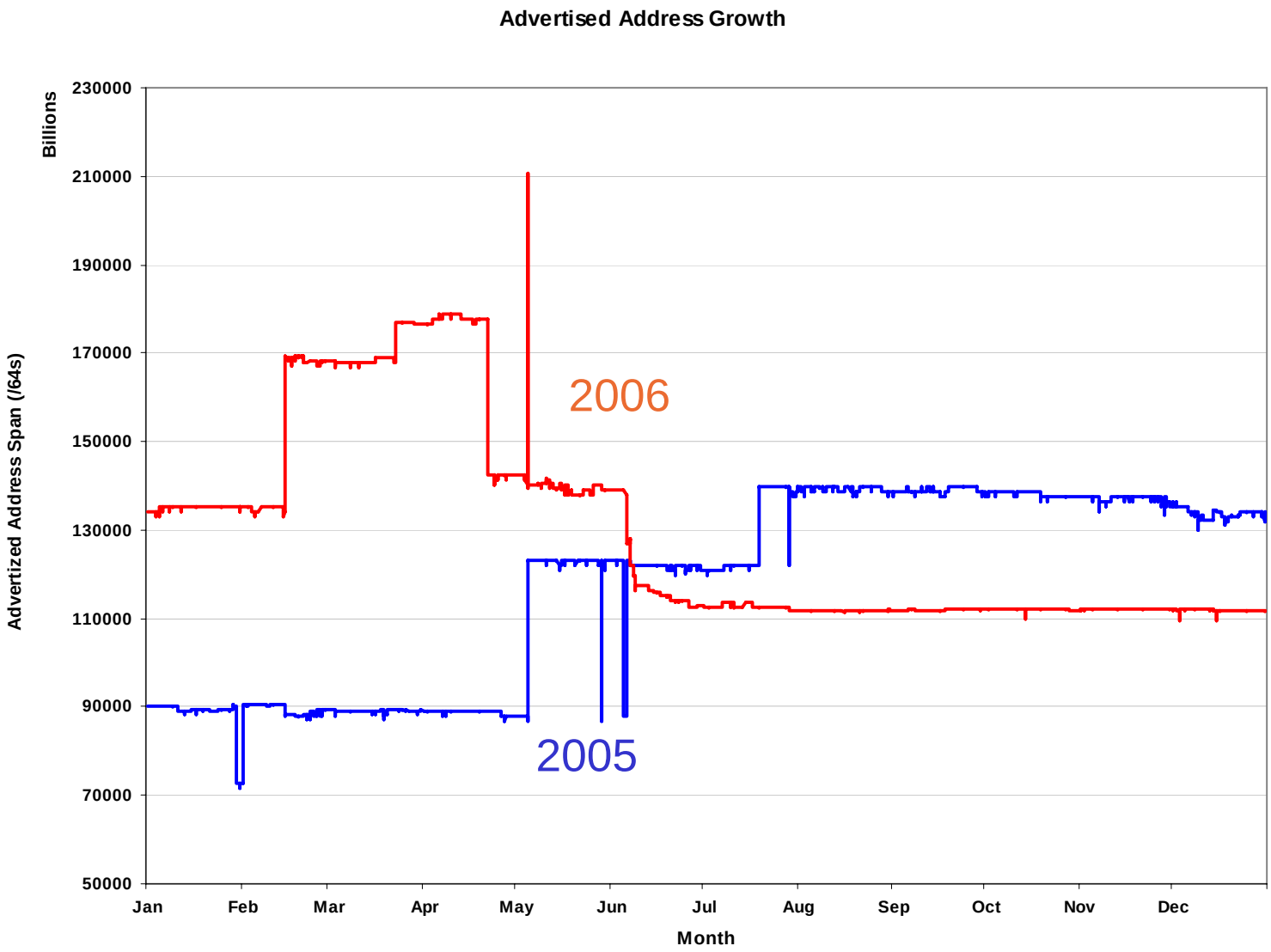
IPv6 in 2006
Advertised Prefix Count



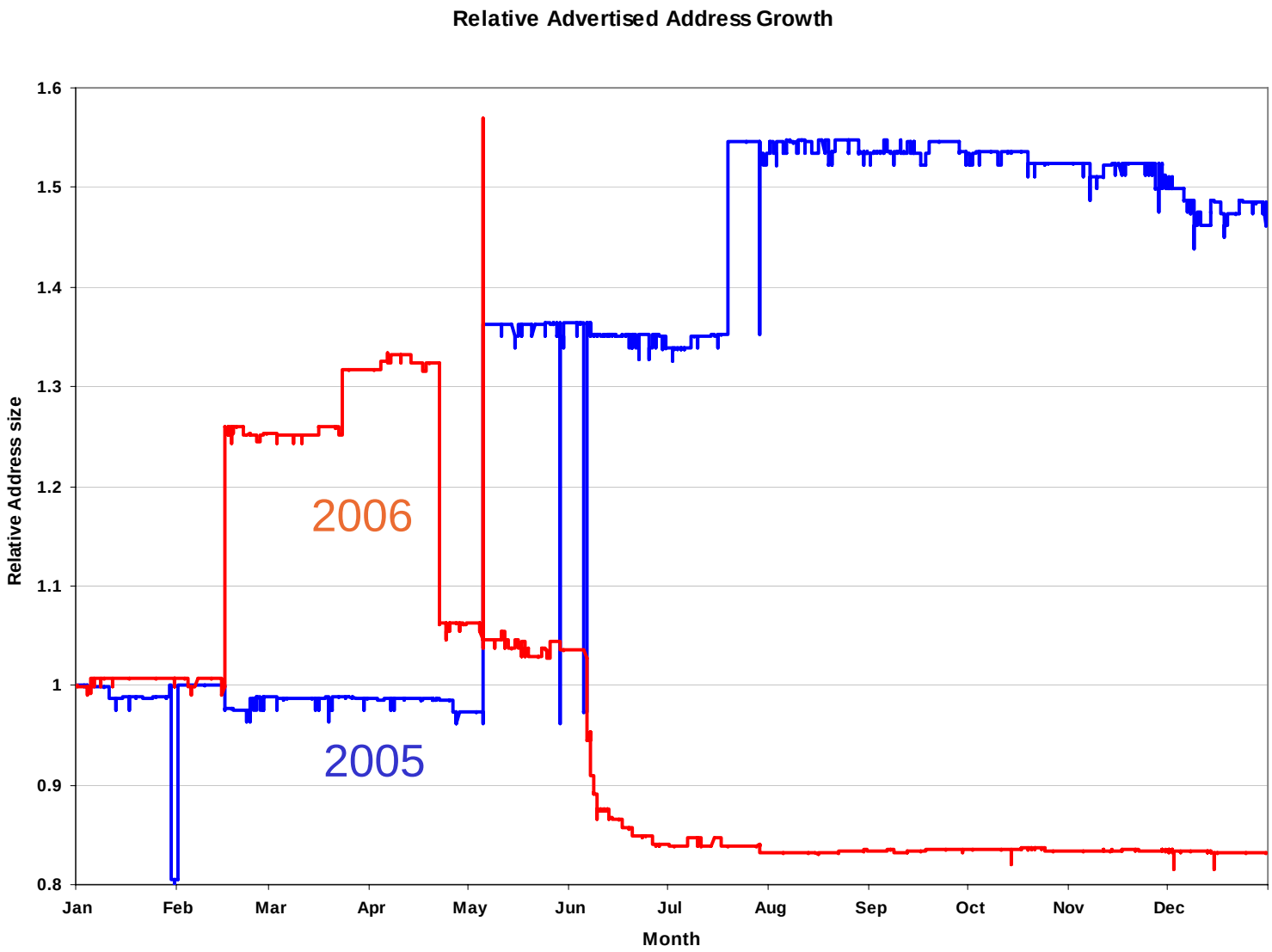
IPv6 in 2006
Advertised Prefix Count



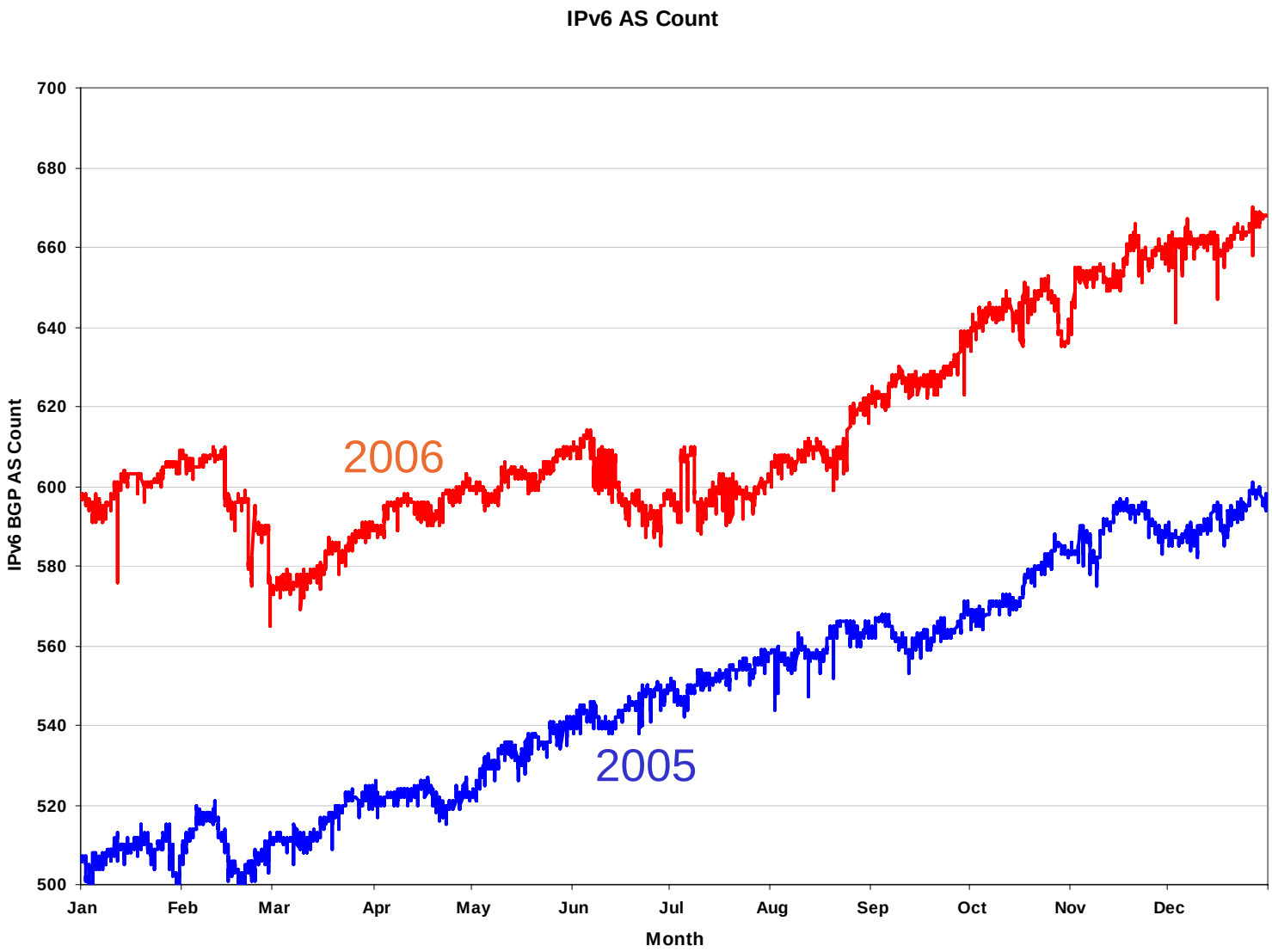
IPv6 in 2006
Advertised Address Span



IPv6 in 2006
Advertised Address Span



IPv6 in 2006
Total Advertised AS Numbers



IPv6 in 2006
Total Advertised AS Numbers

IPv6 – Vital Statistics for 2006

Prefixes	850 – 790	- 7%
Roots	640 – 660	+ 3%
Specifics	210 - 130	-38%
Addresses	13.5 – 11.5 (10**13)	-15%
ASNs	595 – 670	+12%

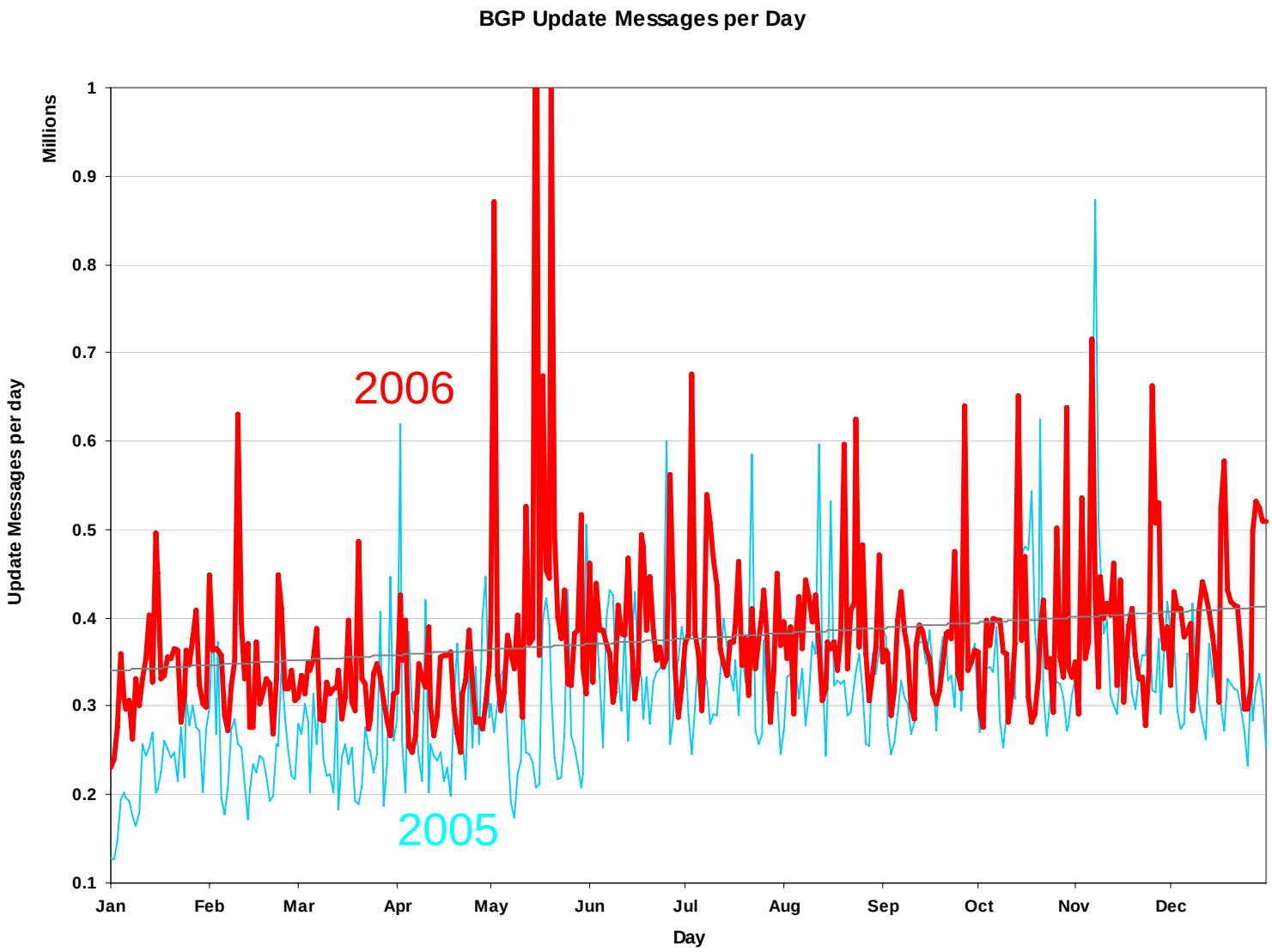
Average advertisement size is getting larger
Average address origination per AS is getting larger
Average AS Path length variable between 3 – 5
AS interconnection degree variable

Through 2006 the IPv6 network saw the withdrawal of the 6Bone prefixes. There was a small increase in the number of ASs, but the derived rate of new AS's supporting IPv6 is slowing as compared to 2005.

