2005 – A BGP Year in Review

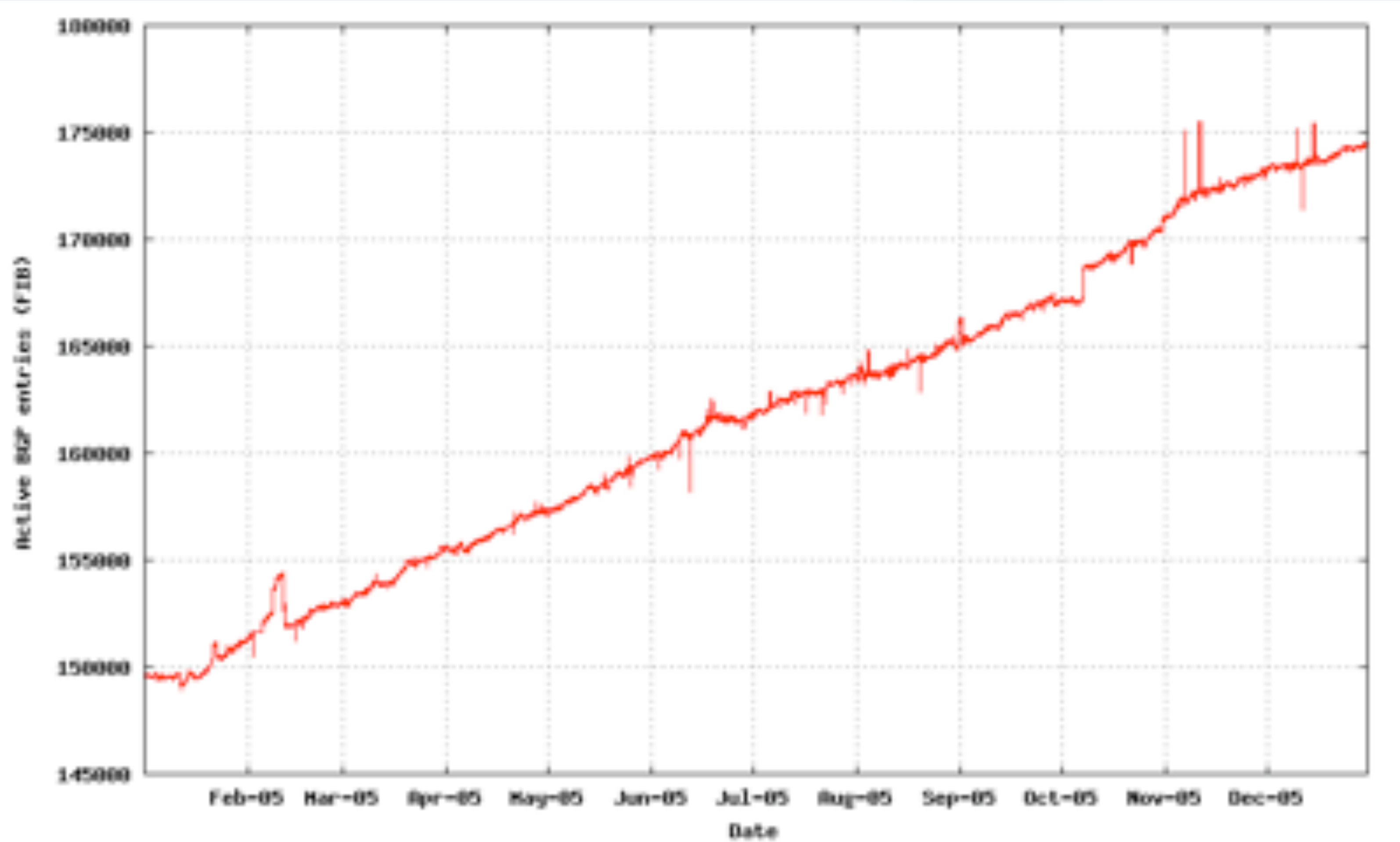
February 2006

Geoff Huston
Research Scientist
APNIC

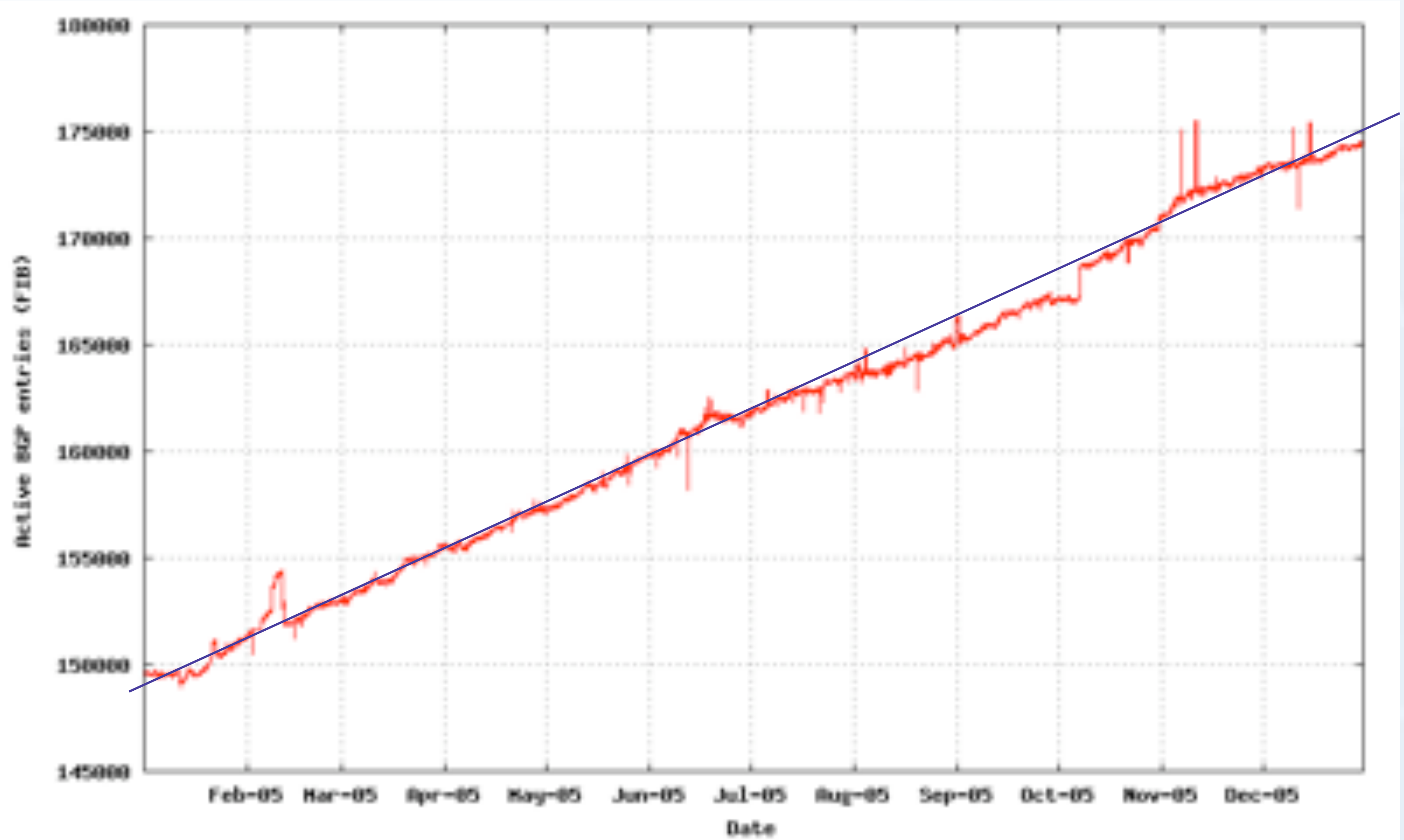
Thanks to

- Stephan Millet of Telstra for assisting with generation of some of the data sets that have been used in this presentation
  - Although any faults in the interpretation of the data are all mine!

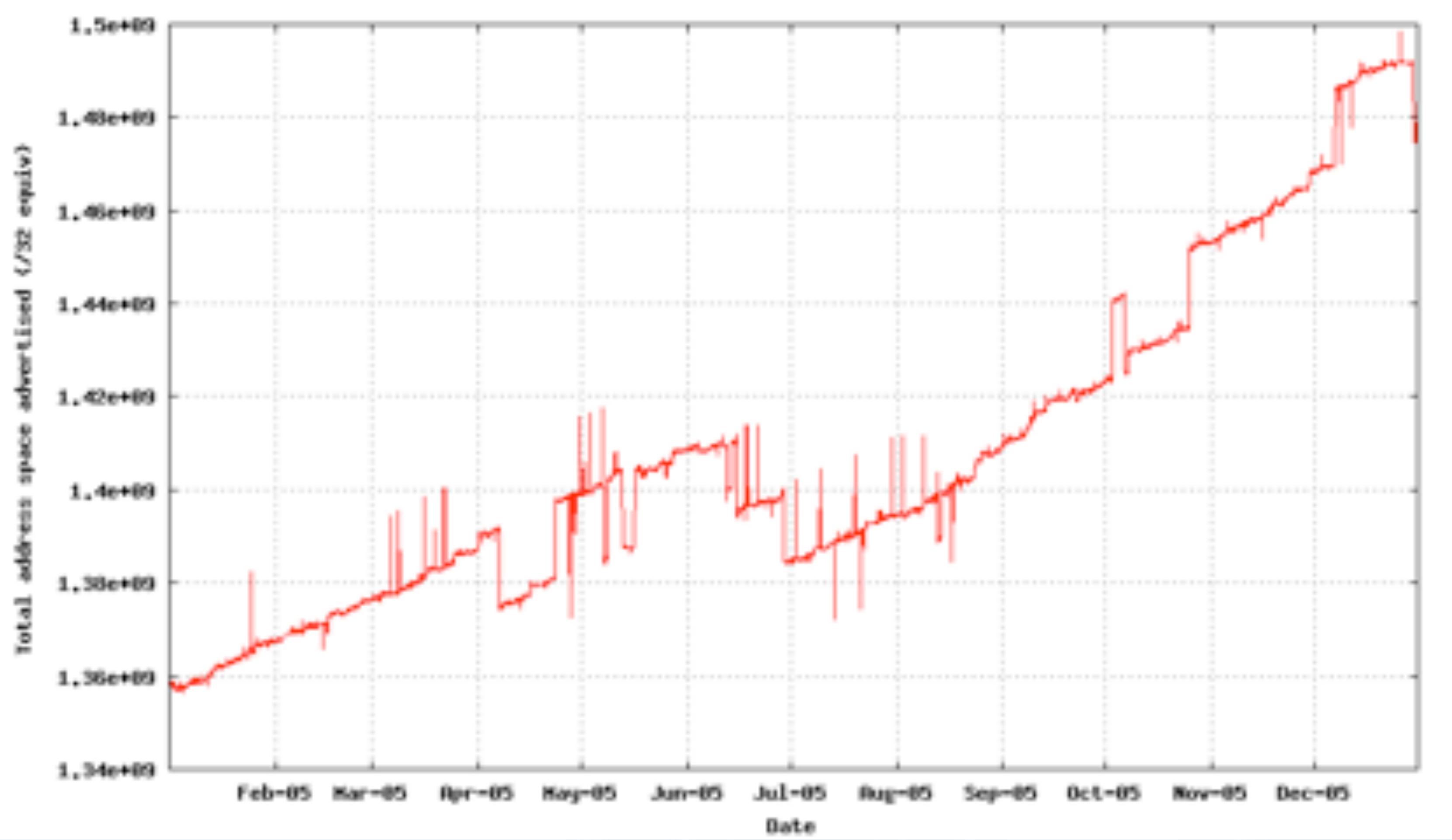
IPv4 in 2005
Total Advertised BGP Prefixes



IPv4 in 2005
Total Advertised BGP Prefixes

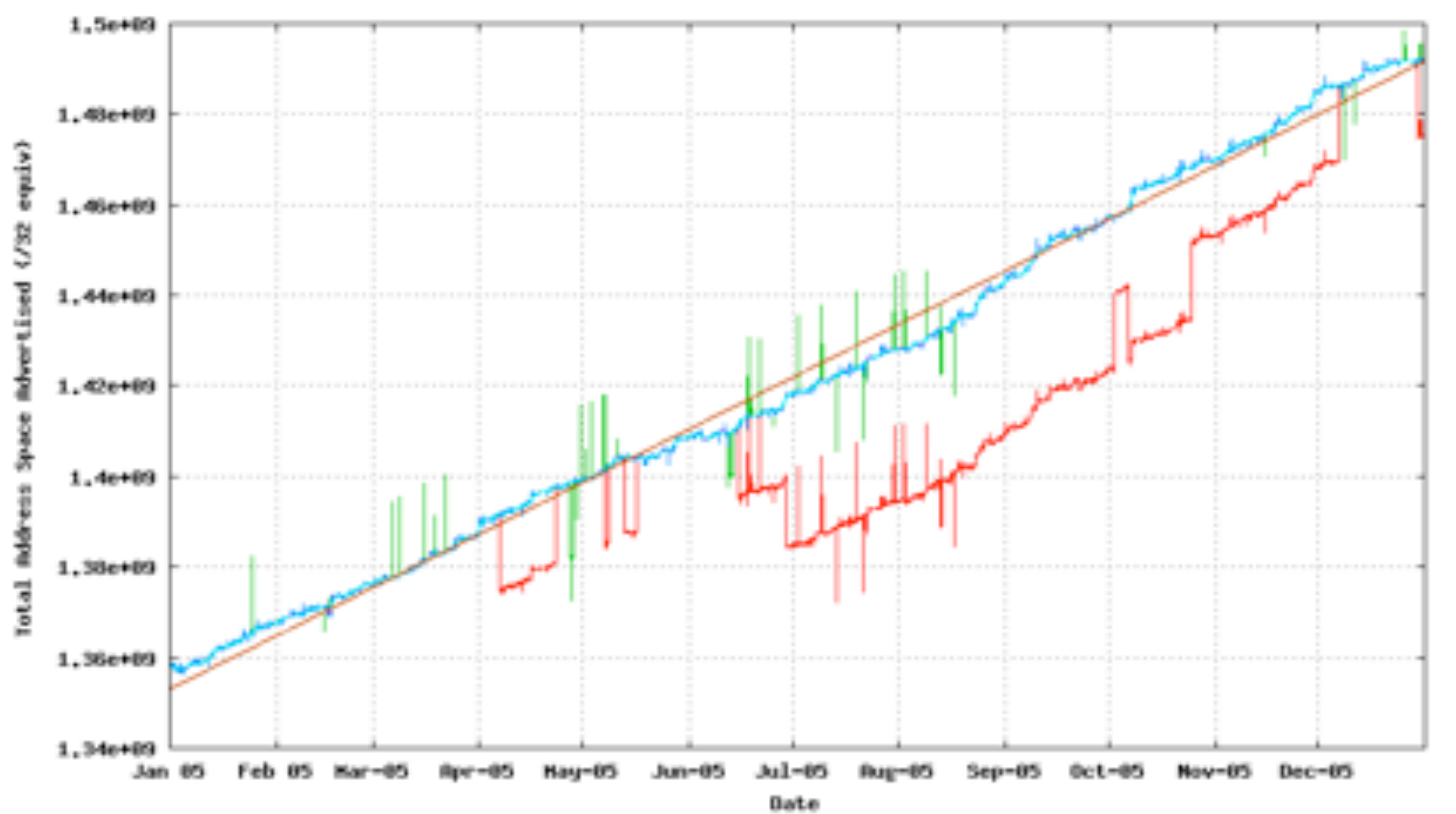


IPv4 in 2005
Total Advertised Address Span

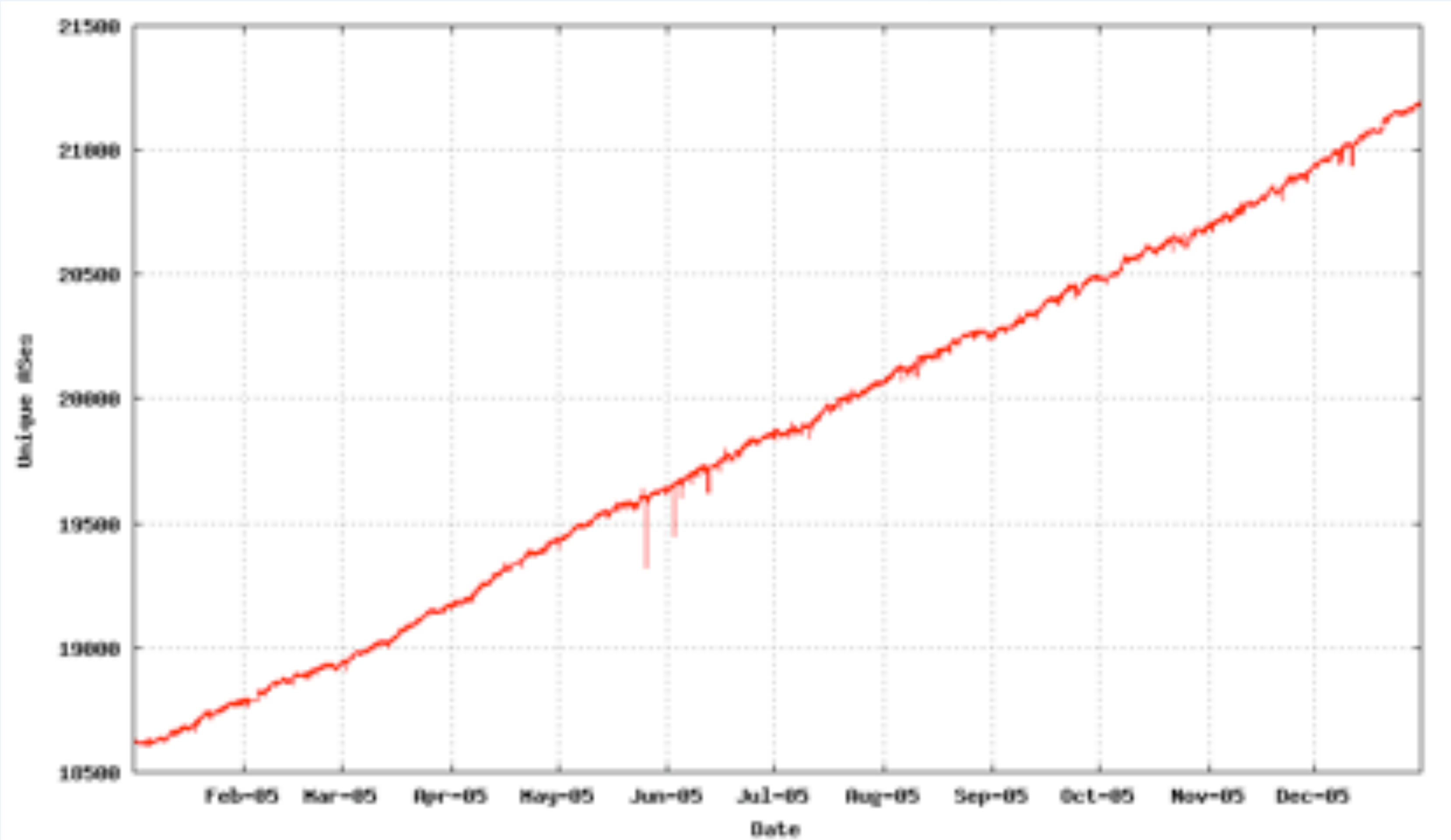


IPv4 in 2005

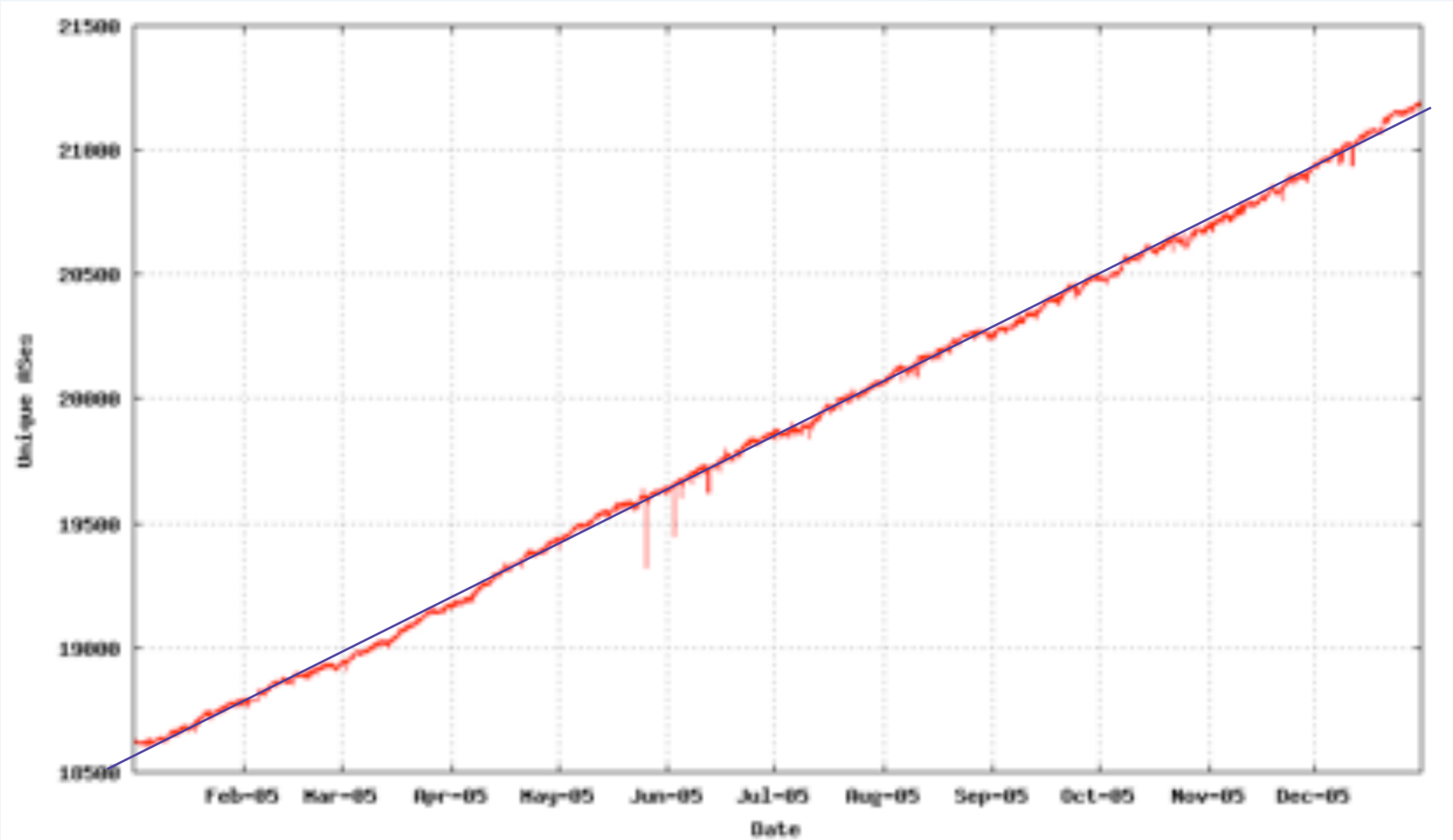
Total Advertised Address Span



IPv4 in 2005
Total Advertised AS Numbers



IPv4 in 2005
Total Advertised AS Numbers



IPv4 – Vital Statistics for 2005

Prefixes	148,500 – 175,400	+18%	26,900
Roots	72,600 – 85,500	+18%	12,900
Specifics	77,200 – 88,900	+18%	14,000
Addresses	80.6 – 88.9 (/8)	+10%	8.3 /8s
ASNs	18,600 – 21,300	+14%	2,600

Average advertisement size is getting smaller
Average address origination per AS is getting smaller
Average AS Path length steady at 3.5
AS interconnection degree up