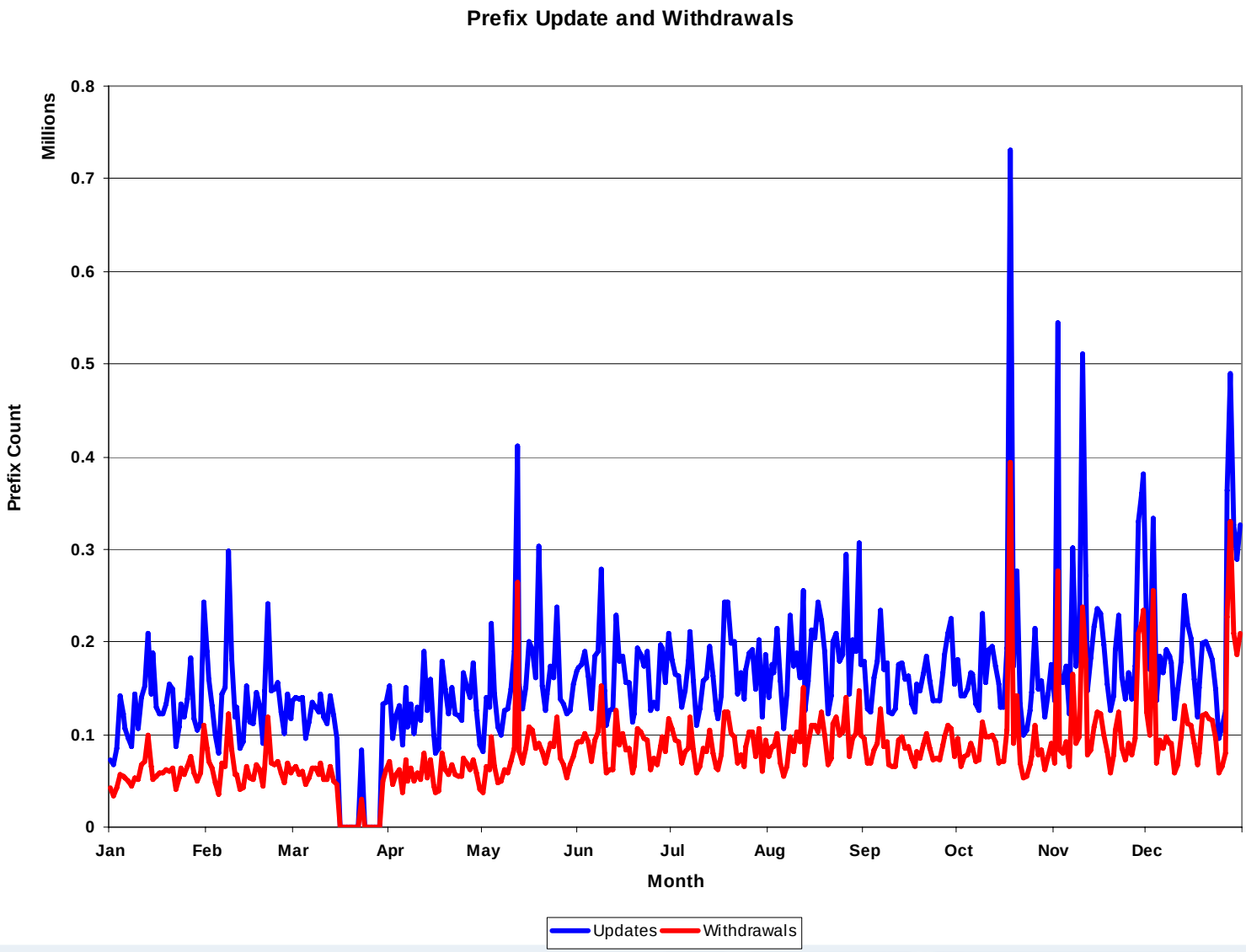
BGP Behaviour - Methodology

- Examine update and withdrawal rates from BGP log records for 2006 from a viewpoint within AS1221
  - Eliminate local effects to filter out non-DFZ BGP updates
  - Look at the relative rate of updates and withdrawals against the table size
  - This looks at the traffic associated with a **SINGLE** BGP peer session