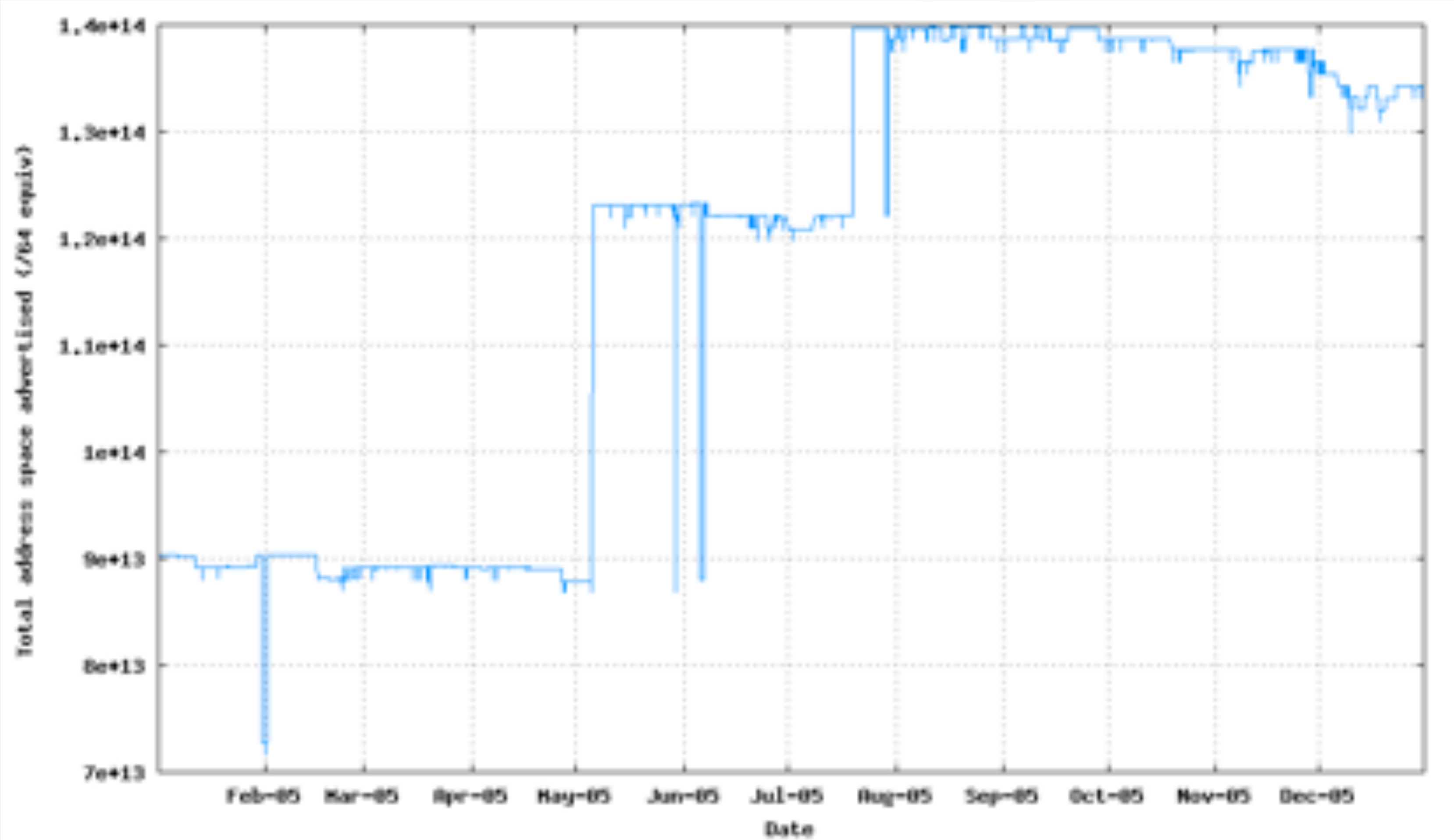
The IPv4 network continues to get denser, with finer levels of advertisement granularity.