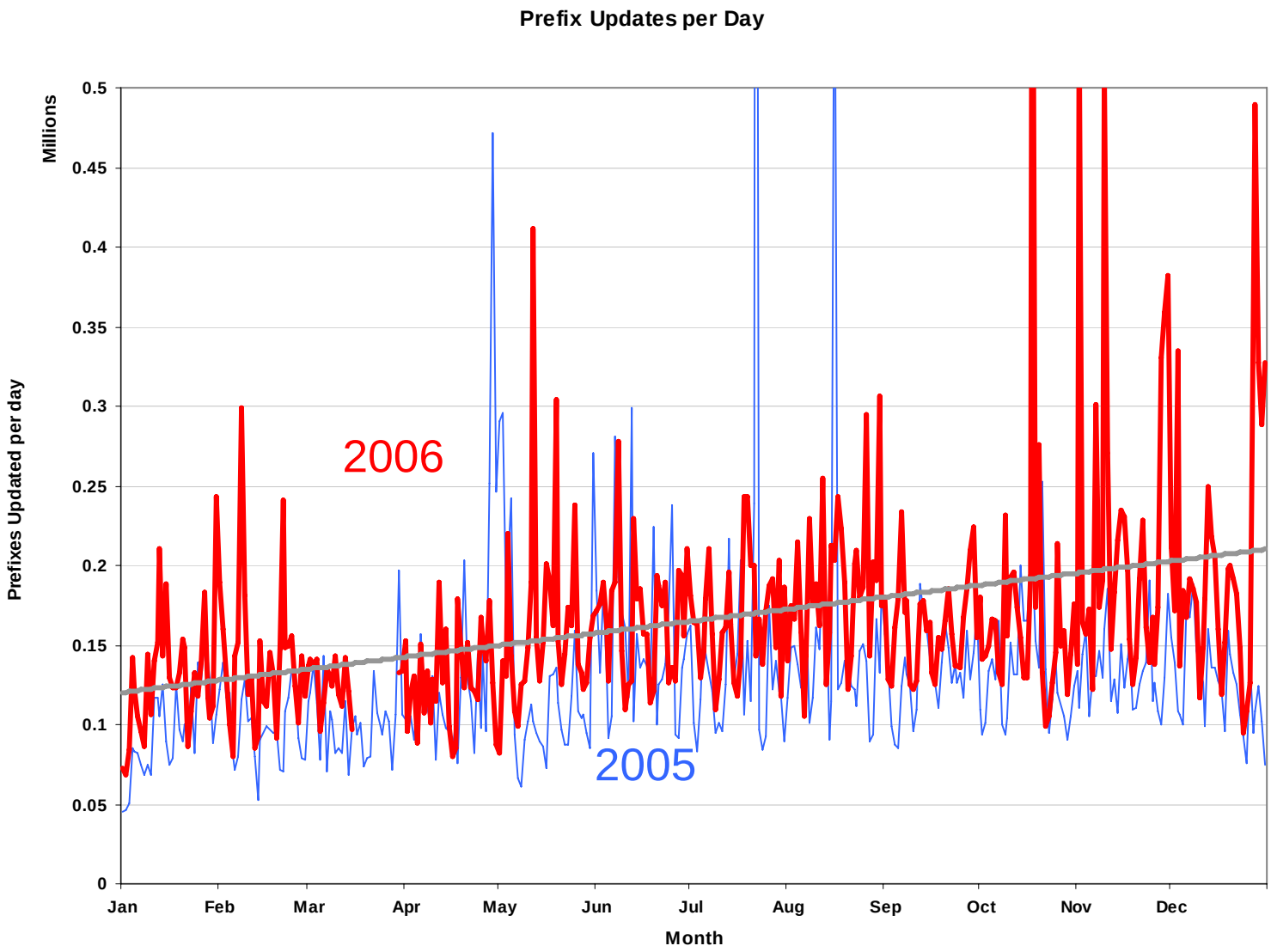
Update Message Rate



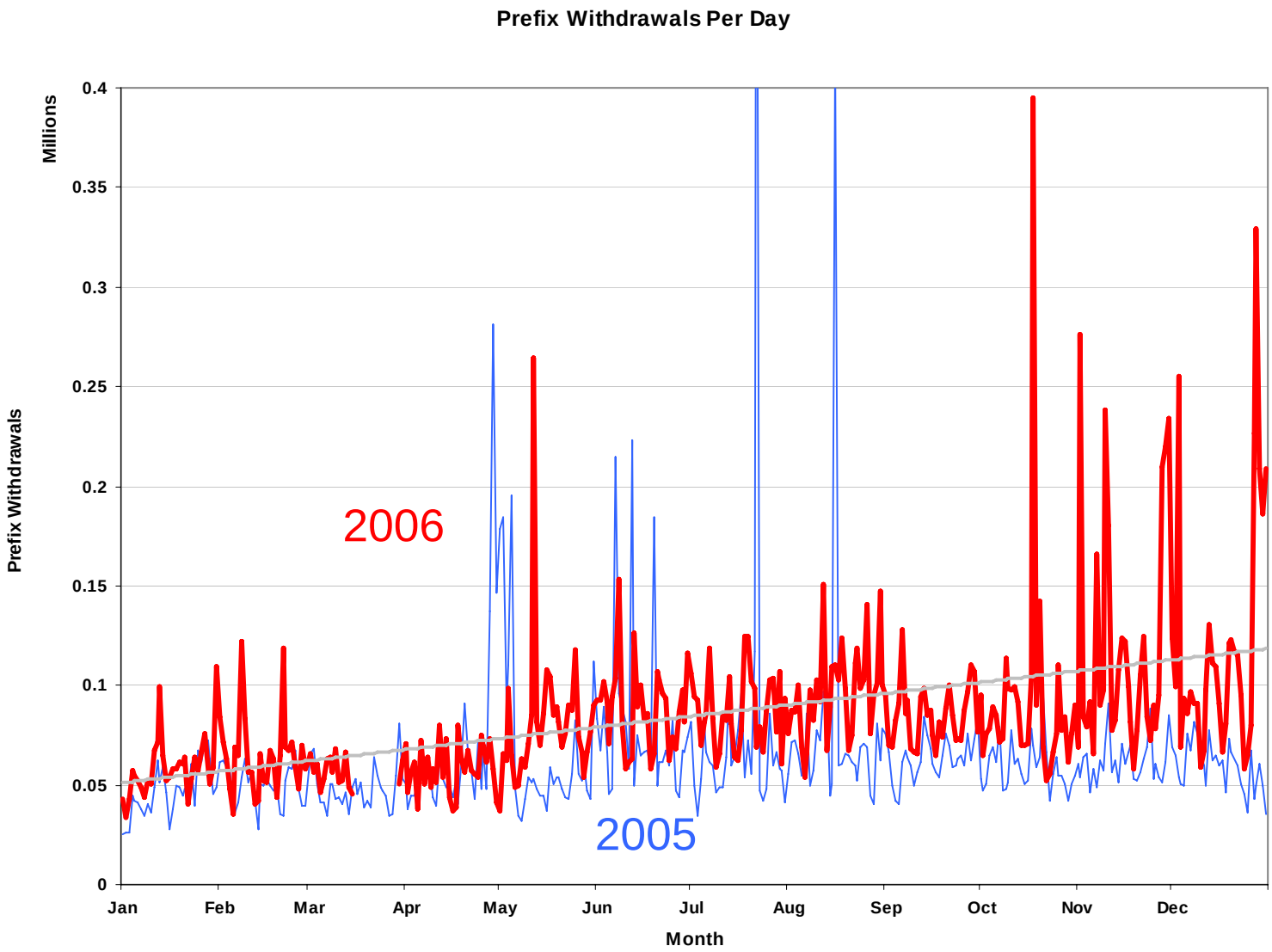
Prefix Update and Withdrawal Rates



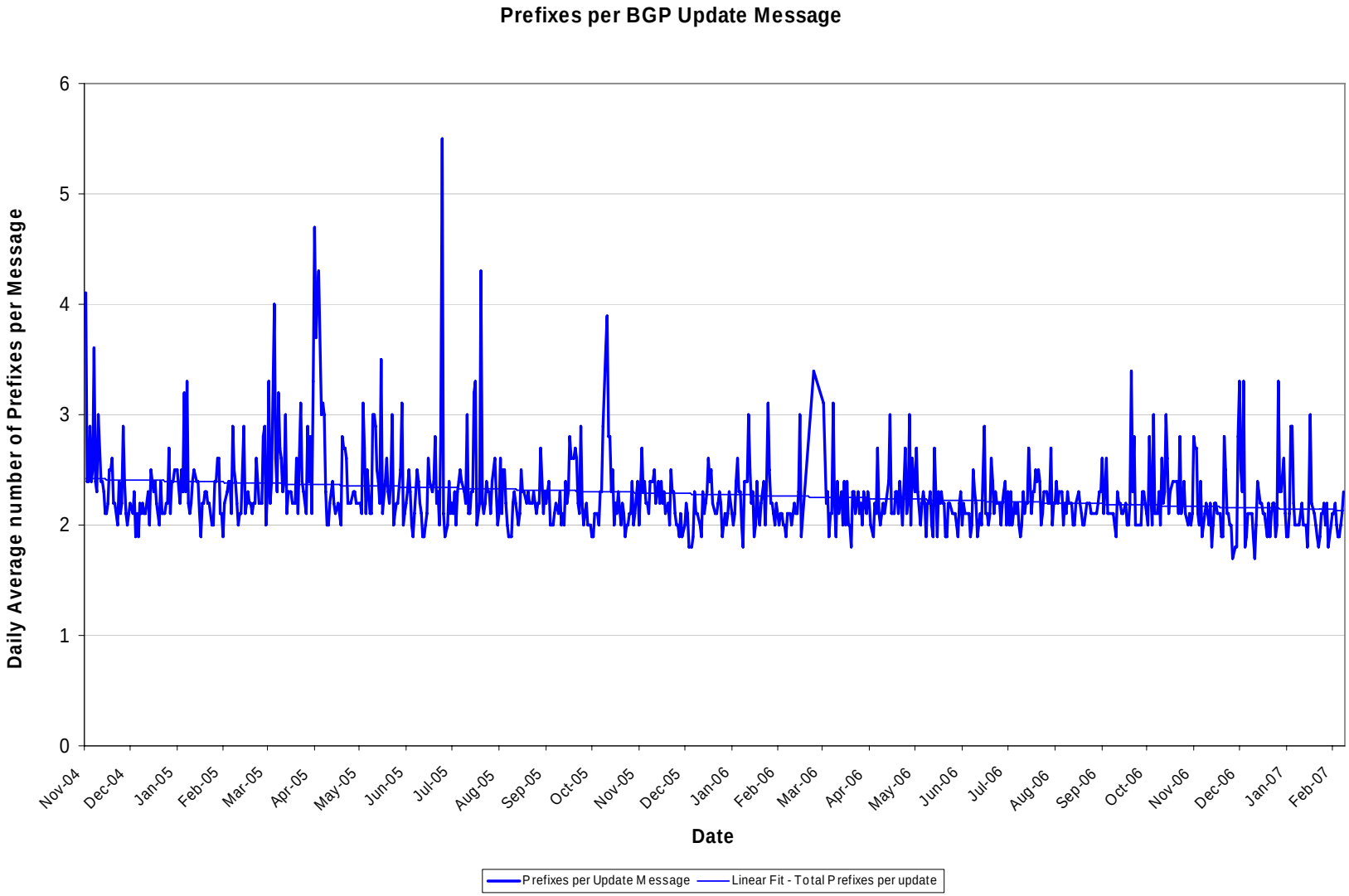
Prefix Update Rates



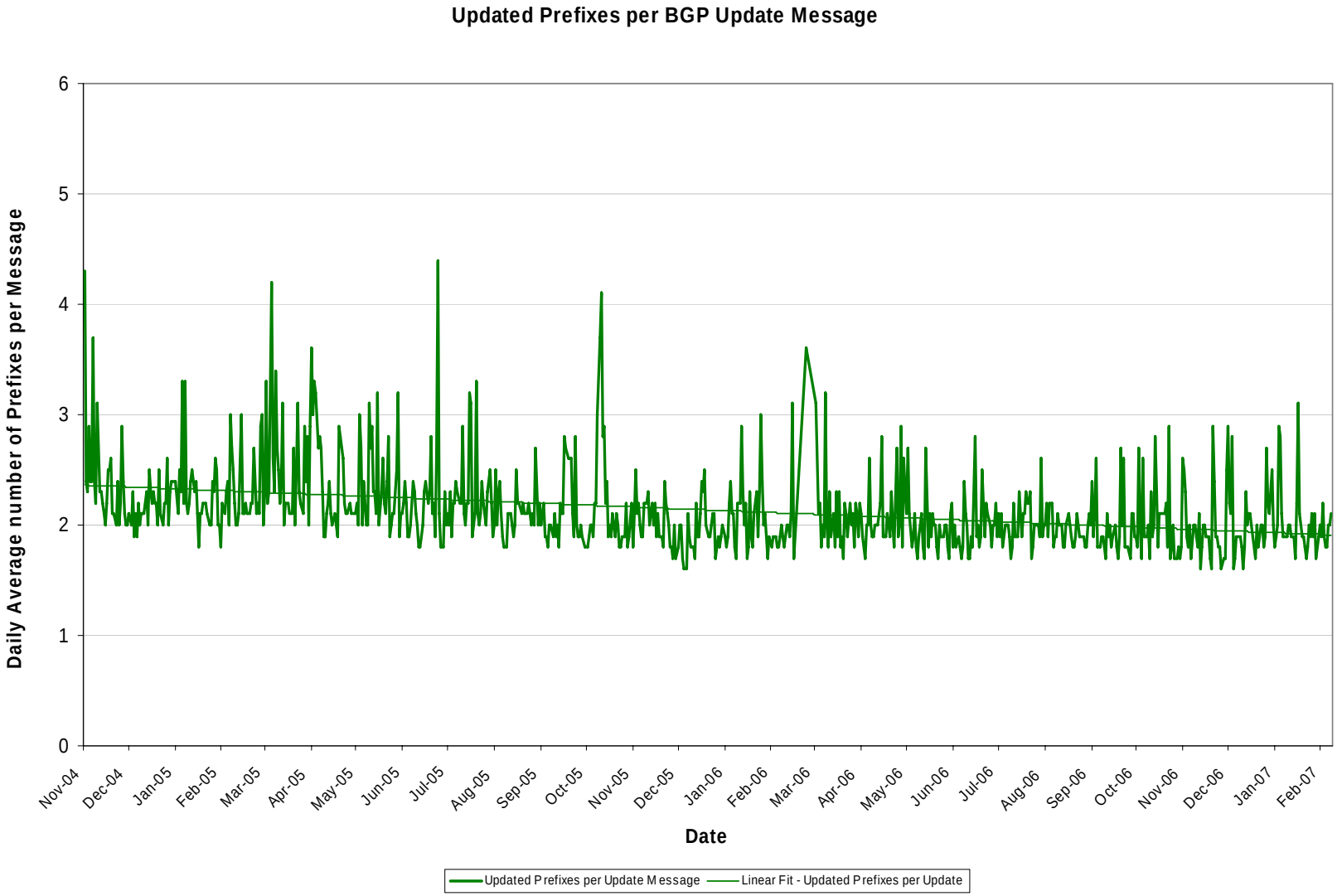
Withdrawal Rates



Average Prefixes per BGP Update



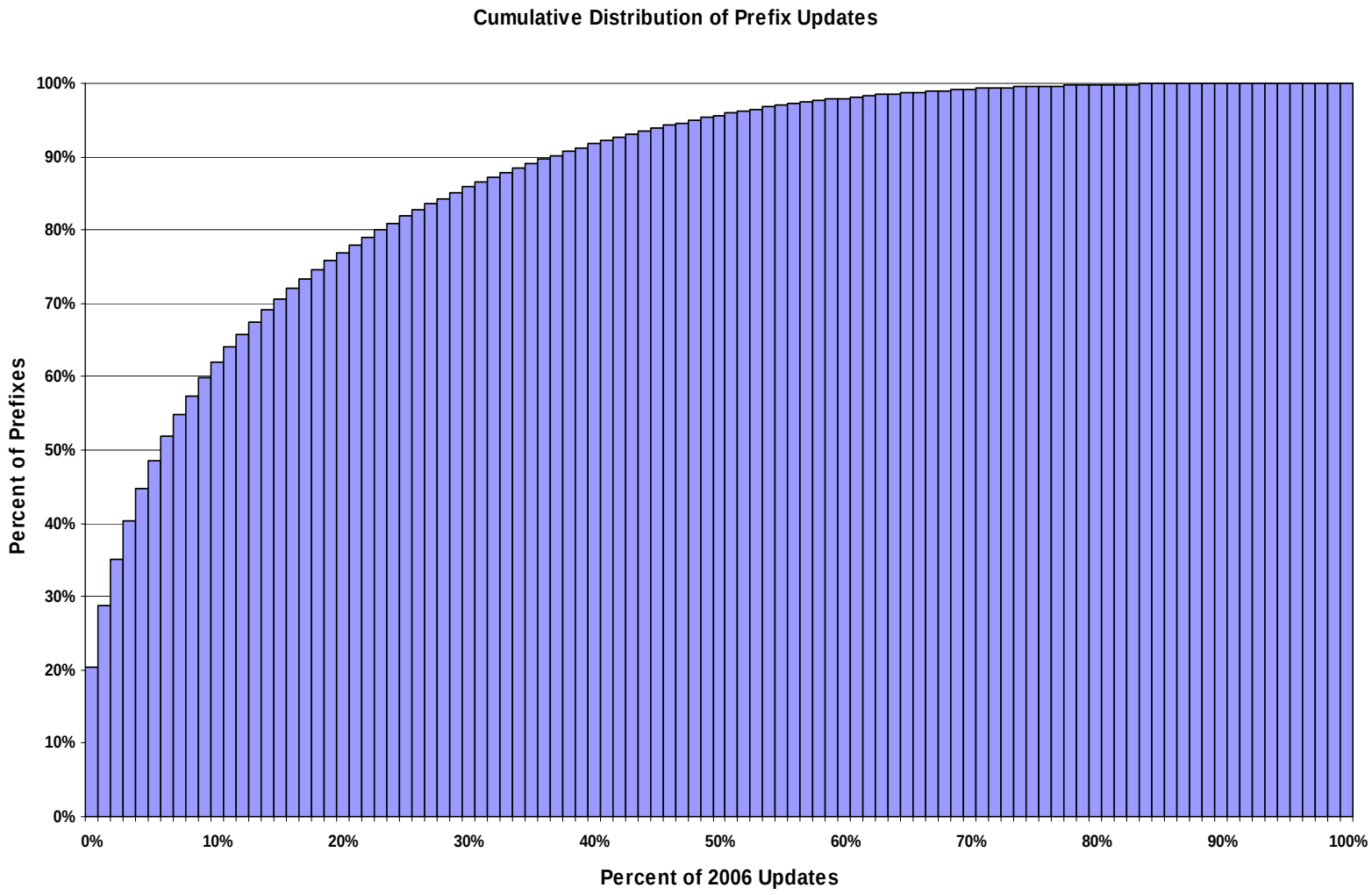
Updated Prefixes per BGP Update Message



Prefix Stats for 2006

- Number of unique prefixes announced: 354,589
- Prefix Updates: 89,582,323 (average = 2.84 per second)
- Prefix Withdrawals: 30,531,219 (average = 0.96 per second)
- Updated prefixes (year end): 203,635
- Withdrawn prefixes: 150,954
- Average Prefixes per BGP Update: 1.95 (down from 2.1 at the start of 2006)

Distribution of Updates by Prefix

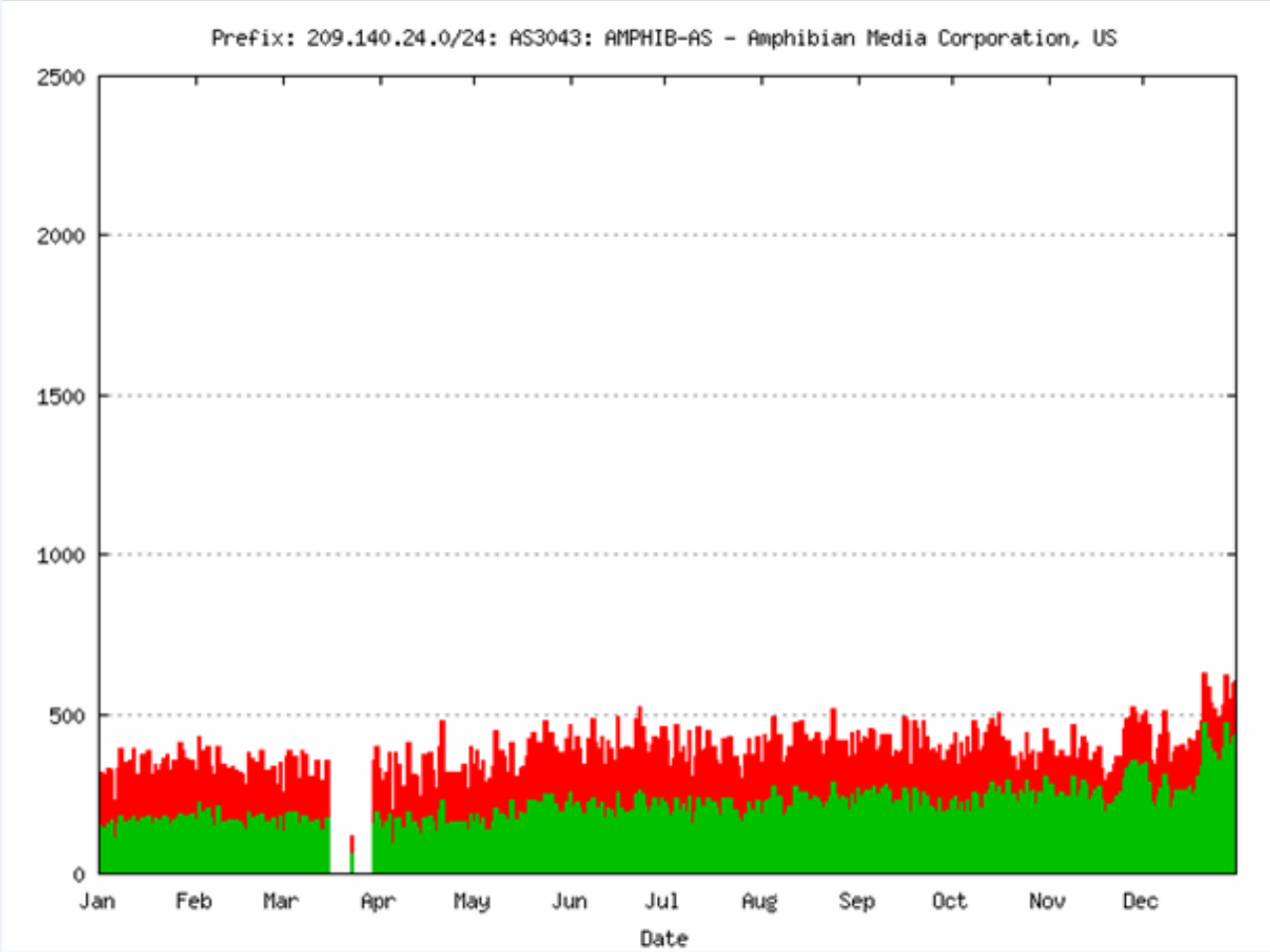


Active Prefixes

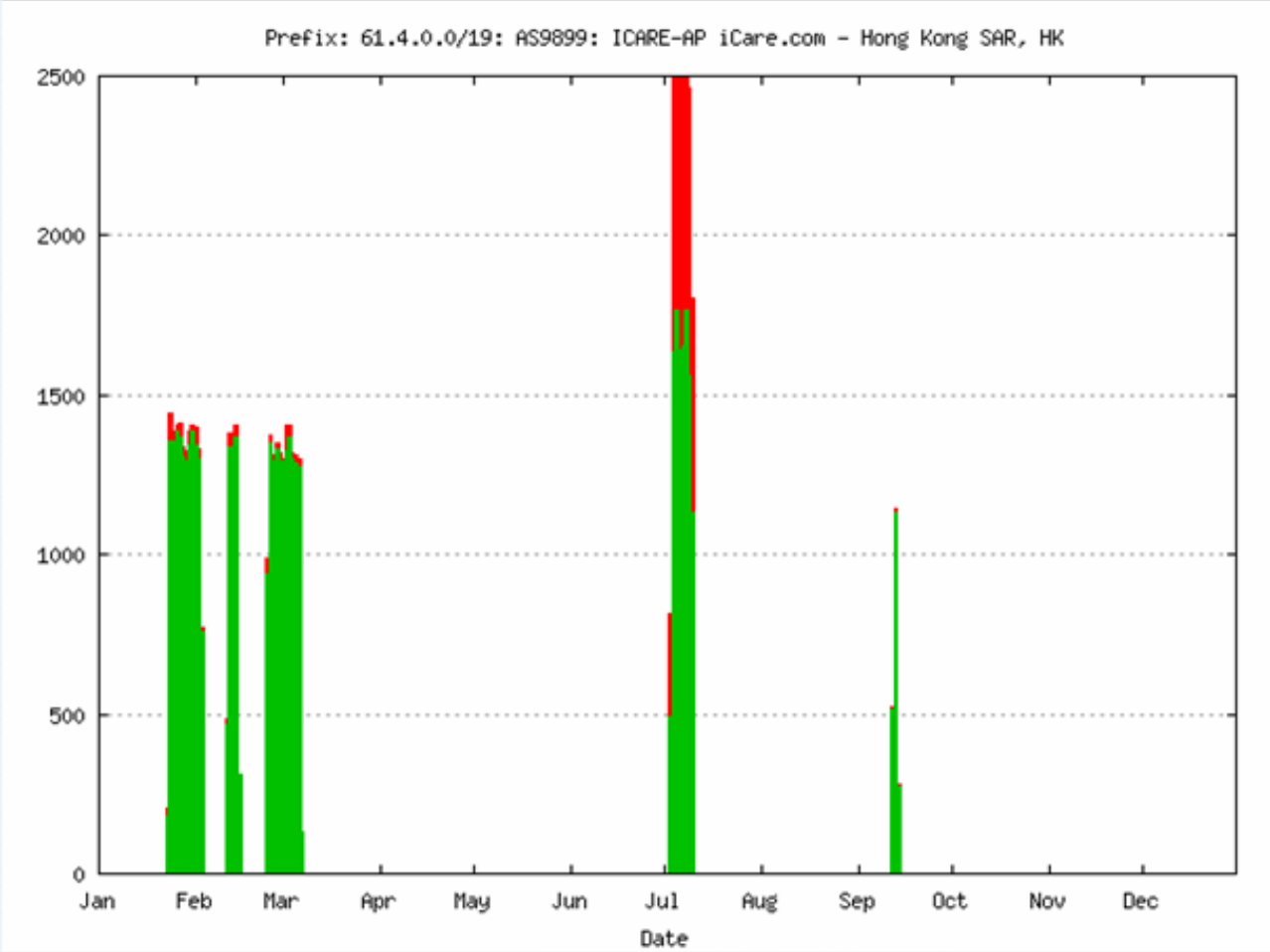
Top 10 Prefixes

	Prefix	Updates	Flaps	Re-Homes
1.	209.140.24.0/24	210,574	151,145	1
2.	61.4.0.0/19	101,901	93,843	35
3.	61.0.0.0/8	89,768	70,863	5,541
4.	81.212.141.0/24	69,688	53,445	12,715
5.	203.199.128.0/19	63,606	51,076	8,592
6.	152.74.0.0/16	61,409	45,532	0
7.	84.205.65.0/24	59,744	44,792	8,454
8.	81.212.149.0/24	59,150	49,159	8,575
9.	193.242.123.0/24	57,717	34,974	16,468
10.	84.205.76.0/24	55,634	41,526	9,110

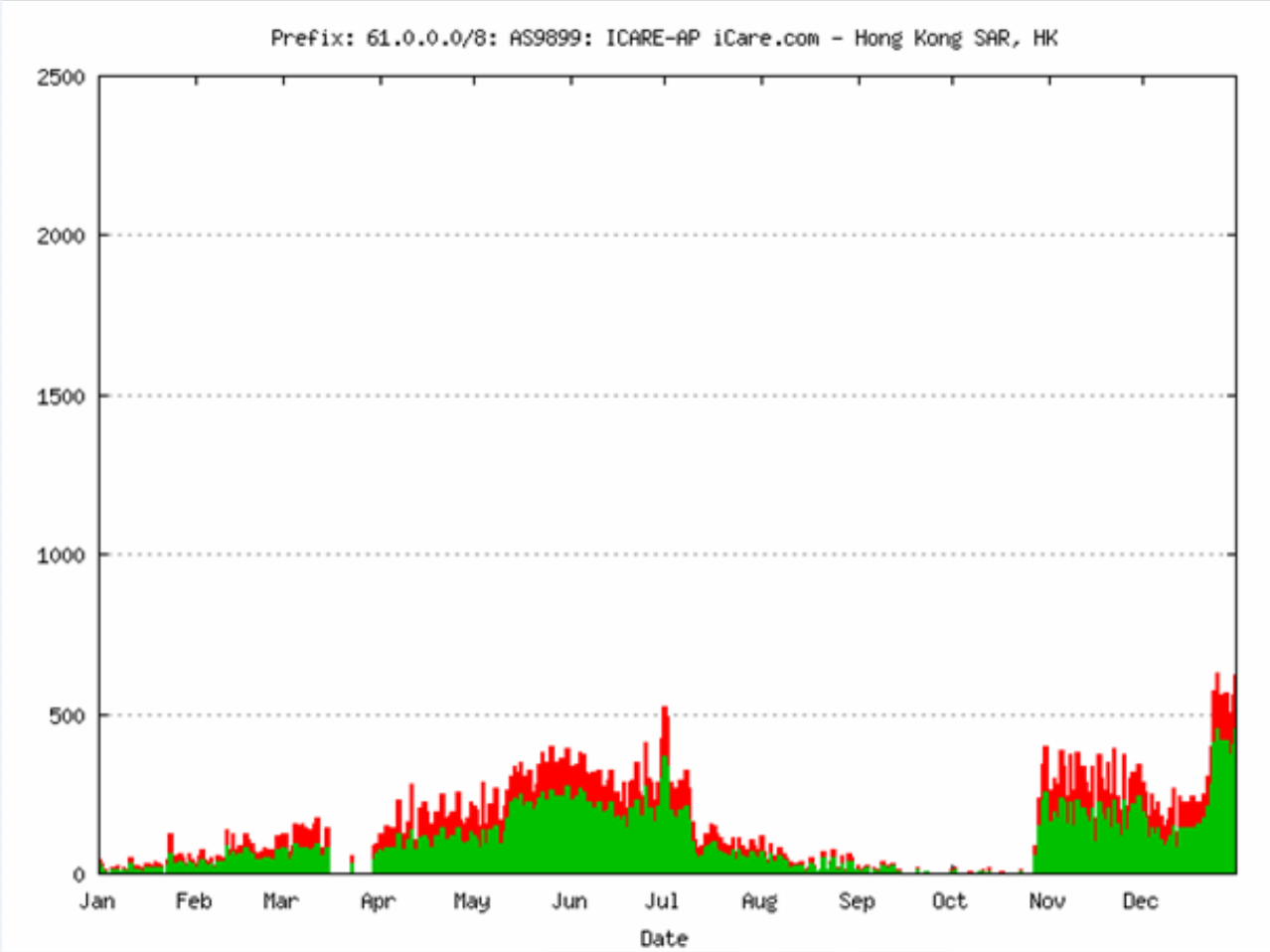
1 – 209.140.24.0/24



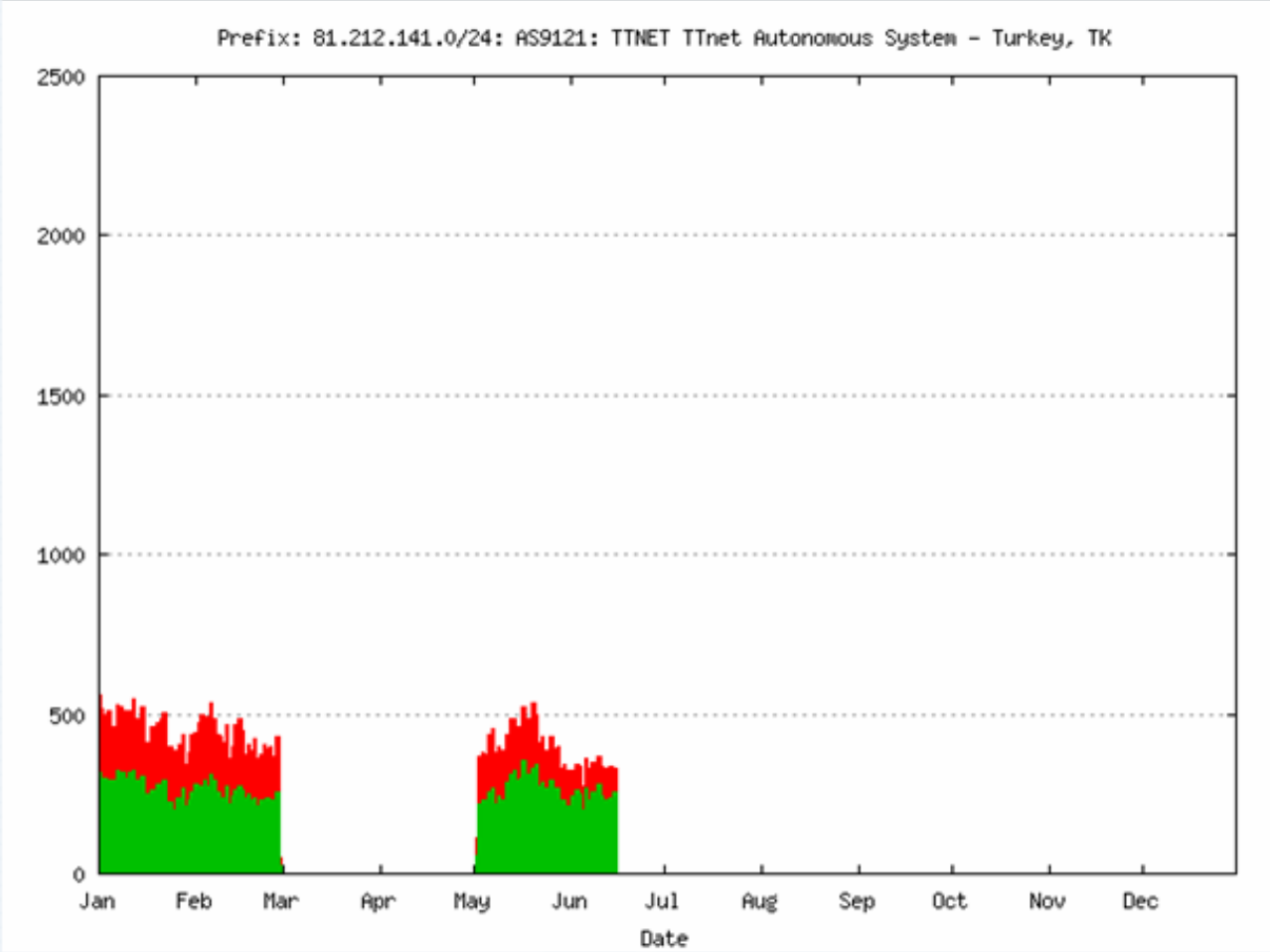
2 - 61.4.0.0/19



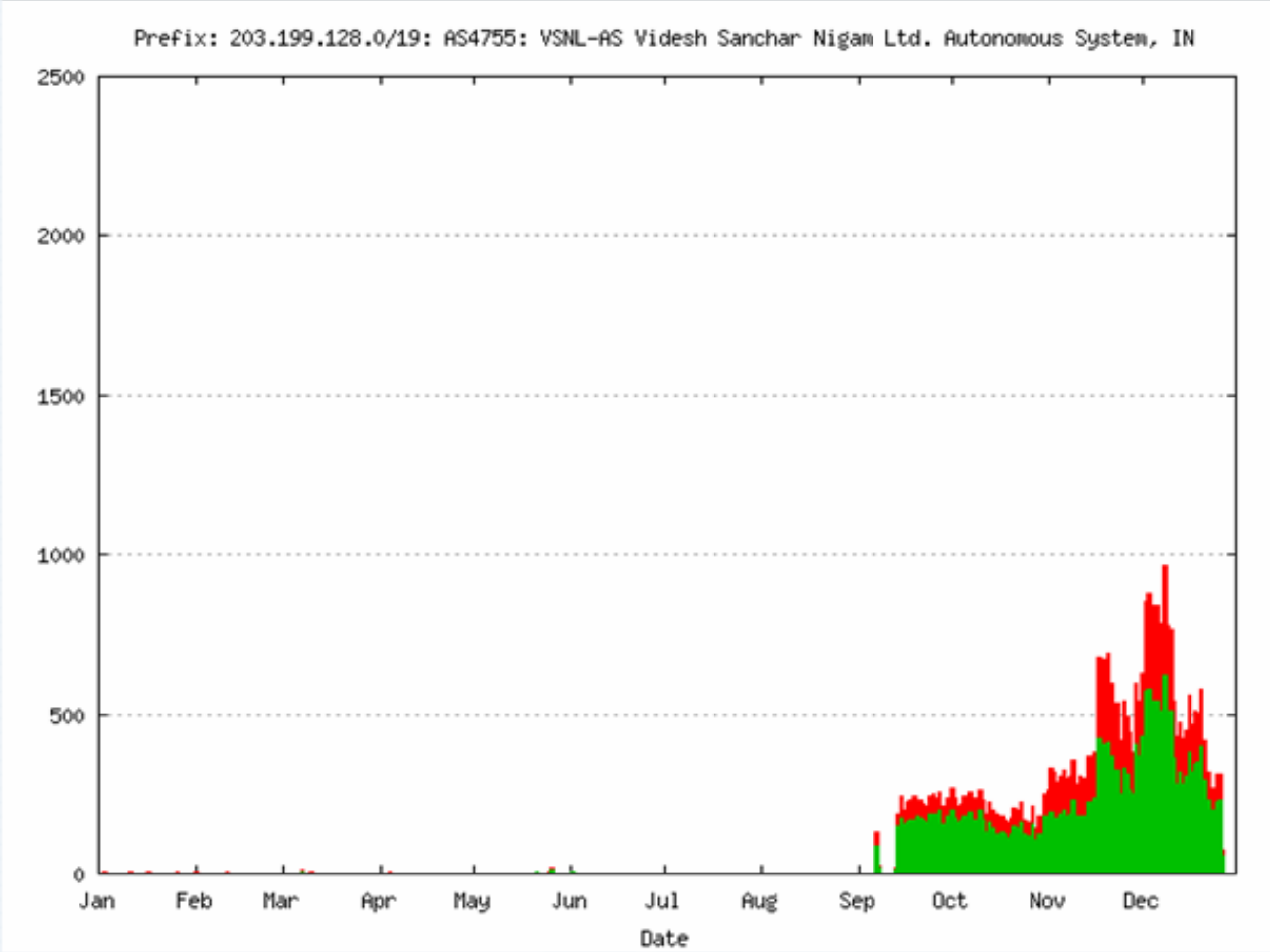
3 – 61.0.0.0/8



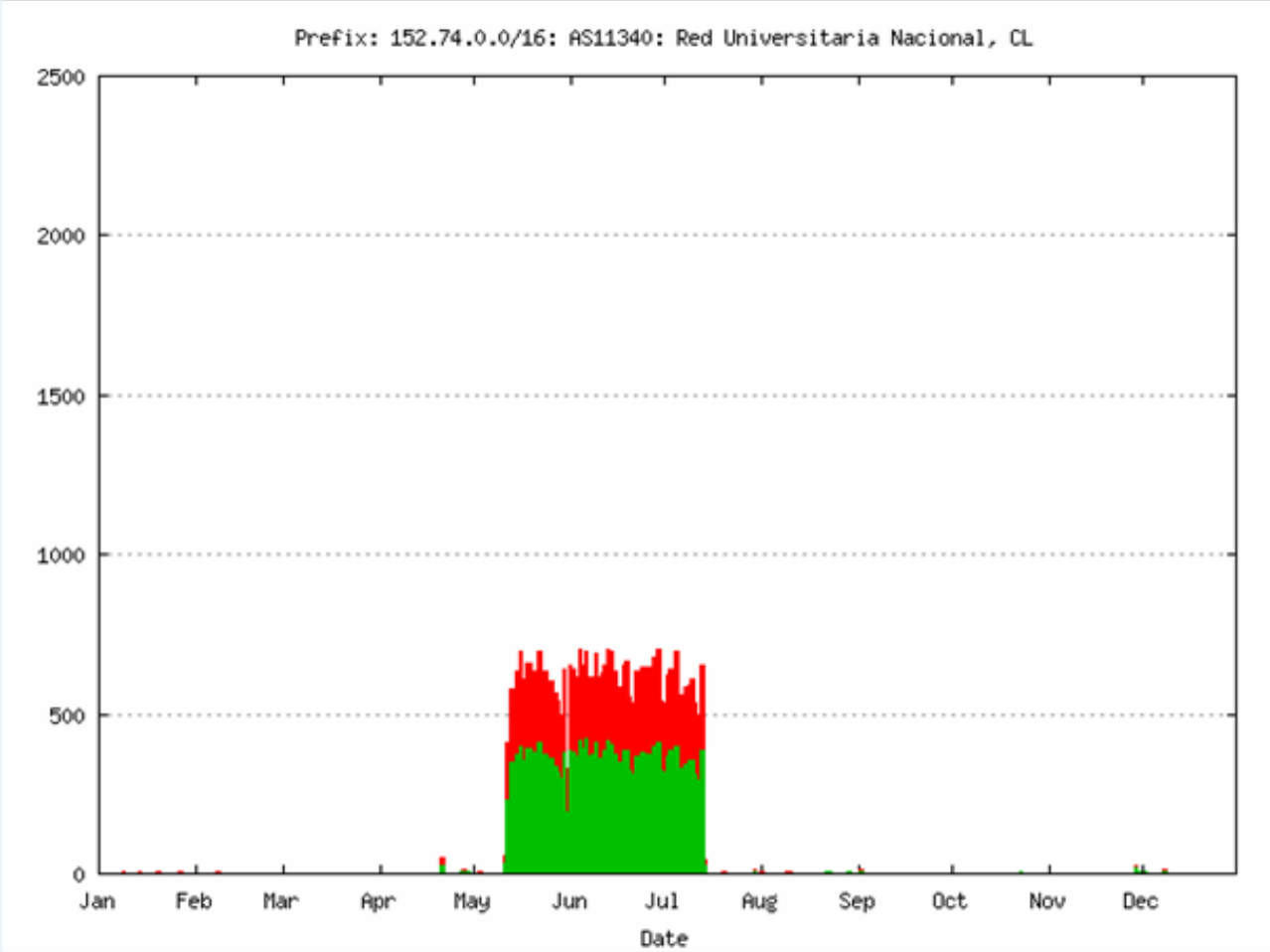
4 - 81.212.141.0/24



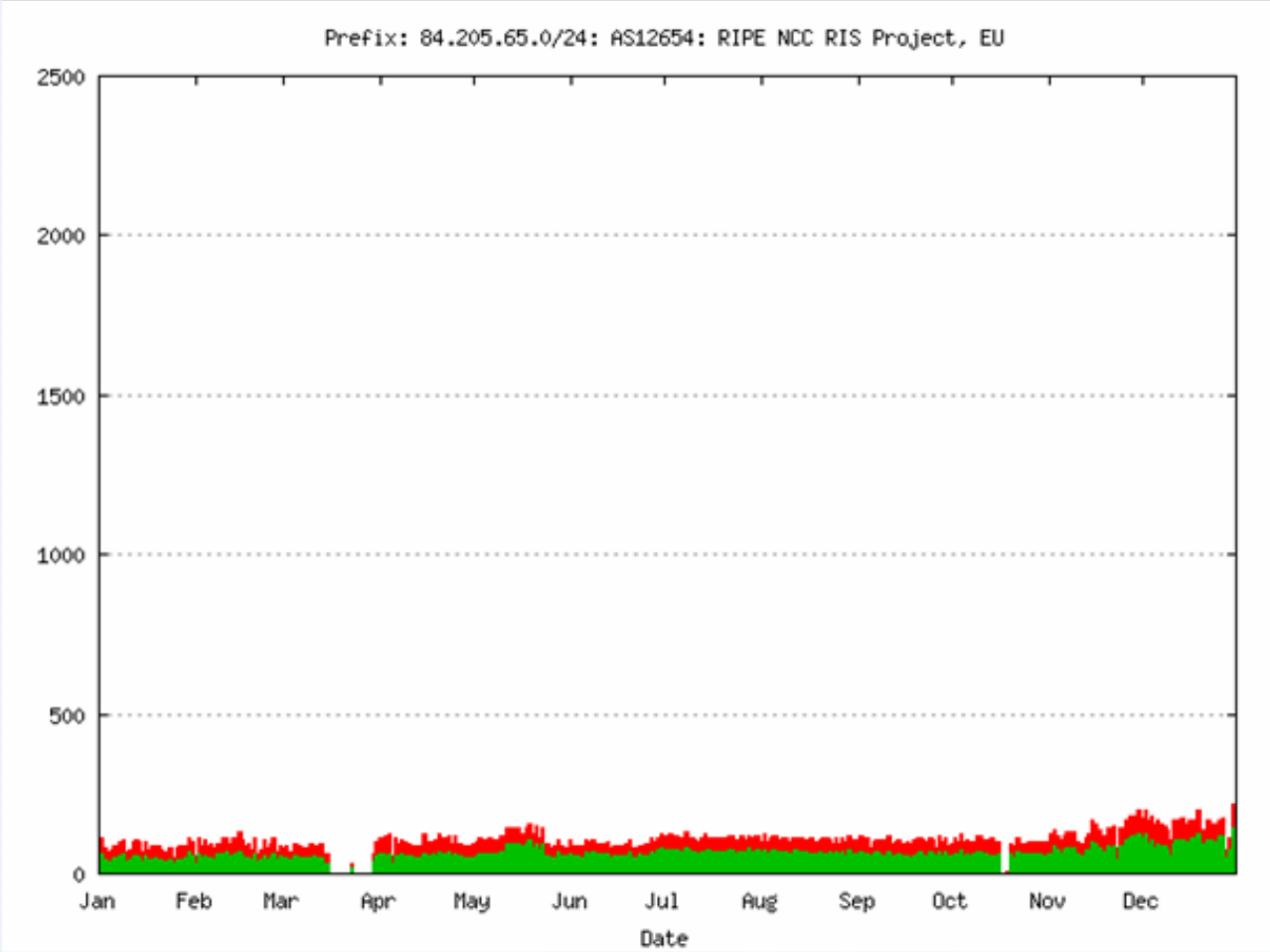
5 – 203.199.128.0/19



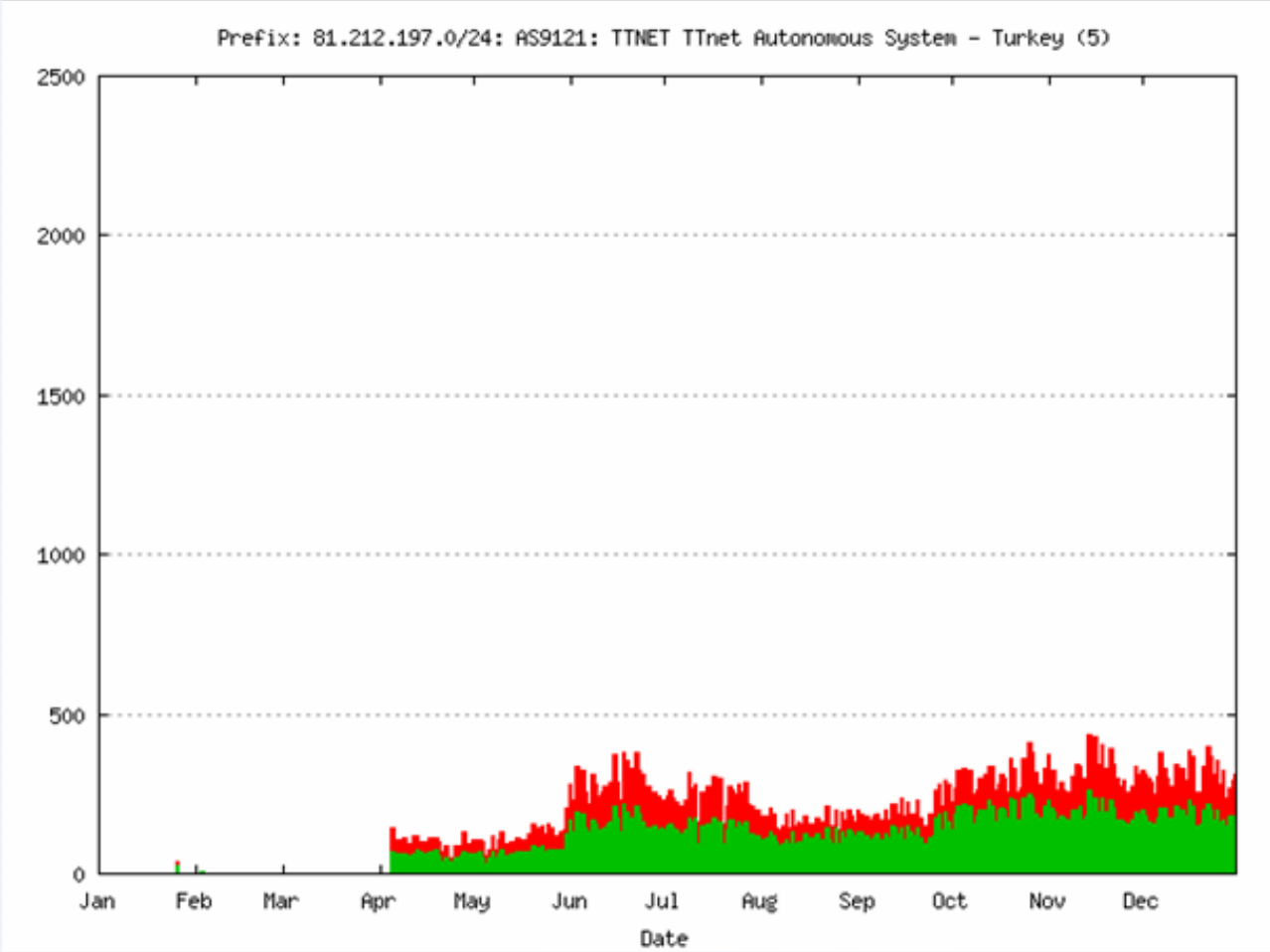
6 – 152,74.0.0/16



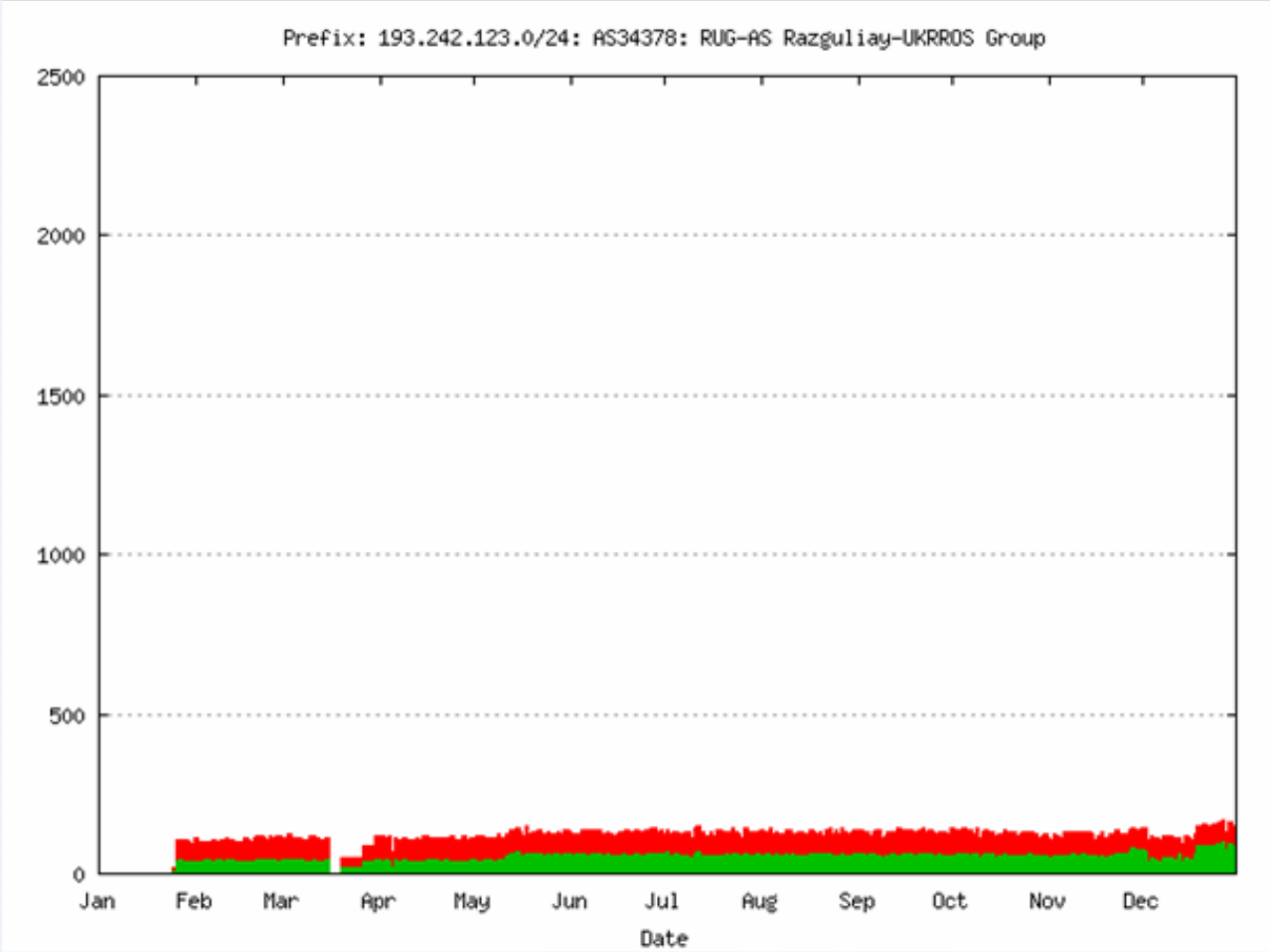
7 – 84.205.65.0/24



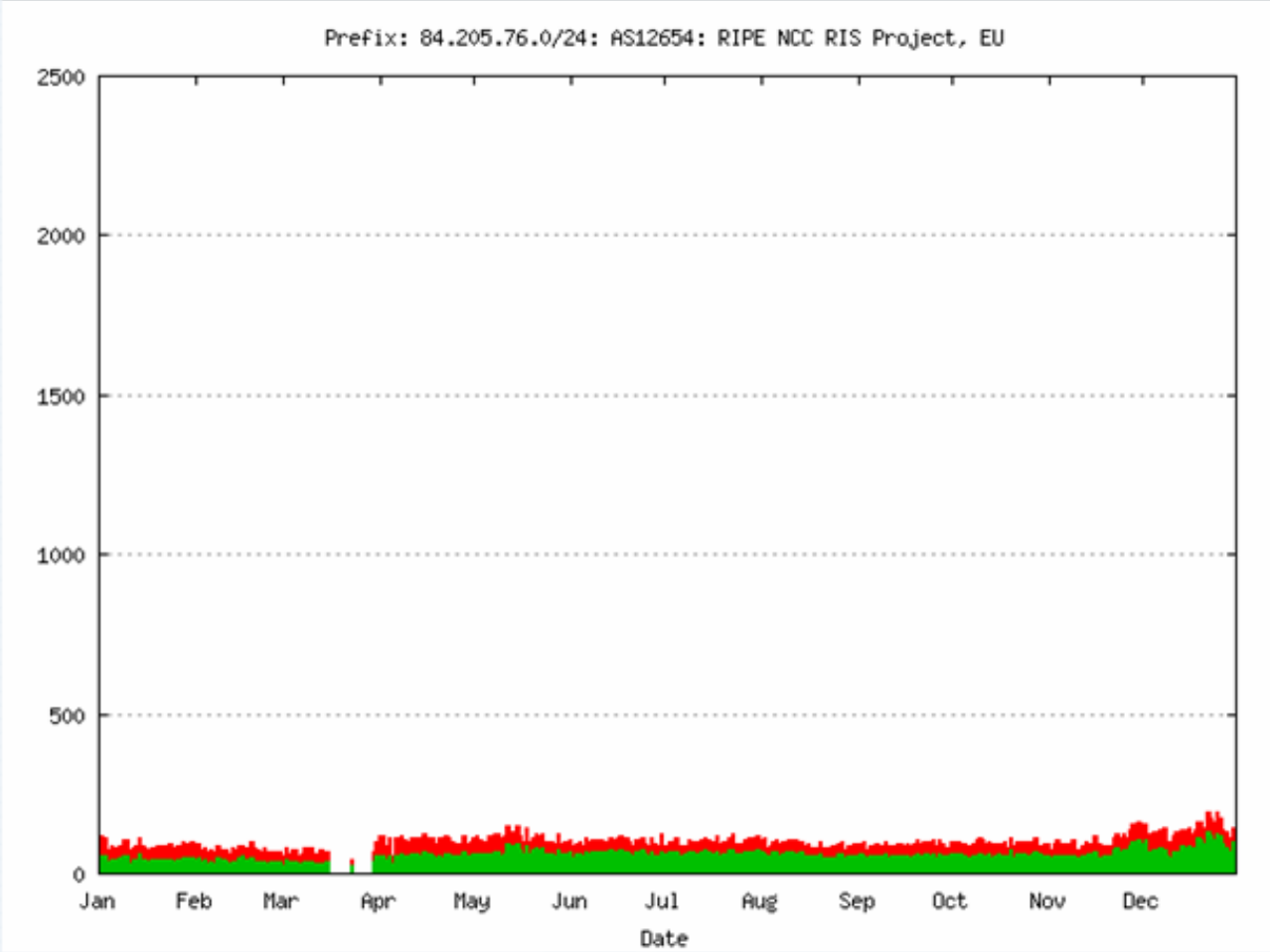
8 - 81.212.149.0/24



9 – 193.242.123.0/24



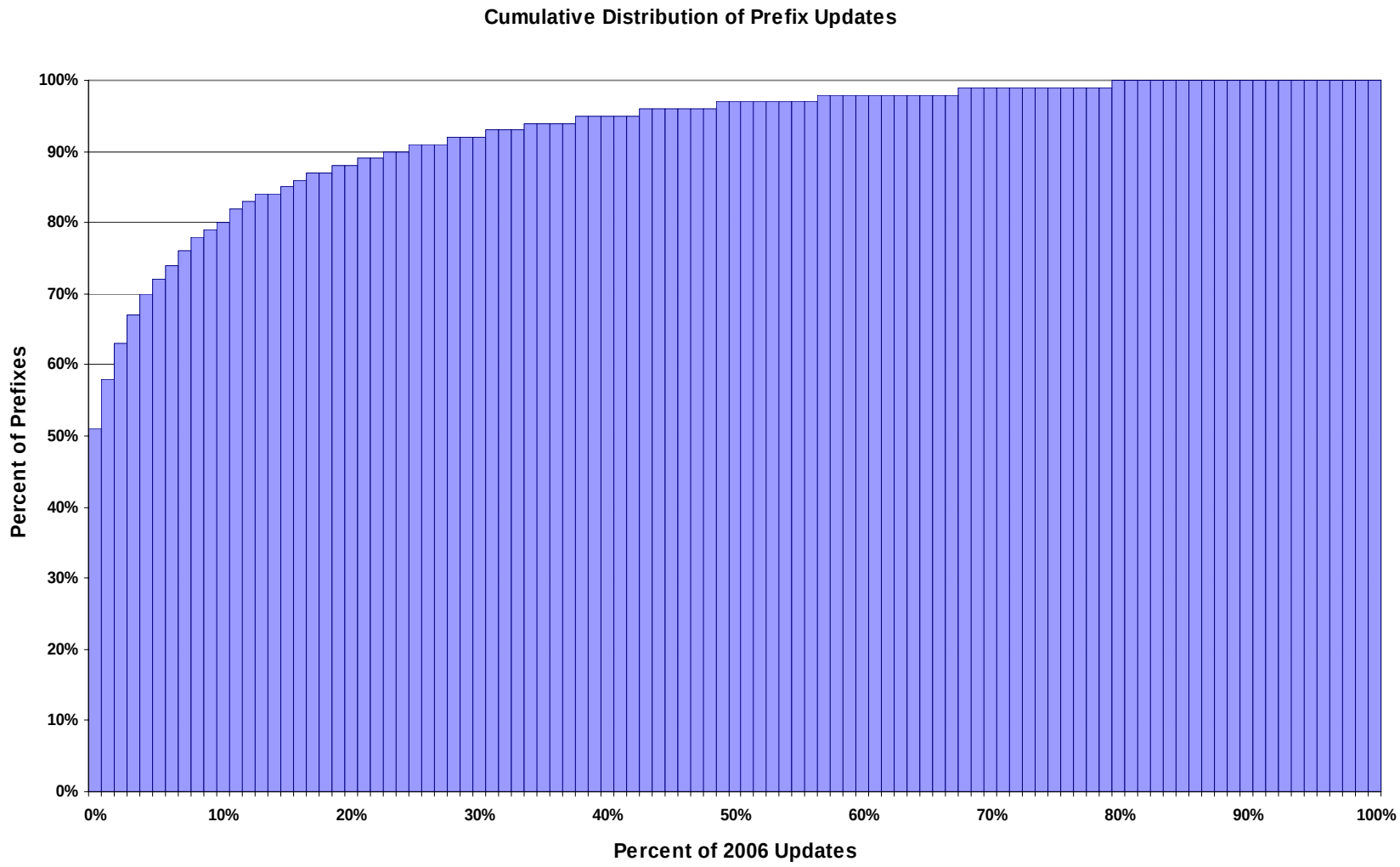
10 – 84.205.76.0/24



10 – 84.205.76.0/24

- Beacon on a 1 hour cycle