More interconnections, more specific advertisements

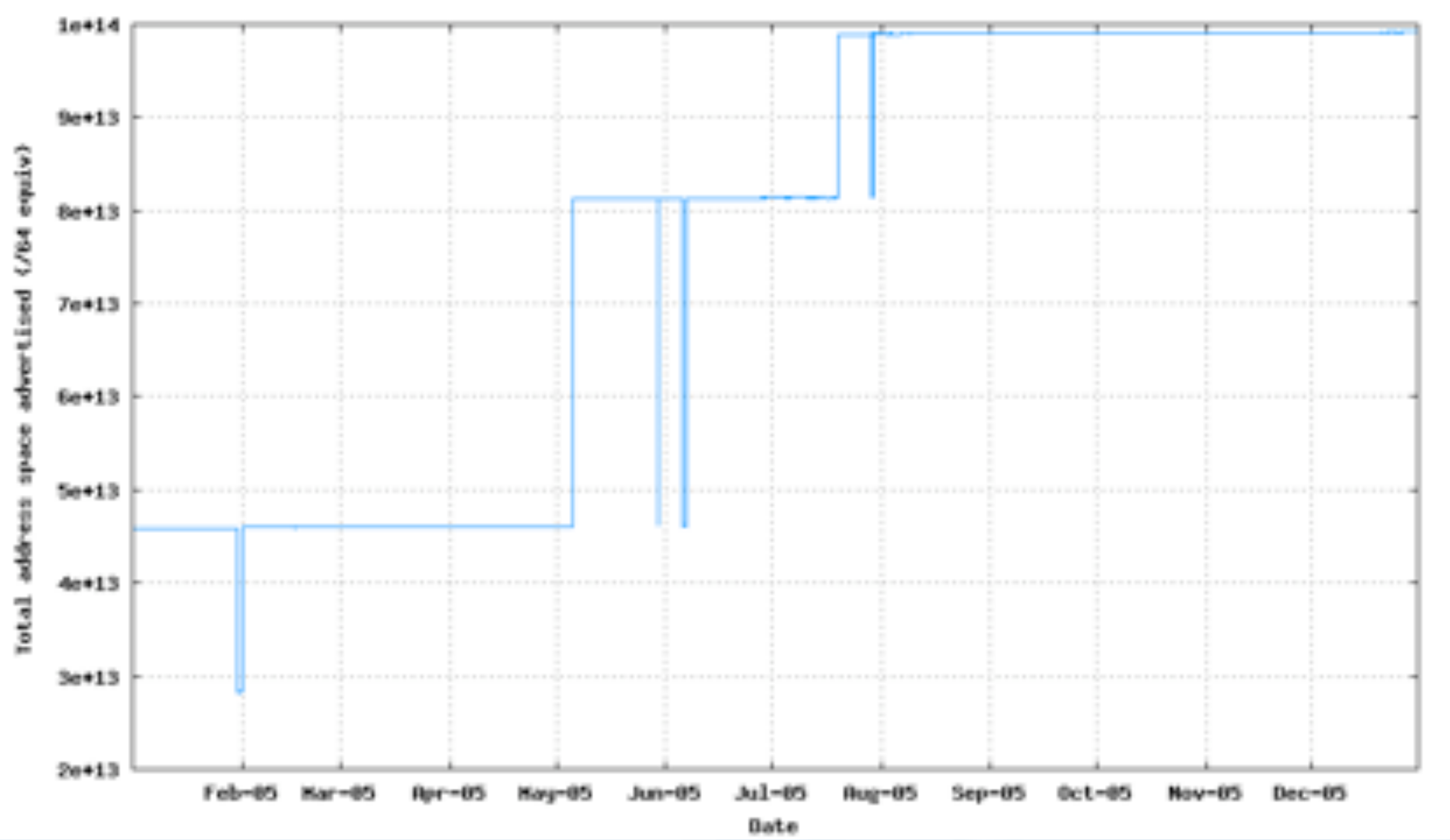
IPv6 in 2005
Advertised Prefix Count

IPv6 in 2005
Advertised Prefix Count

IPv6 in 2005
Advertised Address Span

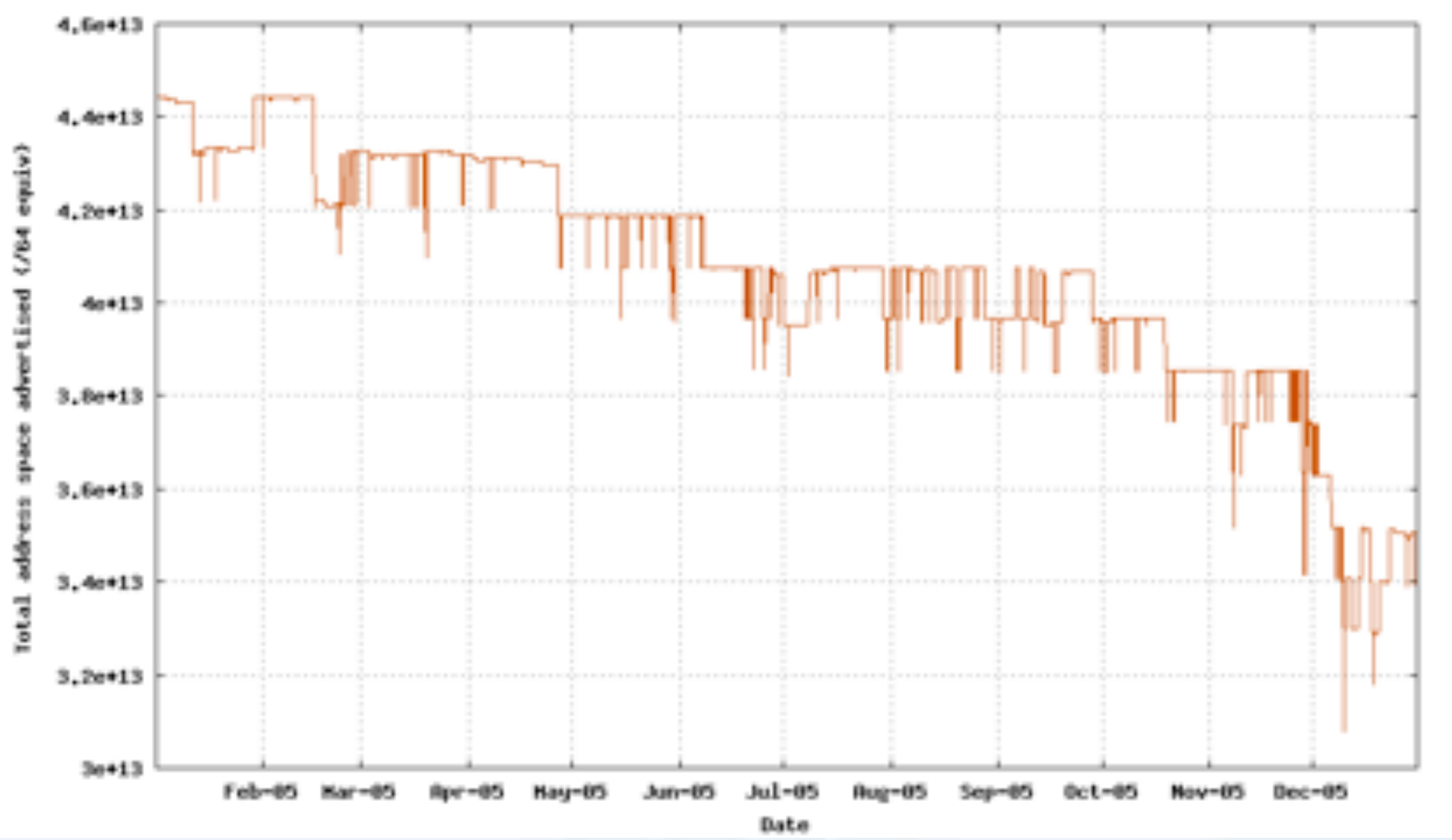


IPv6 in 2005
Advertised Address Span w/o 6Bone

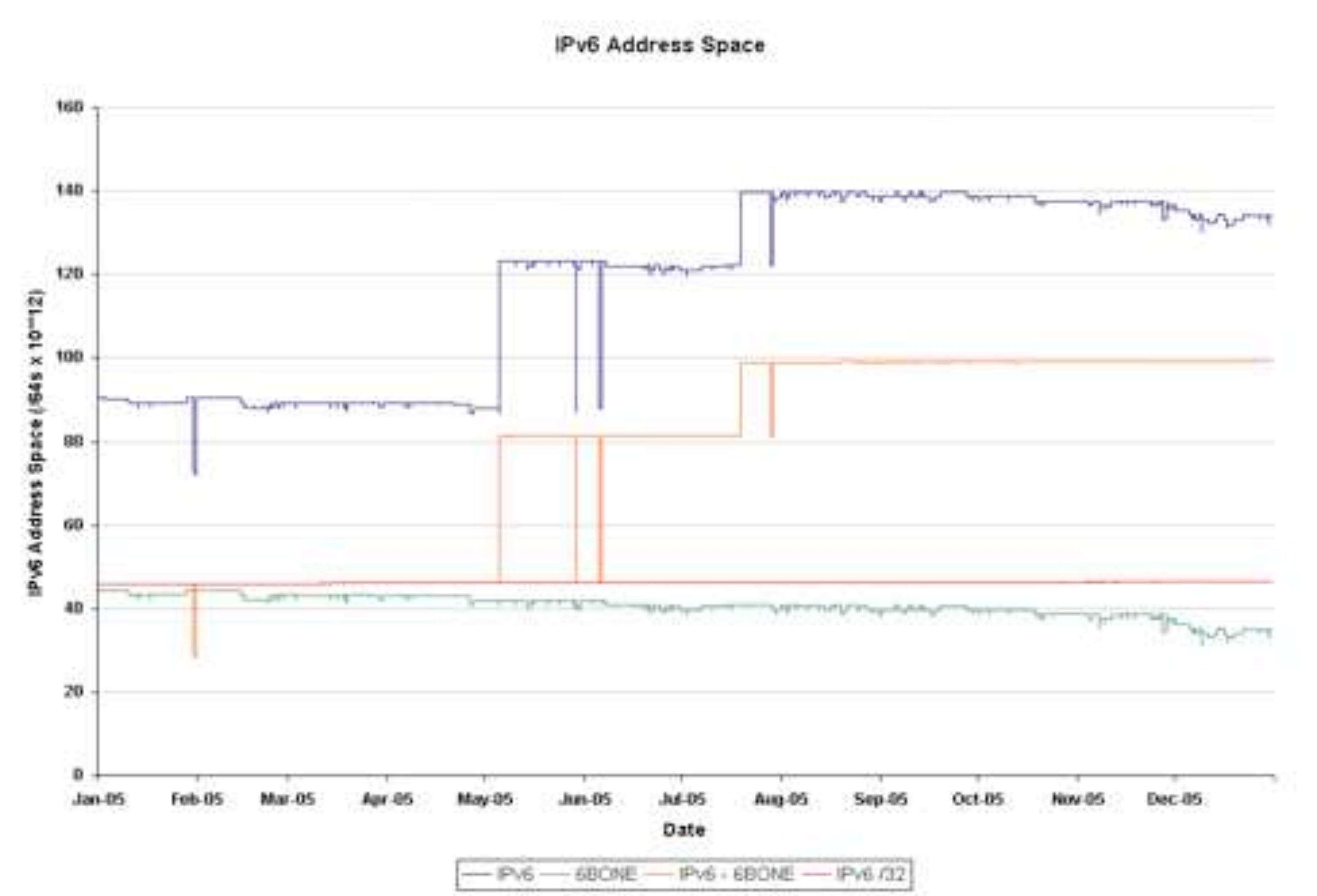


IPv6 in 2005

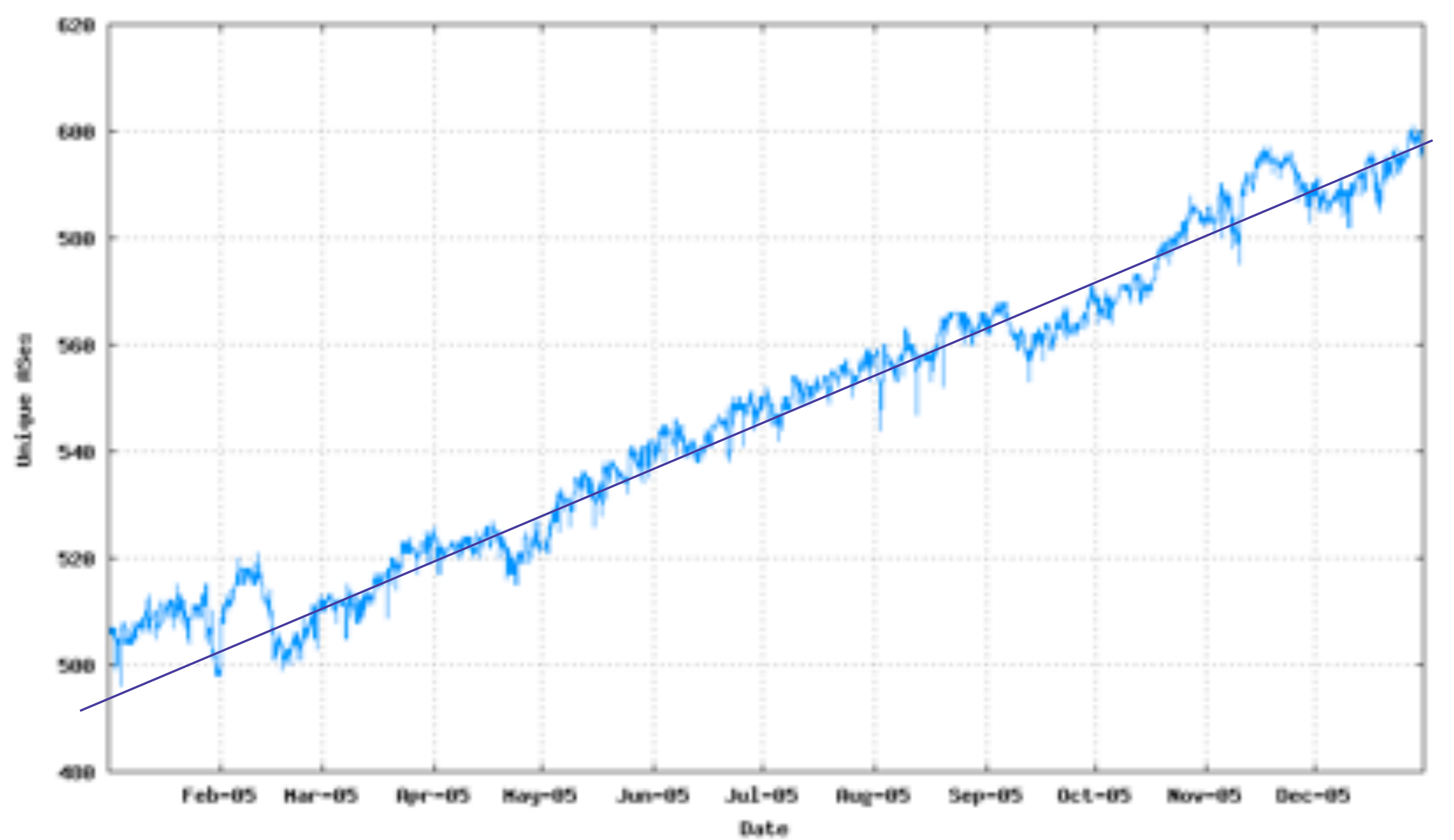
6Bone Address Span



IPv6 in 2005
Combined View of Address Span



IPv6 in 2005
Total Advertised AS Numbers



IPv6 – Vital Statistics for 2005

Prefixes	700 – 850	+21%
Roots	555 – 640	+15%
Specifics	145 - 210	+51%
Addresses	9 – 13.5 (10**13)	+50%
ASNs	500 – 600	+20%

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 2005 the IPv6 network remained small and continued to use a very large proportion of overlay tunnels at the edges. Larger scale trends in network characteristics were not readily discernable from 2005 figures

Vince Fuller's question:

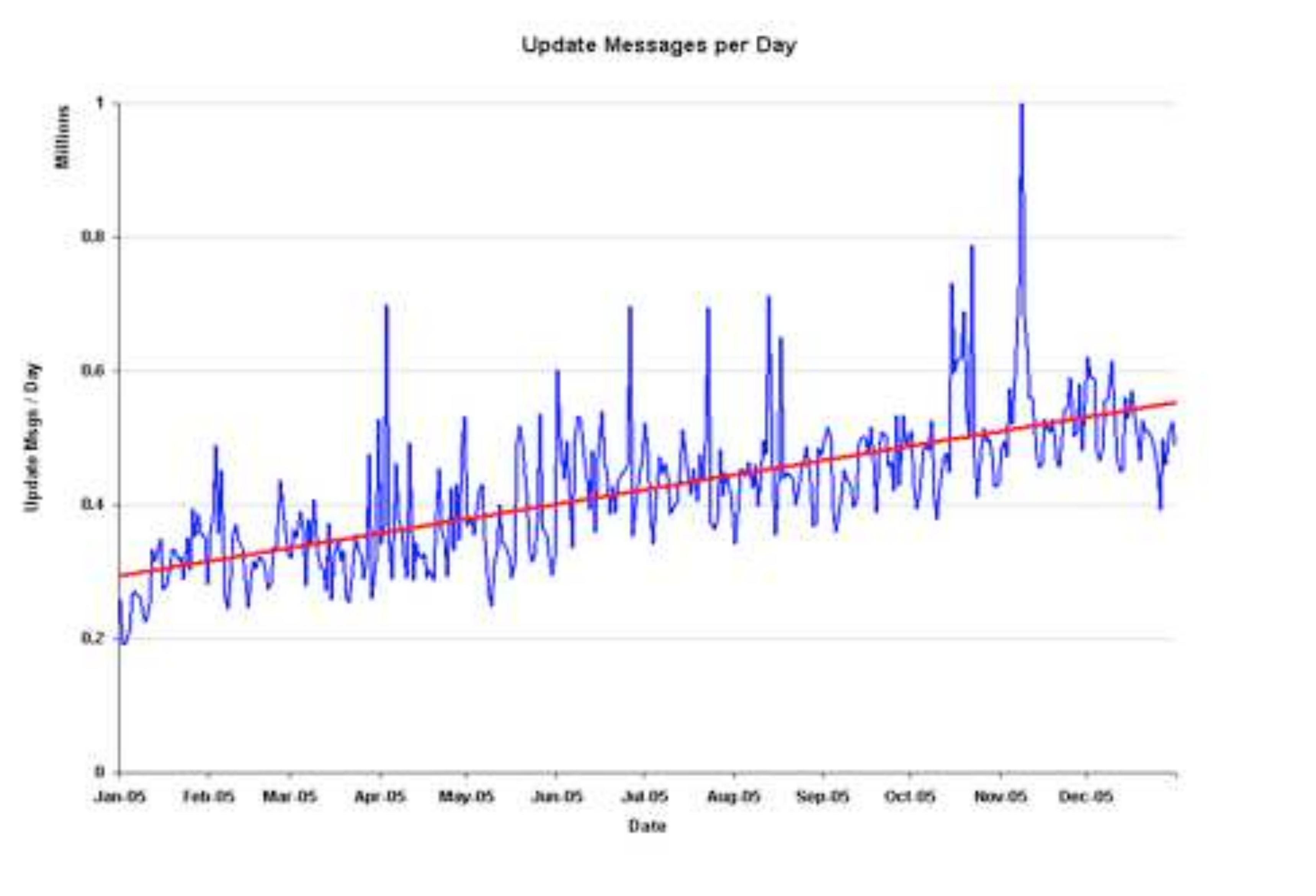
If you were buying a large router suitable for use in a "DFZ" with an expected lifetime of 3-5 years, what would you specify as the number of IPv4/IPv6 prefixes it must be able to handle? And how many prefix updates per second?

personal communication, January 2006

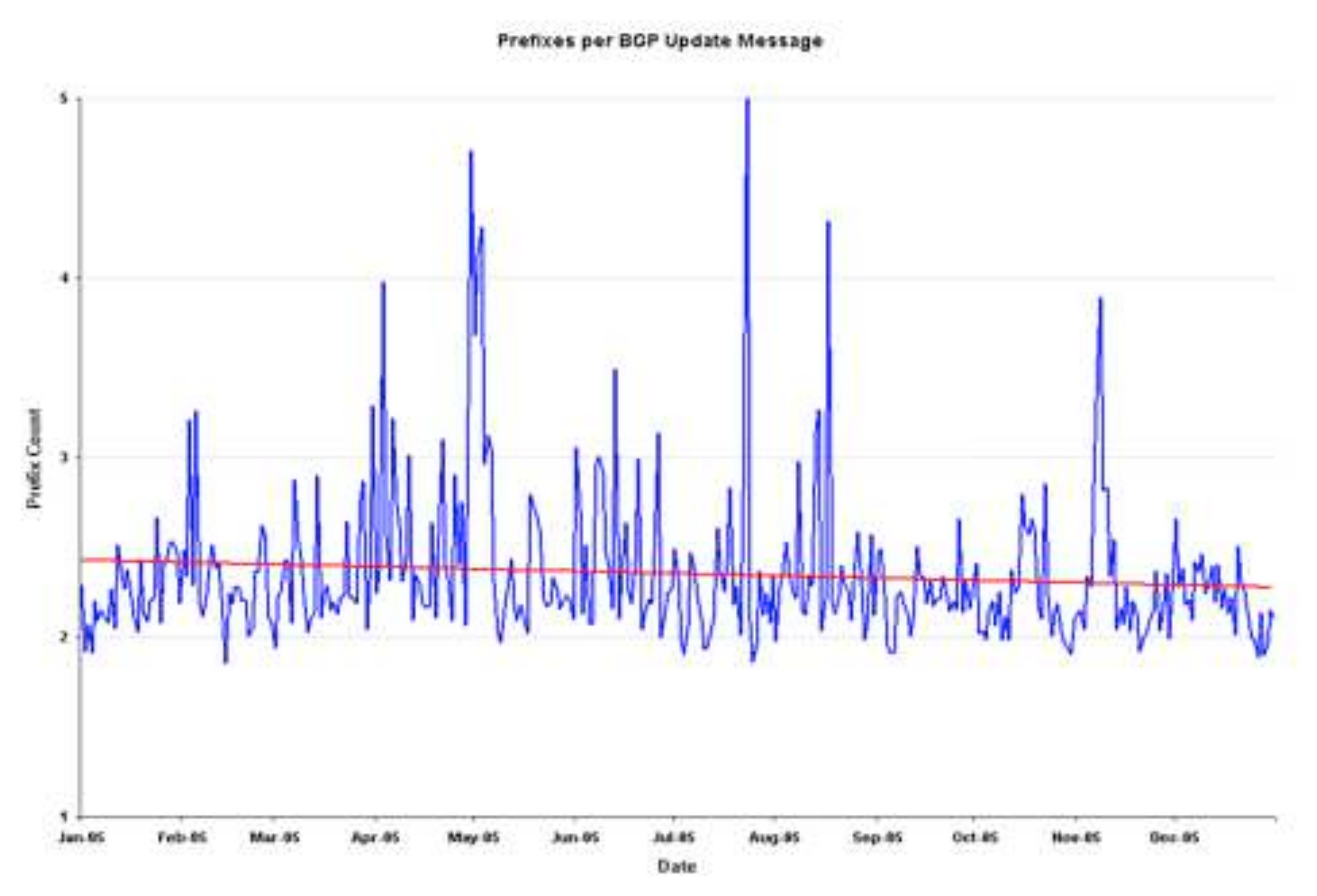
BGP Size Predictions - Methodology

- Examine update and withdrawal rates from BGP log records for 2005 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
- Examine CPU records from a core router in AS1221
  - Again look at the relative processing load against the table size
- Generate a BGP table size predictive model and use this to generate update rate and processing rate predictions

Update Message Rate



Prefixes per Update Message



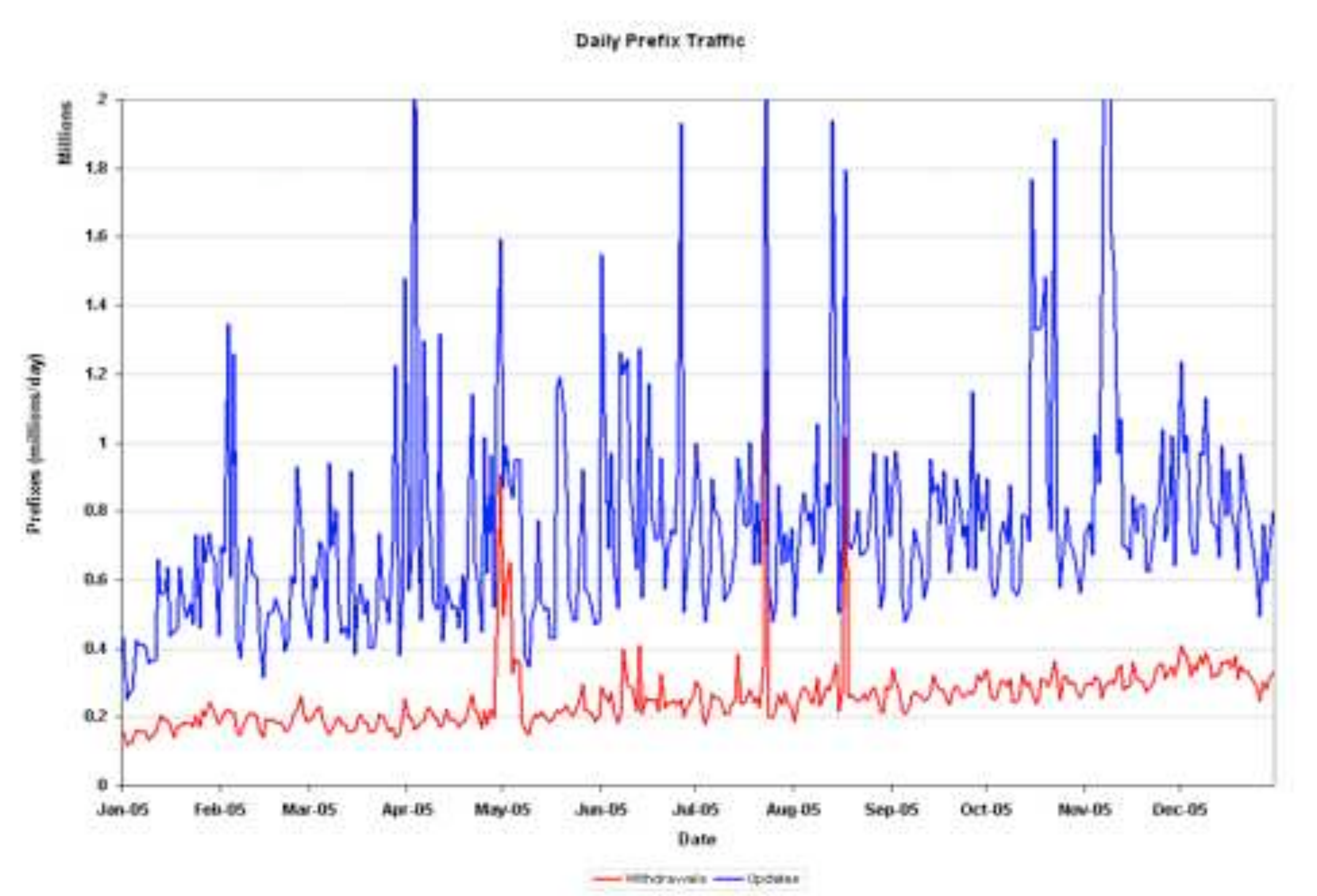
Update Trends across 2005

- Number of update messages per day has doubled across 2005 (Dec 2005 saw approx 550,000 update messages per day)
  - Considering the population size the daily data rate is highly variable – why?
- Number of prefixes per update message is falling from an average of 2.4 to 2.3 prefixes per update
  - Is this attributable to ncreased use of public ASs and eBGP at the edge of the network? (Multi-homing?)
- Is the prefix update rate increasing at a greater rate than the number of prefixes in the routing table?
  - Is there some multiplicative factor at play here?
  - Why is instability increasing faster than the network size?

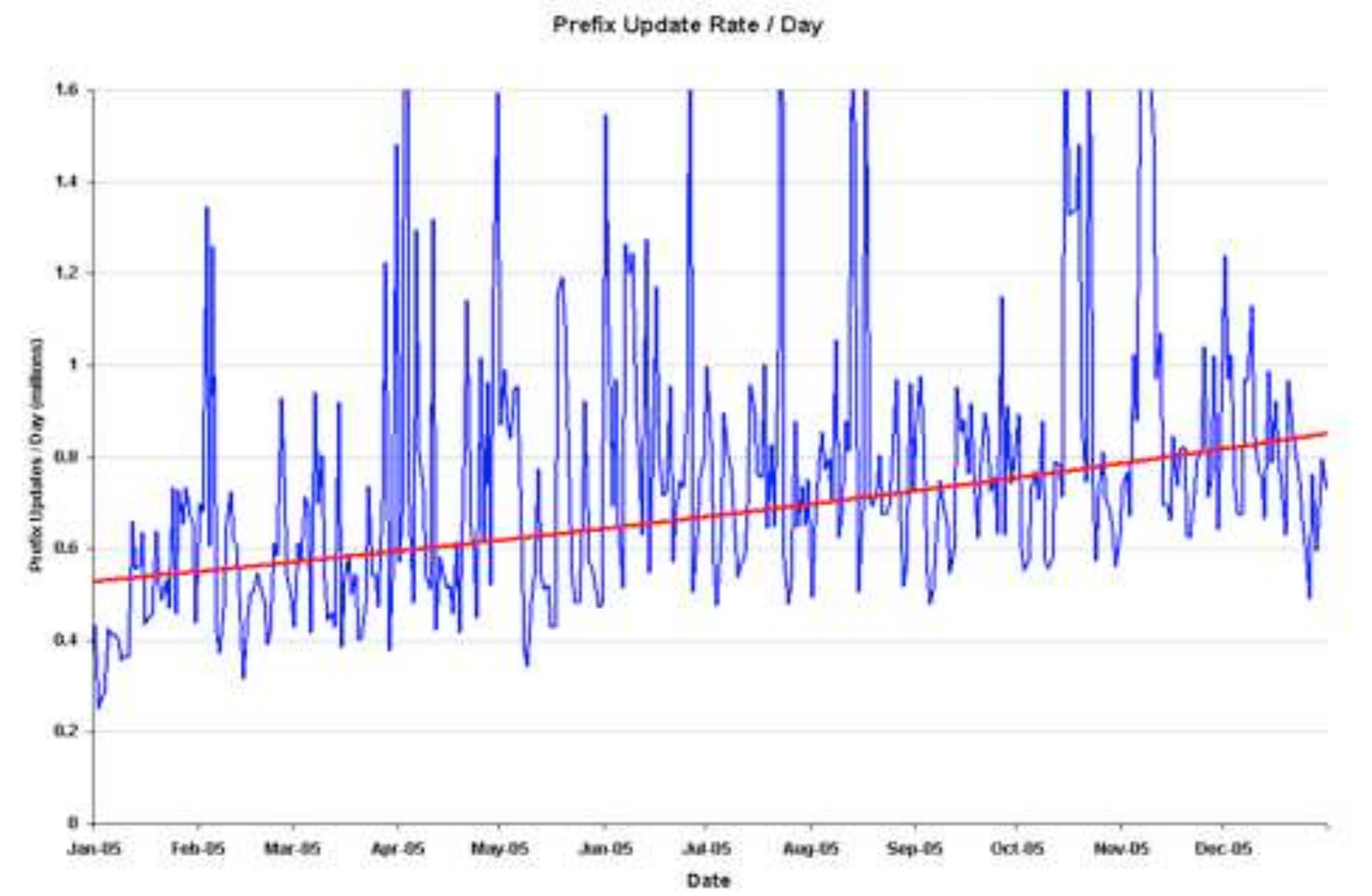
Prefixes vs Updates

- Look at the number of prefixes that are the subject of update messages
- What are the trends of prefix update behaviour?