  - 12 UP and 12 DOWN per 24 hours
  - 4,380 beacon events in 2006
  - 55,634 BGP update events in 2006
  - 4,423 withdrawals
  - 1 Withdrawal at origin caused an average of 11 update messages prior to withdrawal propagation reaching the BGP observer

Distribution of Updates by AS

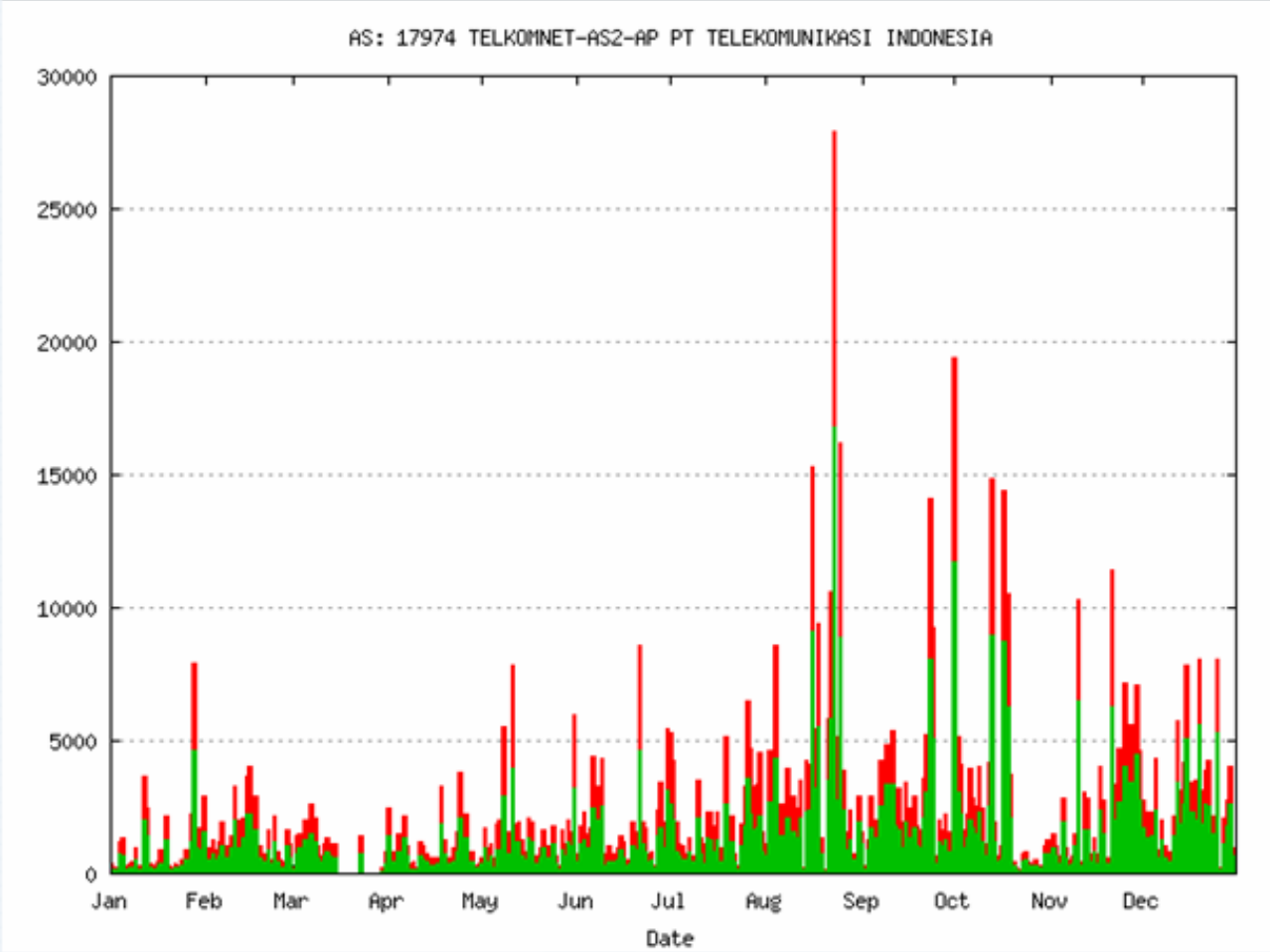


Active ASNs

Top 10 AS

	AS	Updates	Flaps	Re-Homes
1.	17974	1,340,344	983,667	1,819
2.	9121	783,879	542,965	199,409
3.	855	748,611	489,035	8,484
4.	702	517,723	379,880	96,121
5.	15611	517,243	337,669	2,556
6.	8151	425,852	288,042	33,666
7.	12654	396,924	295,083	49,567
8.	4323	393,687	275,477	130,056
9.	4621	370,478	278,650	21,116