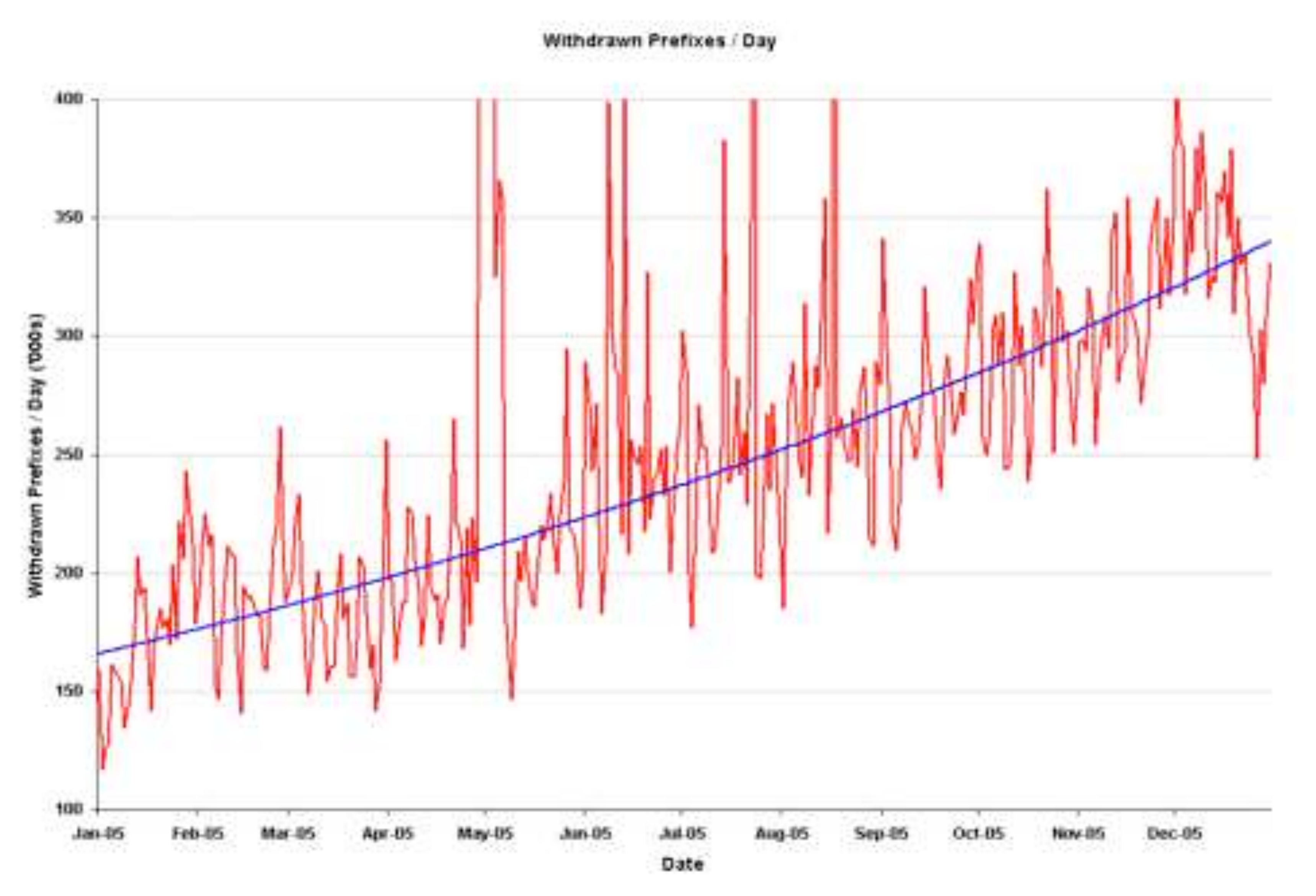
Prefix Update and Withdrawal Rates



Prefix Update Rates



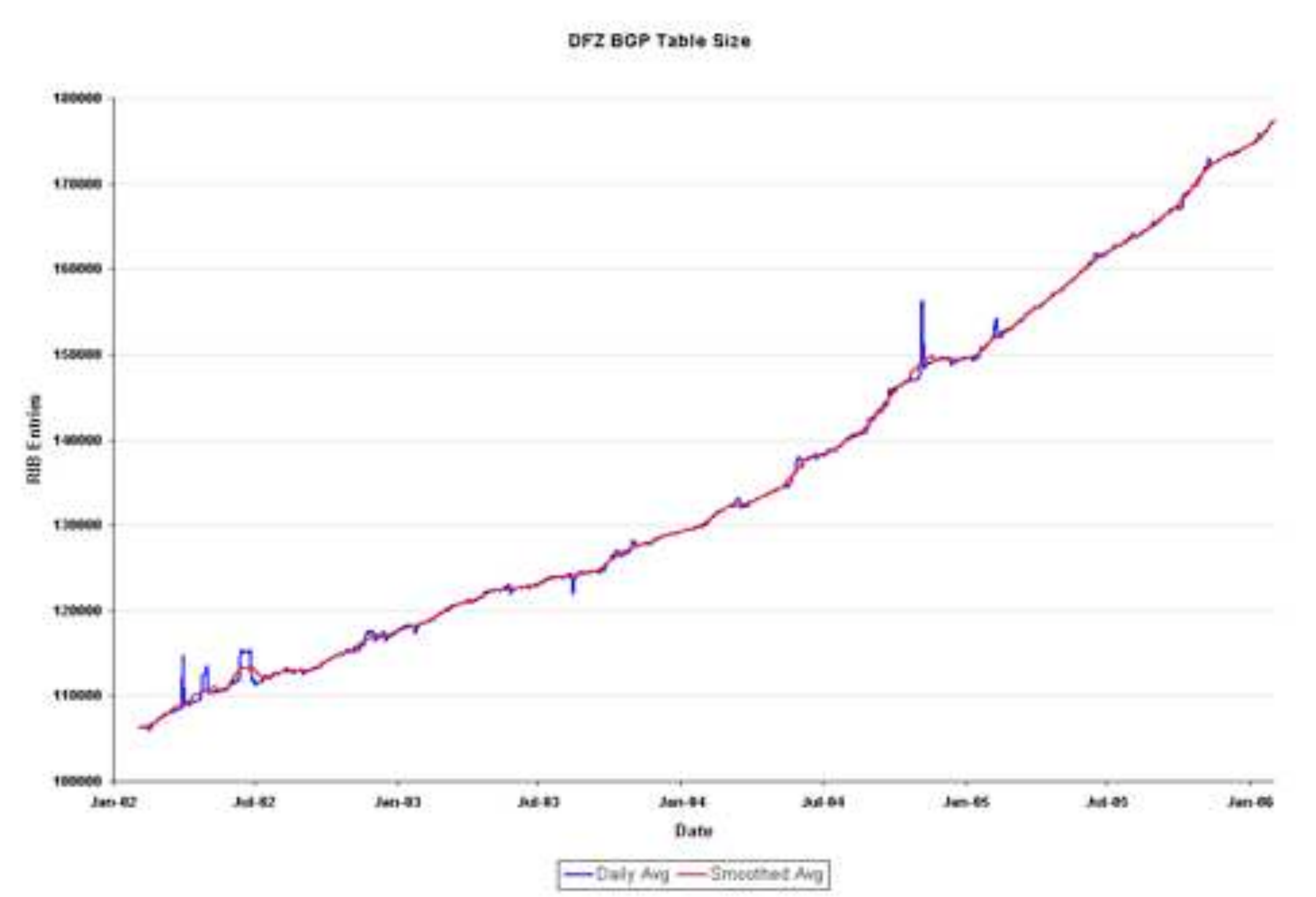
Withdrawal Rates



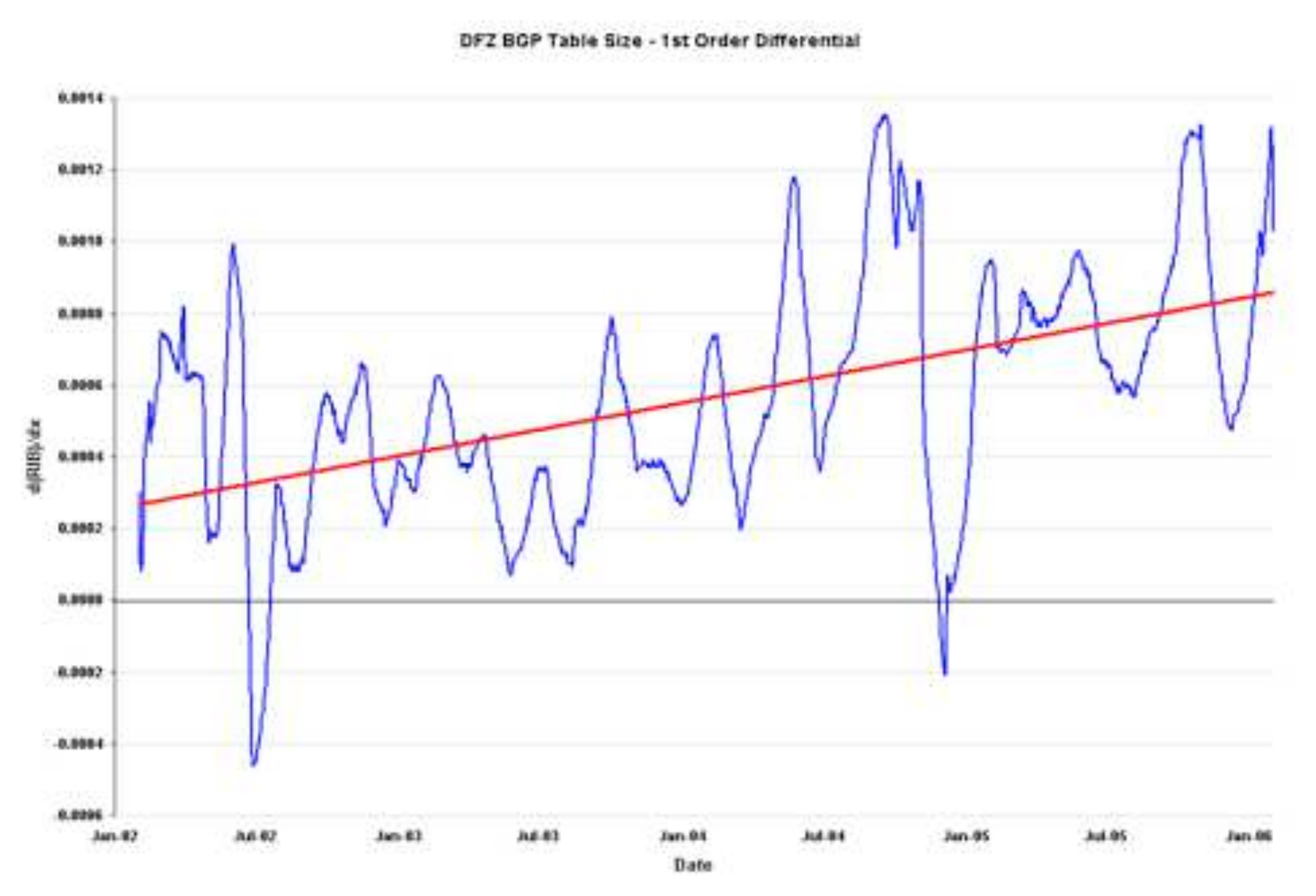
Prefix Rate Trends

- High variability in day-to-day prefix change rates
- Best fit model appears to be exponential – although update and withdrawal rates show different growth rates

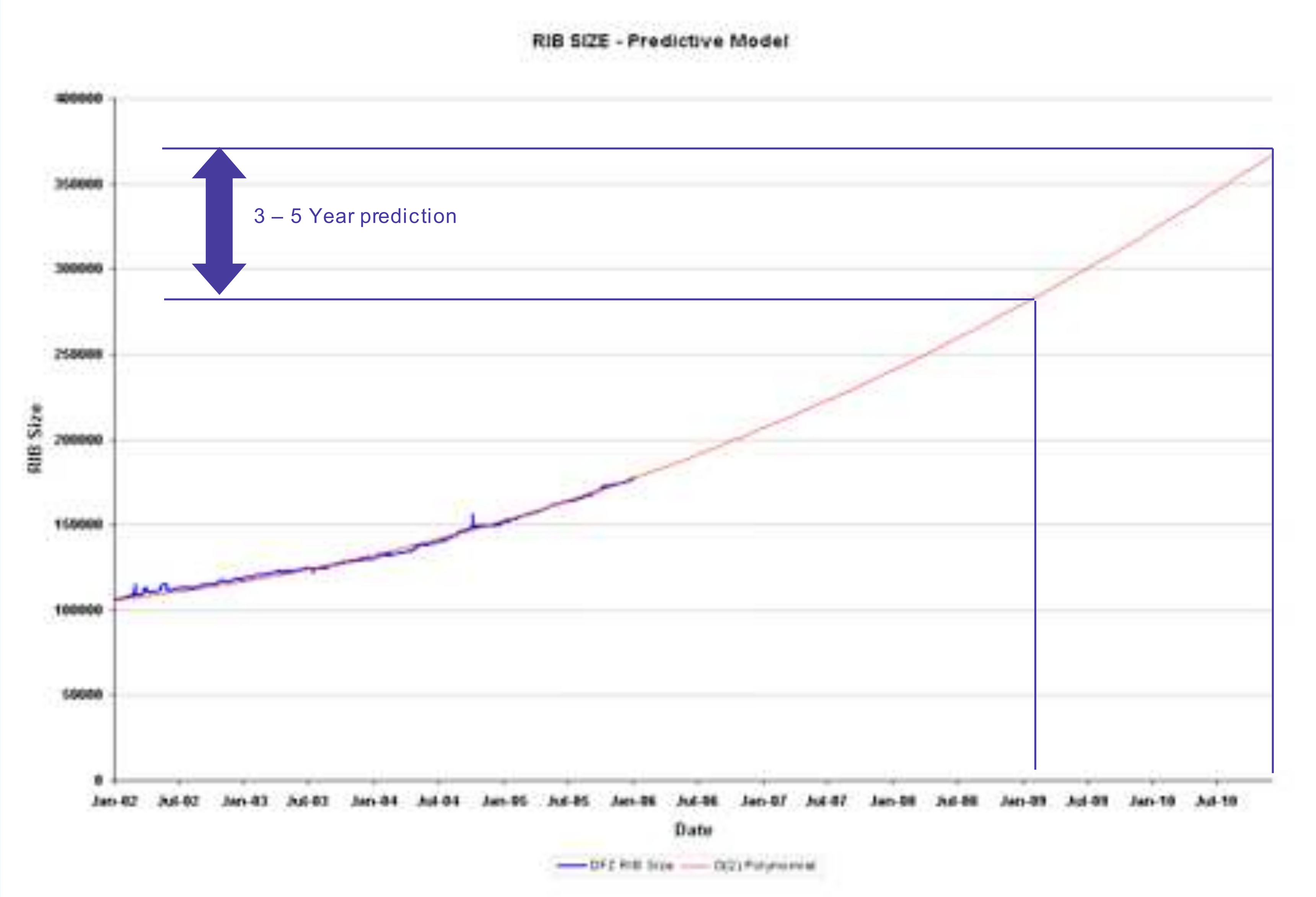
DFZ Prefix Table Size



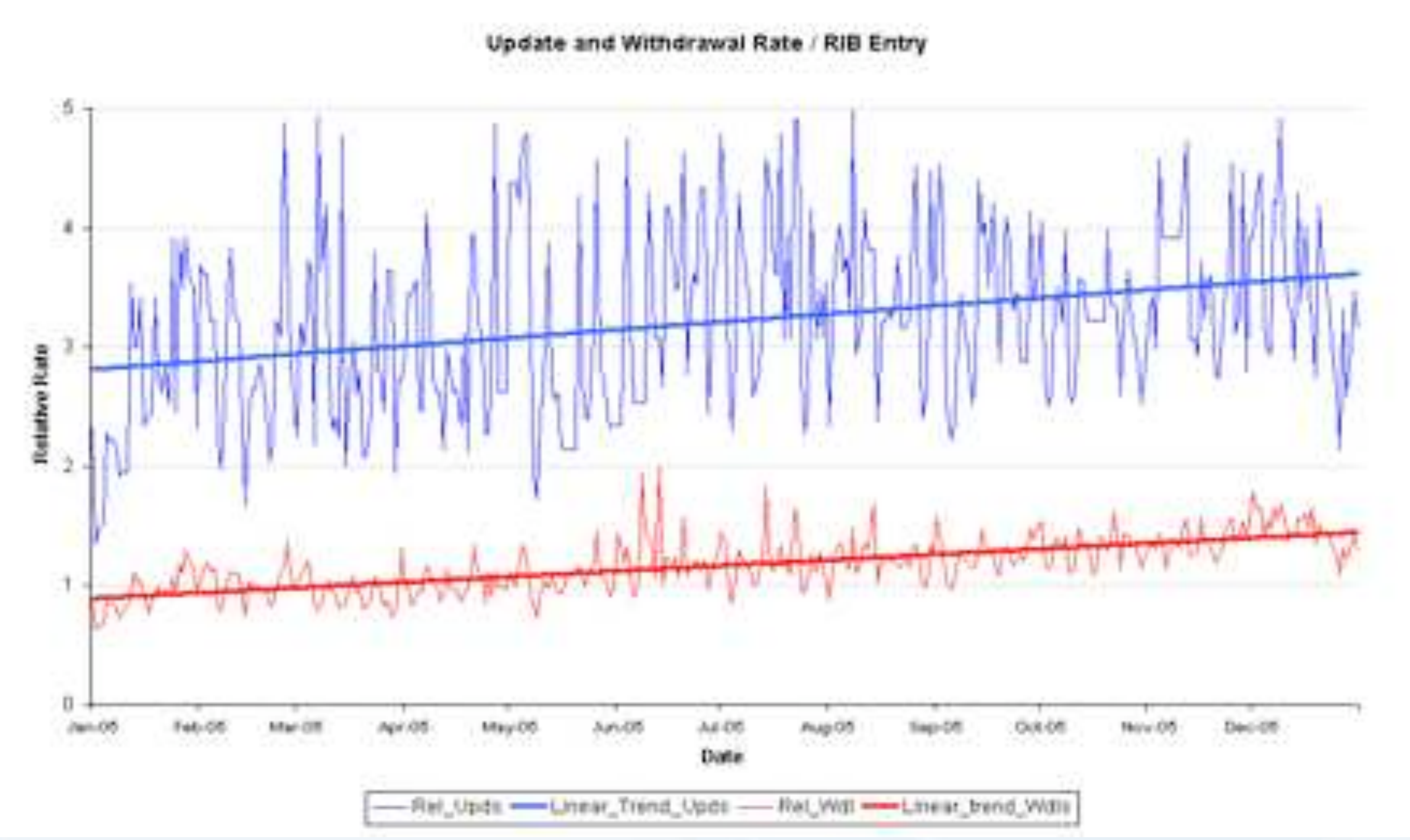
1st Order Differential



DFZ Model as an O(2) Polynomial



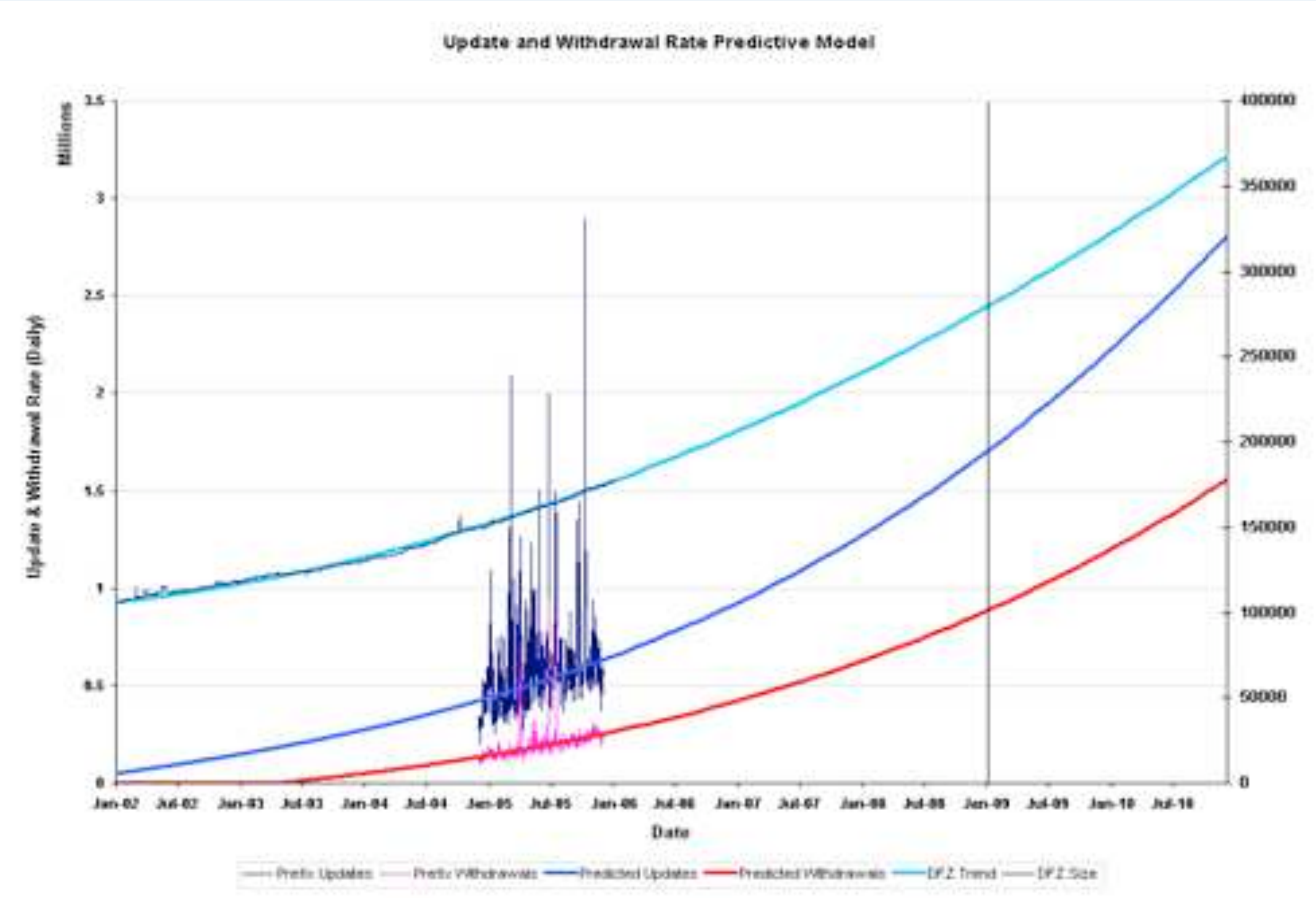
Relative Update / Withdrawal Rates



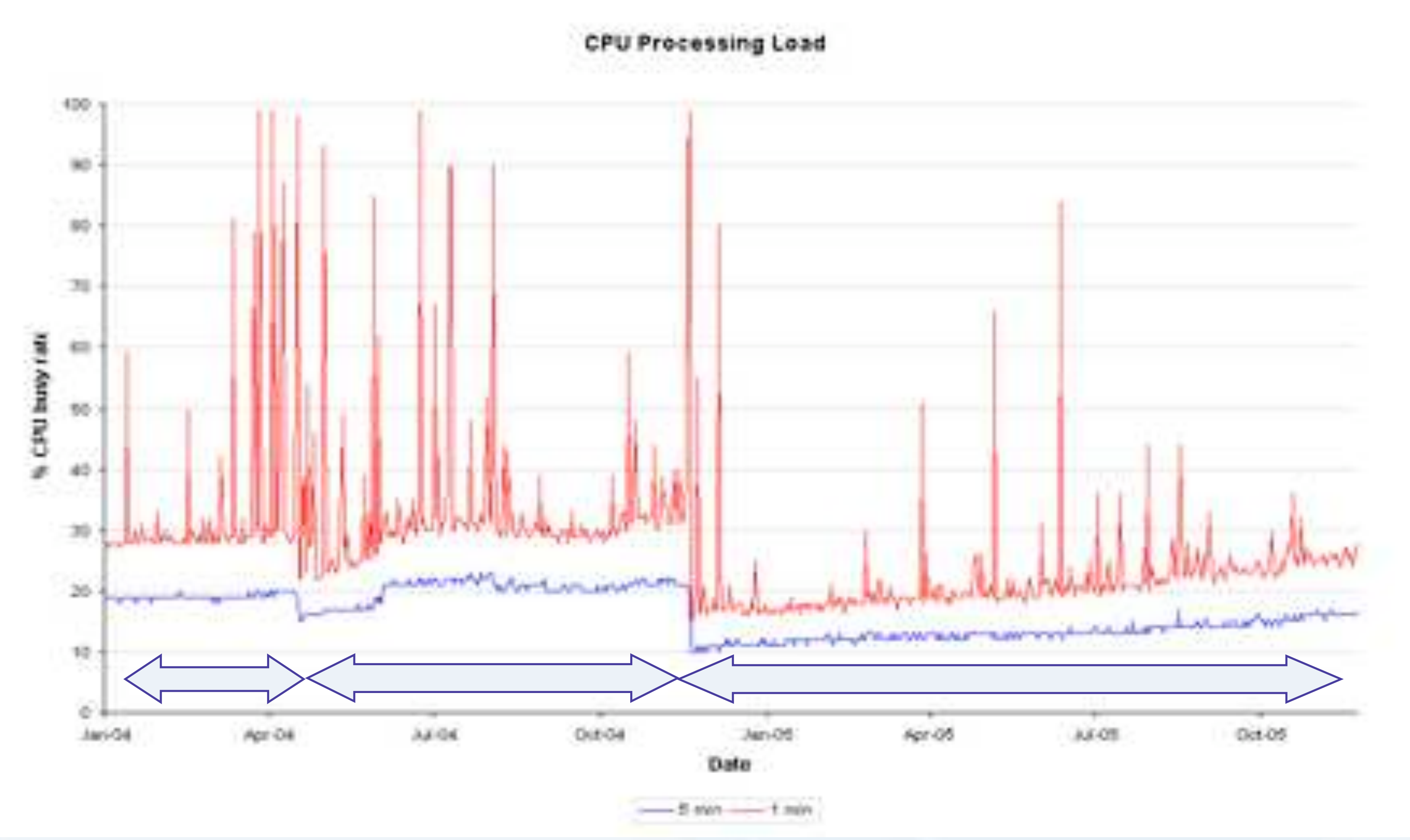
Asia Pacific Network Information Centre

APNIC

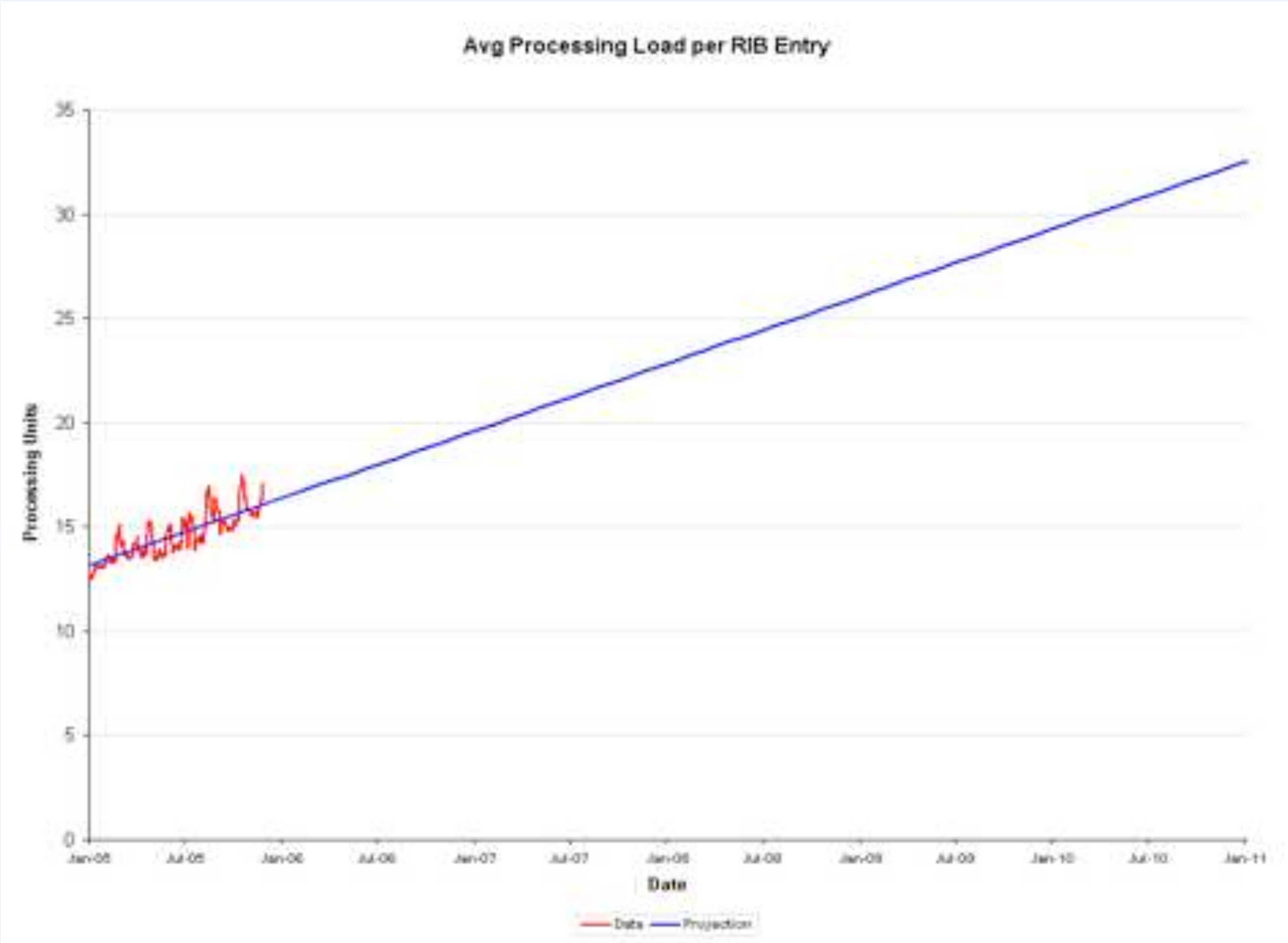
Update Rate Prediction



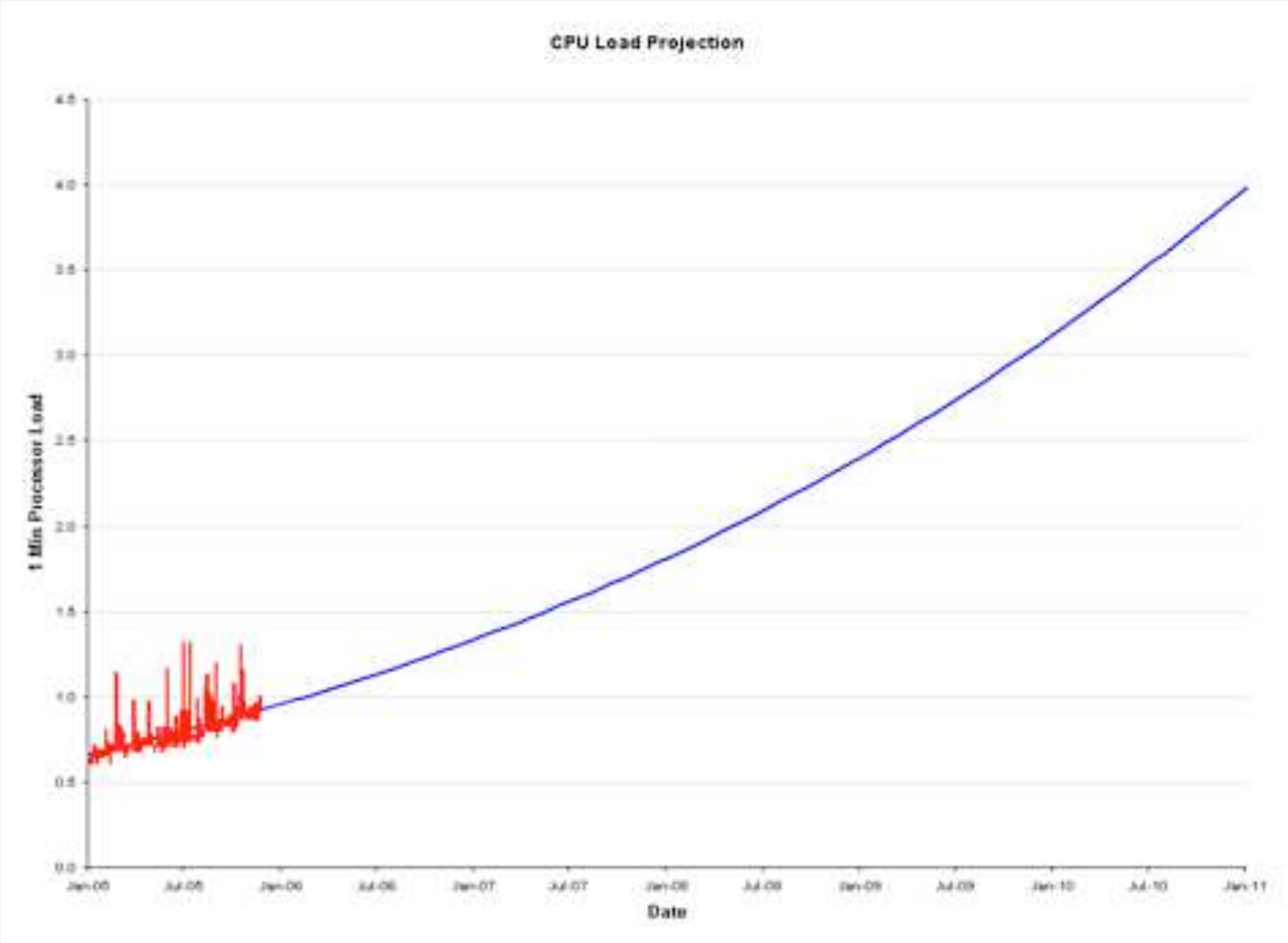
Processing Metrics



Relative Processing Metrics



Projected Processing Load



3-5 Year Predictions for the IPv4 DFZ

- Today (1/1/2006)
  - Table Size **176,000** prefixes
  - Update Rate 0.7M prefix updates / day
  - Withdrawal Rate 0.4M prefix withdrawals per day
  - 250Mbytes memory
  - 30% of a 1.5Ghz processor
- 3 Years (1/1/2009)
  - Table Size **275,000** prefixes
  - Update Rate 1.7M prefix updates / day
  - Withdrawal Rate 0.9M withdrawals per day
  - 400Mbytes Memory
  - 75% of a 1.5Ghz processor
- 5 Years (1/1/2011)
  - Table Size **370,000** prefixes
  - Update Rate 2.8M prefix updates / day
  - Withdrawal Rate 1.6M withdrawals per day