10.	17557	368,689	248,680	121,211

1 – AS 17974



AS17974 Upstreams

- 17974 TELKOMNET-AS2-AP PT TELEKOMUNIKASI INDONESIA
  Adjacency: Upstream: 1 Downstream: 0
- Upstream Adjacent AS list

  AS7713 TELKOMNET-AS-AP PT TELEKOMUNIKASI INDONESIA

  Upstream: 5

  AS9237 HUTCHCA-AS Corporate Access (HK) Ltd.

  AS11919 LORAL-SKYNET-AR - Loral Skynet Network Services, Inc.

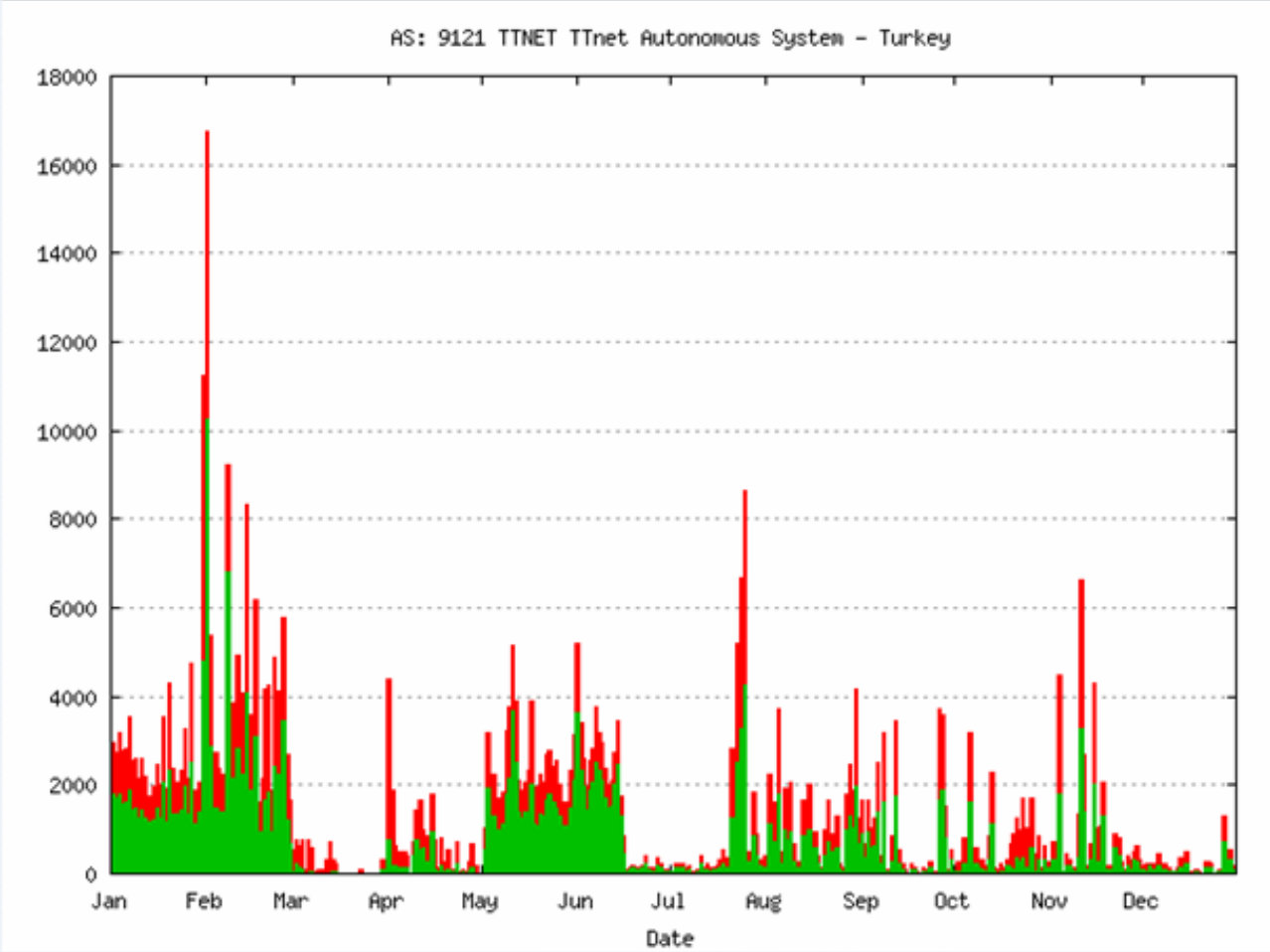
  AS24077 TMHK-TRANSIT-AS-HK-AP TMHK Global Transit

  AS7473 SINGTEL-AS-AP Singapore Telecom

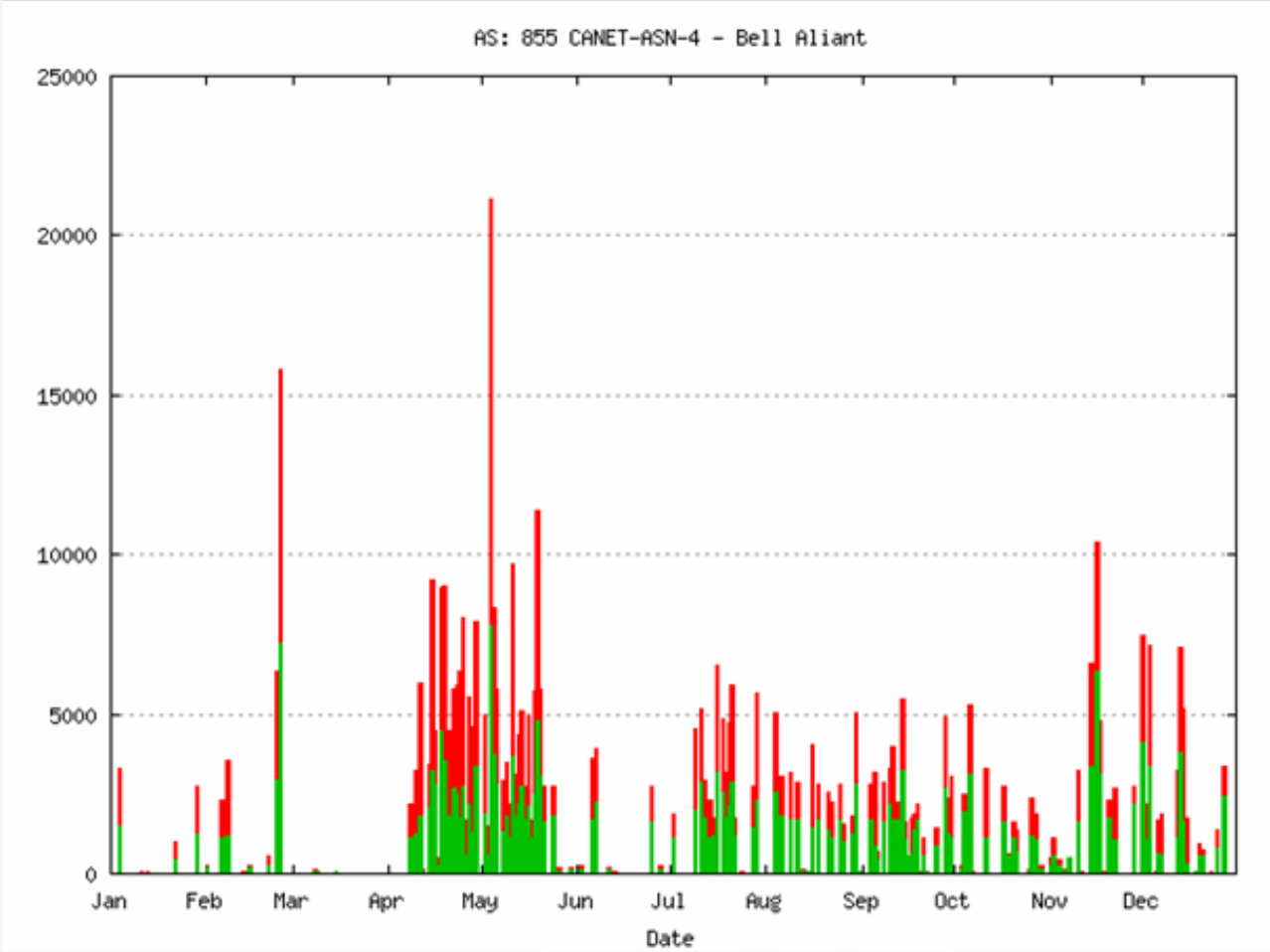
  AS7632 MEGHANTARA-AS-AP PT. Meghantara

  Traffic engineering across multiple upstreams?

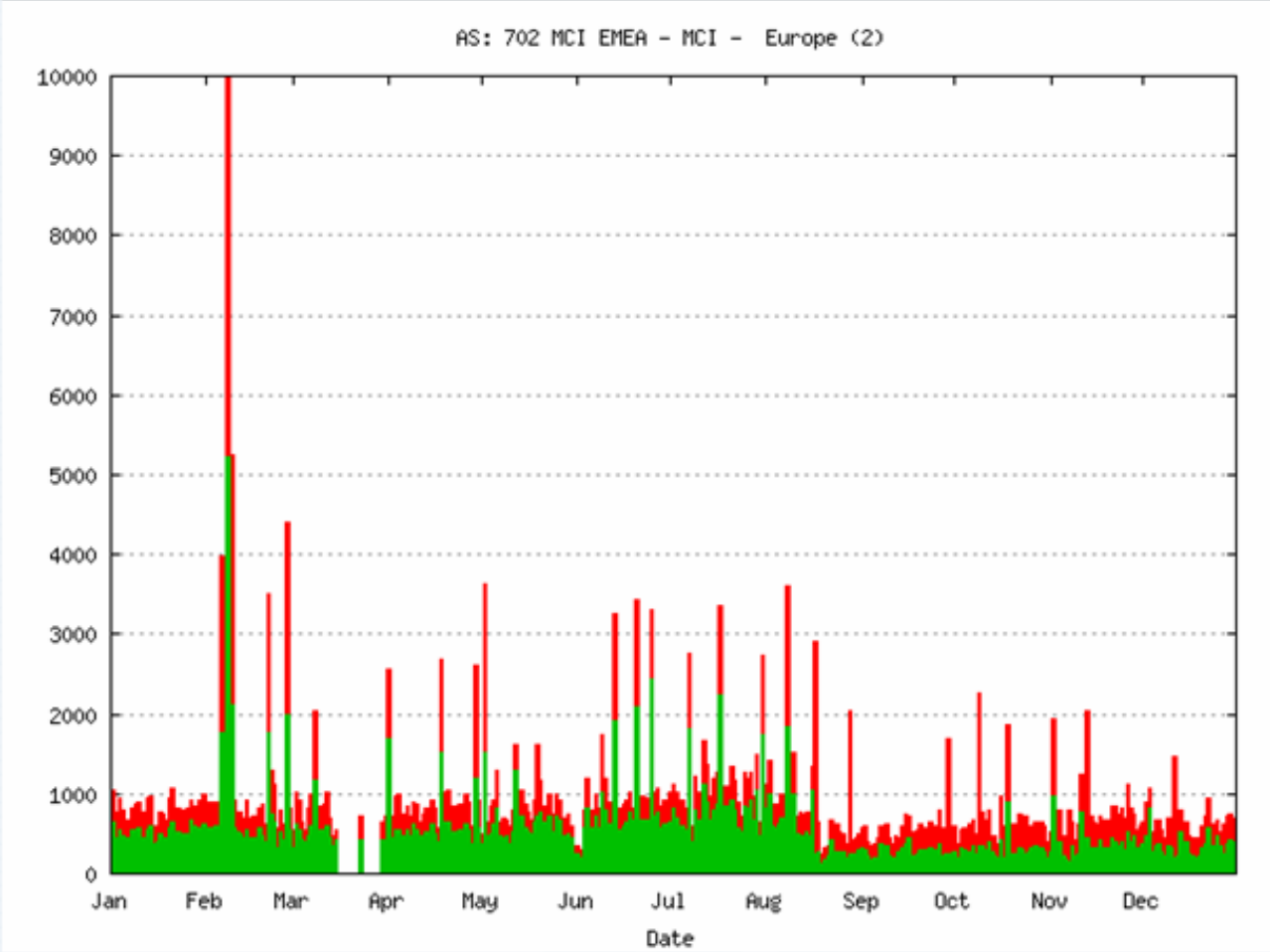
2 – AS 9121



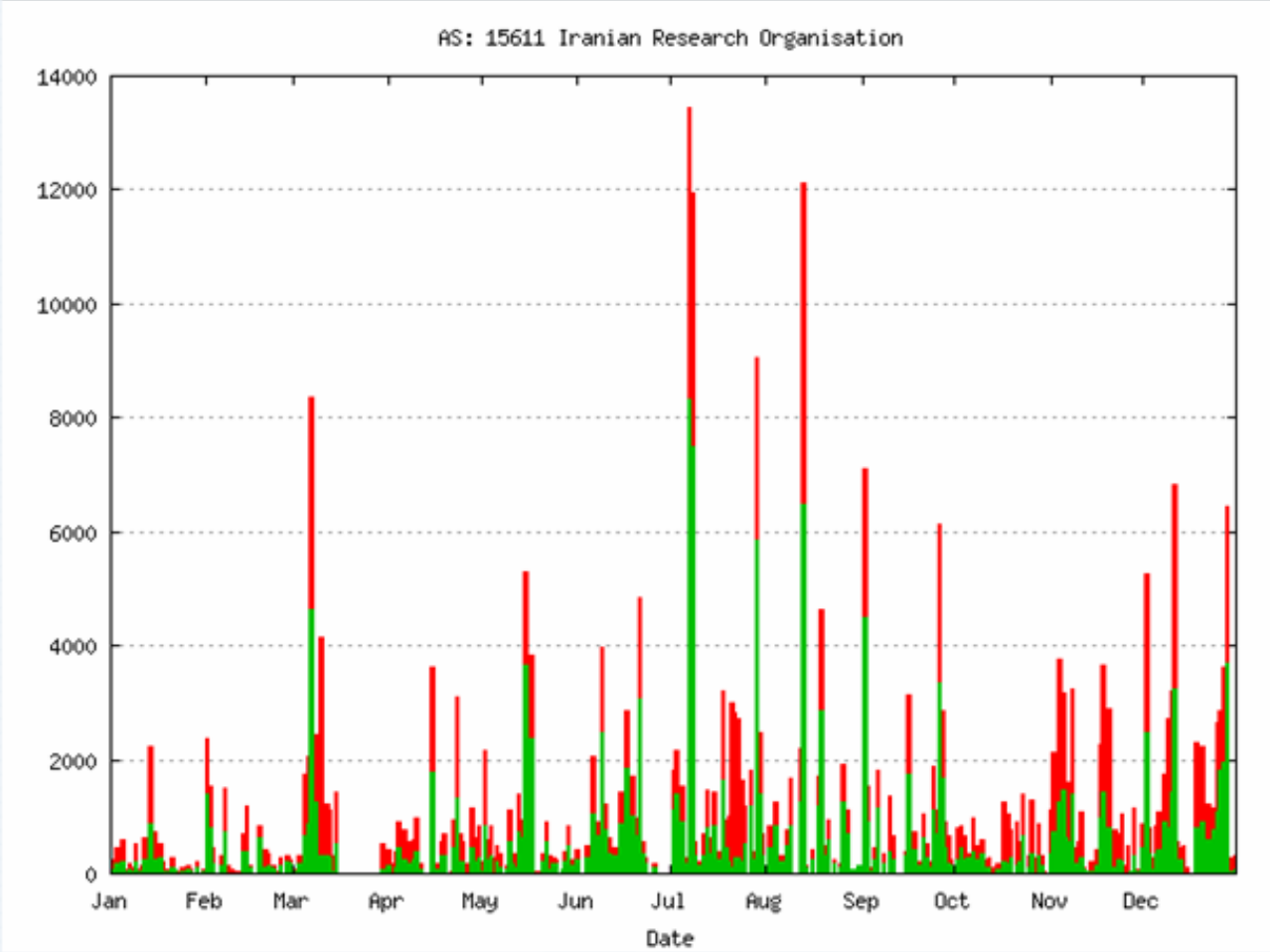
3 – AS 855



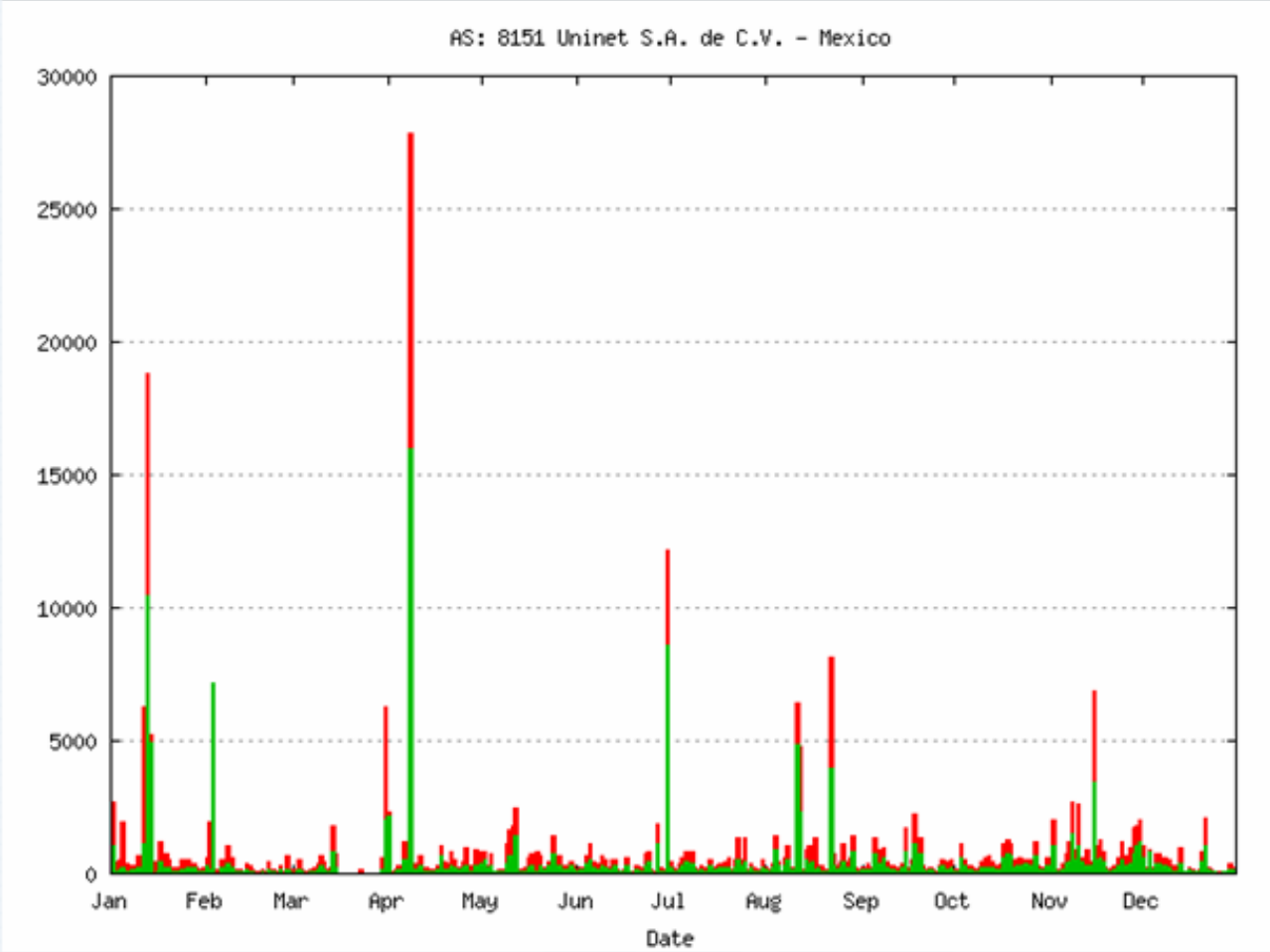
4 – AS 702



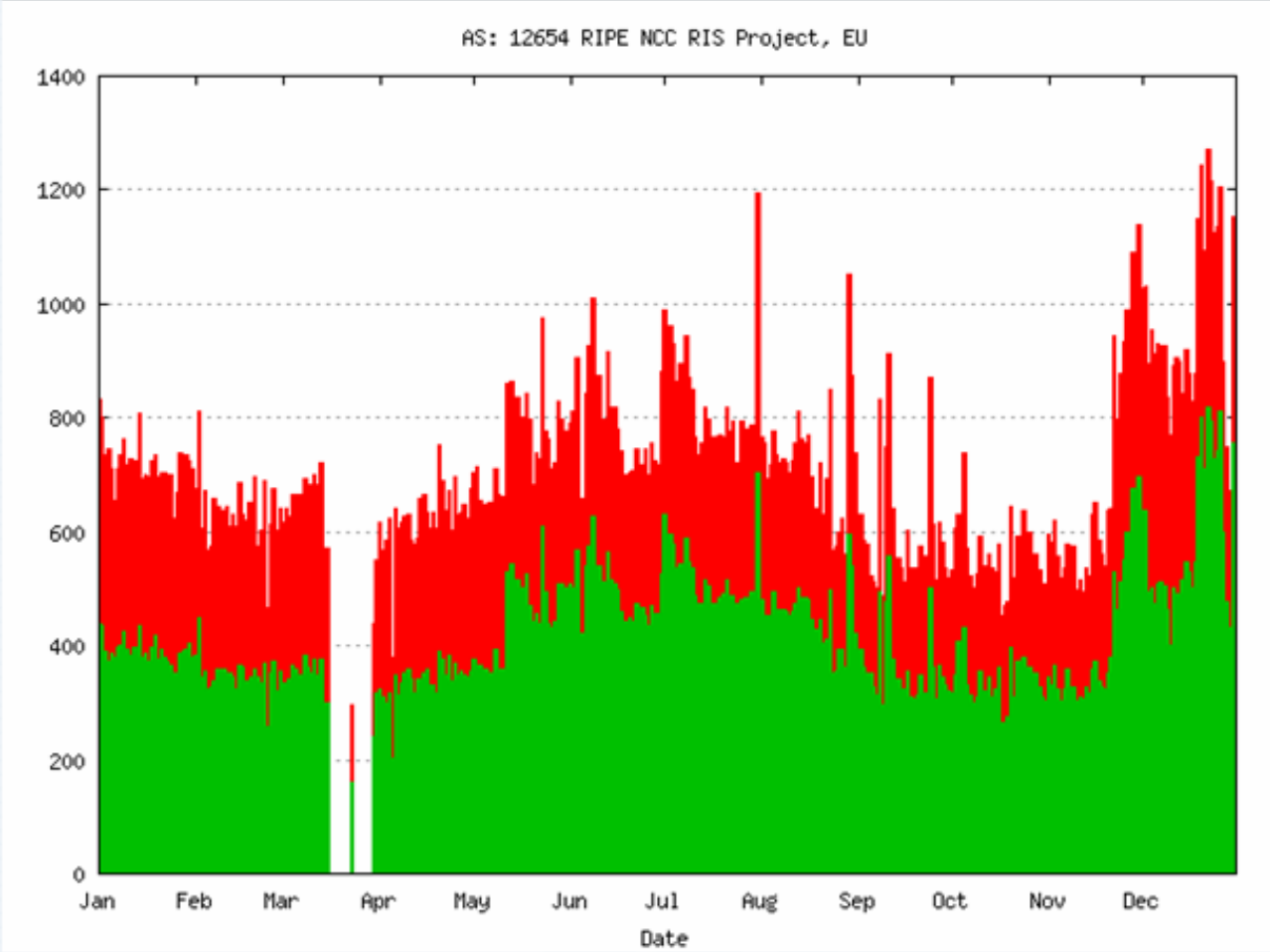
5 – AS 15611



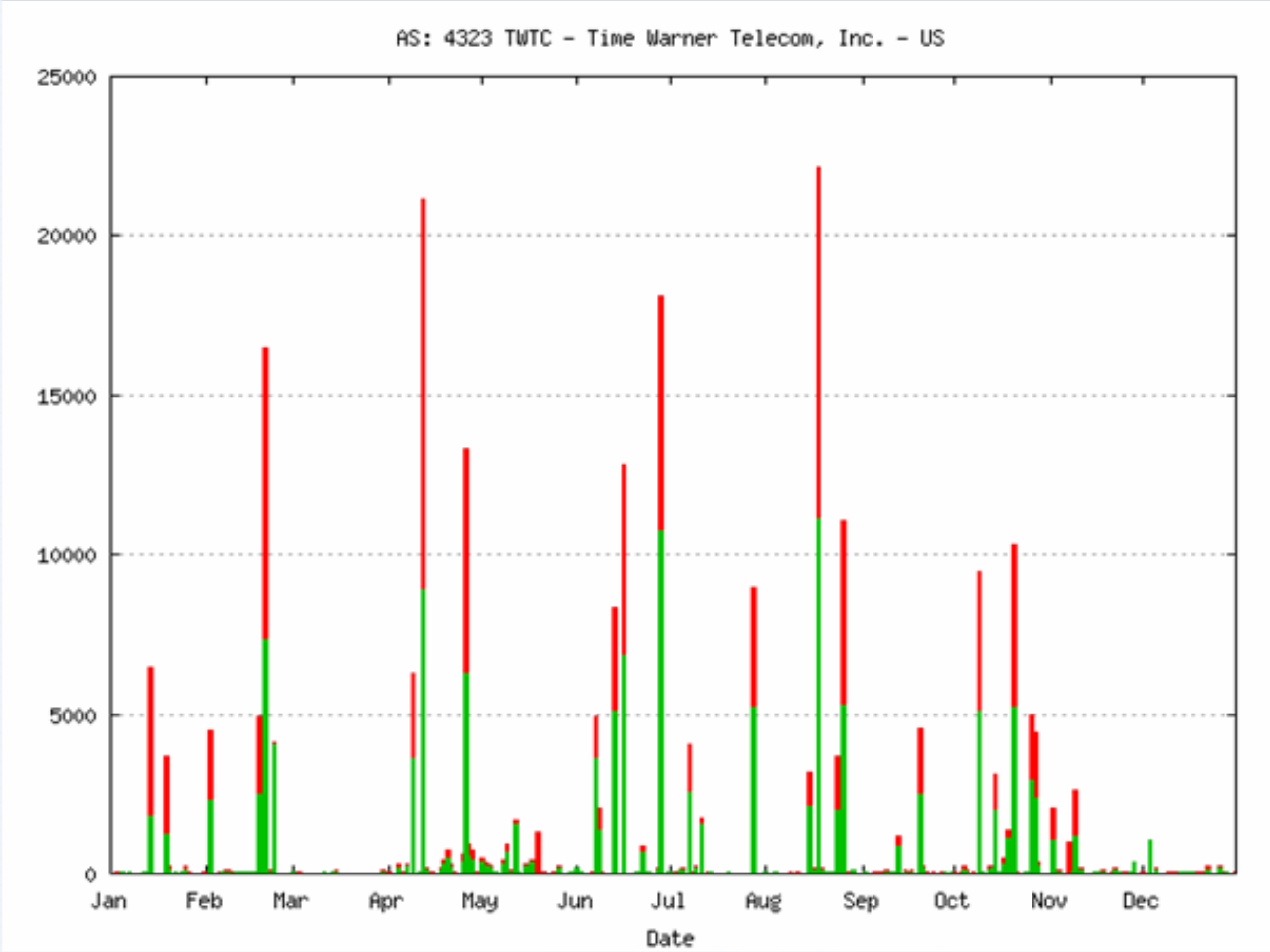
6 – AS 8151



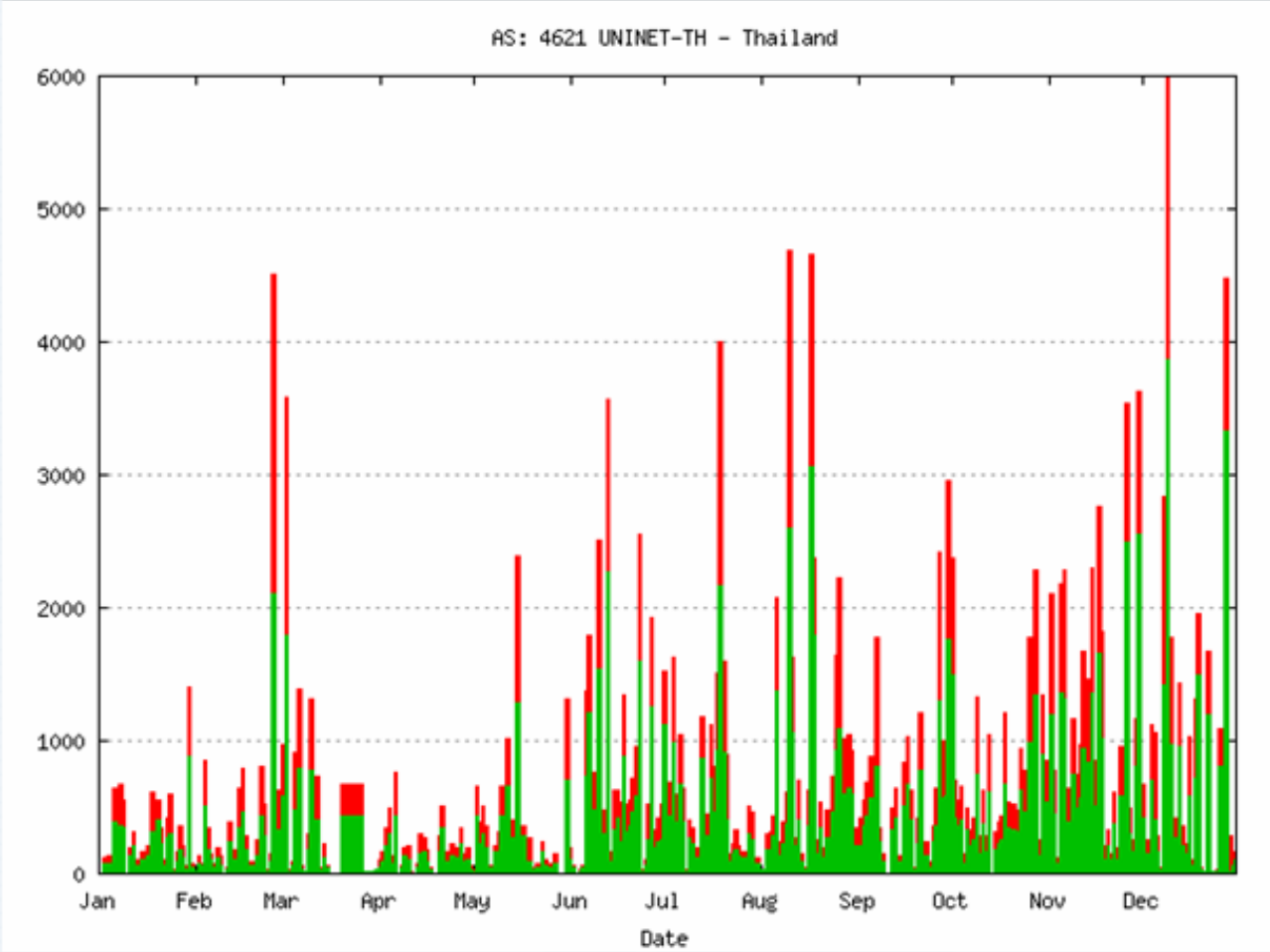
7 – AS 12654



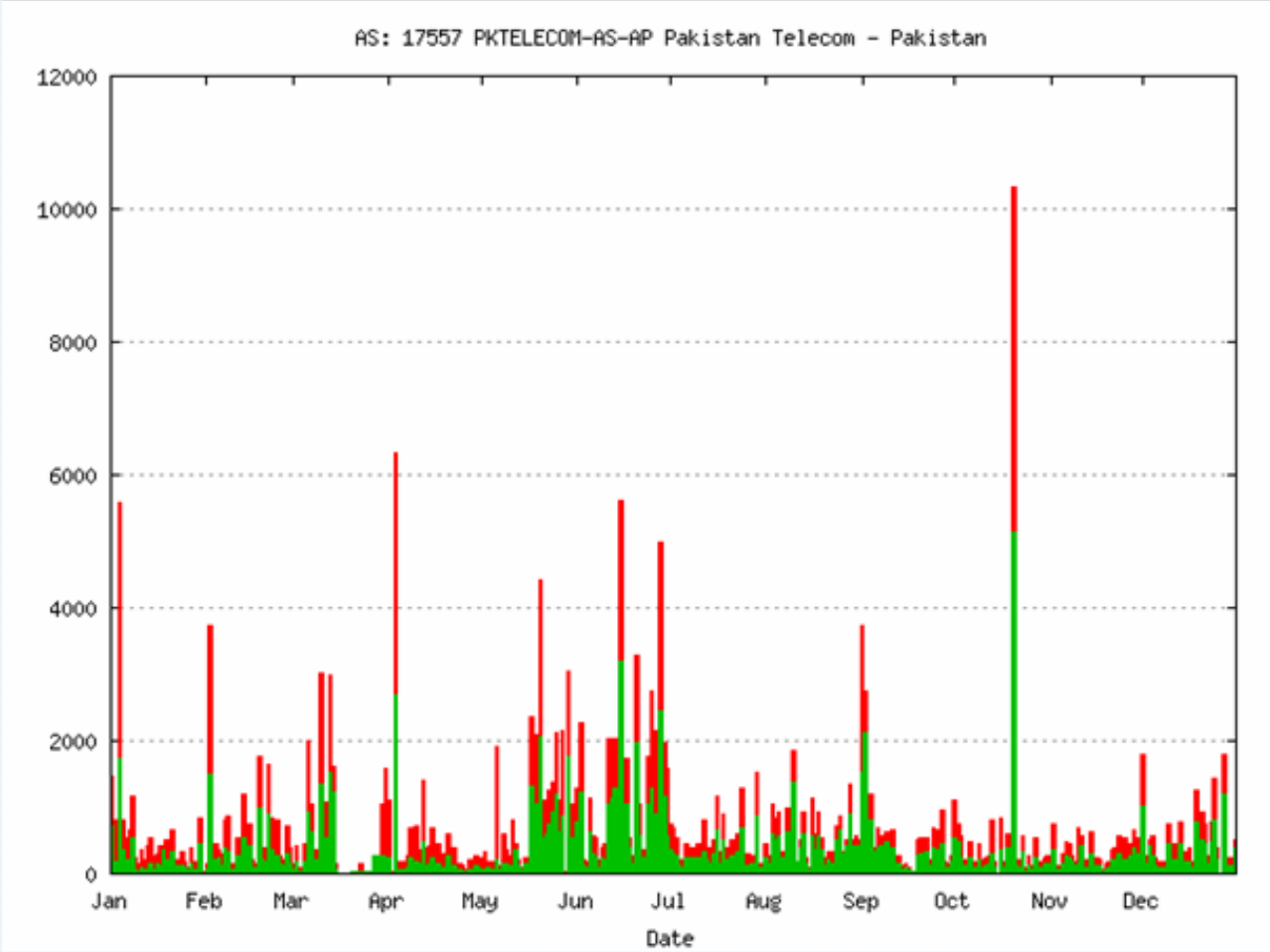
8 – AS 4323



9 – AS 4621



10 – AS17557



So what's going on?

- It would appear that the BGP update rate is being strongly biased by a small number of origins with two forms of behaviour:
  - Traffic Engineering - consistent update rates sustained over weeks / months with a strong component of first hop change and persistent announce and withdrawal of more specifics
  - Unstable configuration states – a configuration which cannot stabilise and for a period of hours or days the update rate is extremely intense

How “good” is this data?

- Its just one (ordinary) router’s view of a rather complex routing world, not an aggregated view of a larger routing environment. There is some ‘locality’ component in the data.
- Its not located the within the world’s richest connectivity (it may be understating the routing load)
- The data is very noisy (e.g. 150,000 short term (leaked?) prefixes)
- The data is heavily skewed by a ‘heavy tail’ distribution (small number of prefixes and ASs appear to be the subject of a large number of updates)
- So any effort at generating trend data is biased by the small number of these “intense updaters” (making predictive models even more uncertain than normal)

So what's going on?

- It would appear that the BGP update rate is being strongly biased by a small number of origins with two forms of behaviour:
  - Supernova

    Multi-Homing & Traffic Engineering - bursting update rates sustained over weeks / months with a strong component of first hop change and persistent announce and withdrawal of more specifics
  - Background Radiation

    Unstable configuration states – a configuration which cannot stabilise and for a period of hours or days where the convergence to withdrawal causes continual updates

Where is this heading?

- Can we make BGP “scale” better or are we forced to look at a new routing structure?
- Making BGP “scale”:
  - Is there a more effective mechanism for damping unstable routes and paths and /or damping convergence to withdrawal?
  - Can we encourage widespread use of standard mechanisms that limit the propagation of BGP advertisements?
  - Should we consider alternate ways of BGP coping with withdrawal?
    - Does the “origin withdrawal” attribute added to BGP protocol specification make sense?
    - Should we consider “alternate reachability” selective advertisements that address withdrawal / update patterns in BGP convergence by changing the BGP protocol behaviour?

Changing BGP

- It's now a large system with massive deployment inertia
- Any 'change" will require piecemeal deployment capability with benefits realized by those who deploy
  - Which implies that use of backward compatible incremental change with piecemeal deployment is the only feasible approach here
  - The 32-bit ASN transition is a useful case study in changing BGP
    - Capability negotiation for peer setup
    - Transitive opaque attributes to signal additional capabilities (such as origin withdrawal)
    - Local changes to BGP processing

Some themes for further study

- How well do we understand BGP today?
  - More observation points
  - Investigation of known BGP pathologies
  - Control points as observation benchmarks
- How well do we understand the BGP of tomorrow
  - What metrics provide reasonable indicators?
  - How stable is the time series data?
  - What is the confidence interval of 3 – 5 year predictors
- How well do we understand the impacts of incremental change to BGP?
  - Modelling connectivity and behaviours
  - Simulation and direct experimentation

Thank You