  - 550Mbytes Memory
  - 120% of a 1.5Ghz processor

However…

- These are **very** low end predictors
  - The router needs to cope with per second peak update rates, not average loads
  - It's the capability to keep the forwarding fabric in sync with the network topology that is the critical factor – its speed under peak load that counts
  - These projections assume unaltered BGP
    - For example, secure BGP protocol sessions, additional security-related payload factors, incremental workload to validate security payloads, and related aspects are not factored in
- It would be prudent to include a significant additional capability margin for these factors.

DFZ router sizing for 3 – 5 years
A more conservative estimate:

- 500,000 entries in the RIB
- Update rate of up to 6M prefix updates /day
- Short term peak update rate 100 x average daily rate (7000 prefix updates /sec)
- 2 Gbytes route processor memory (or more, depending on DFZ peer count)
- 5GHz processor for route processing

+ Security processing overheads

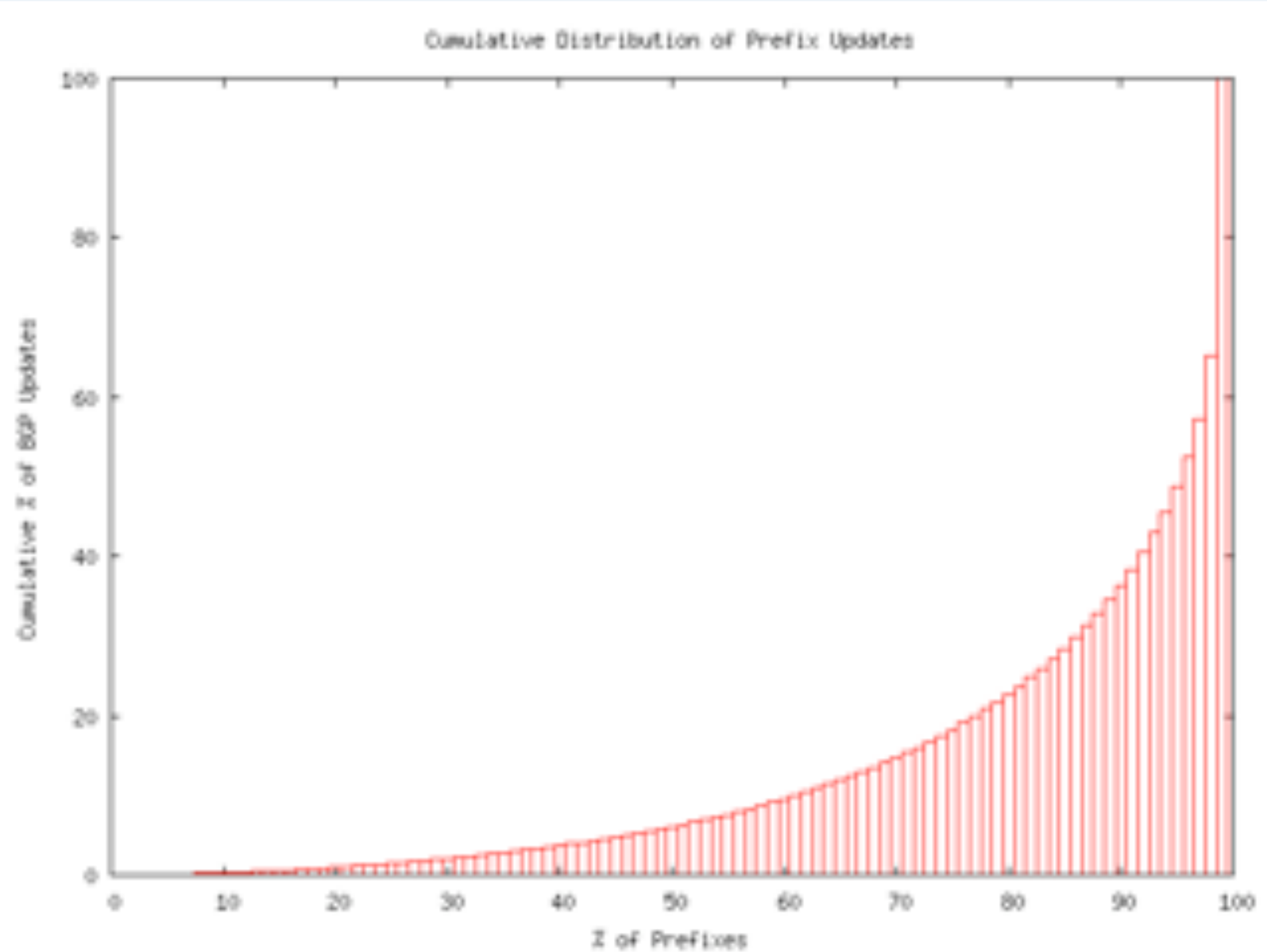
What’s the uncertainty factor?

- Are we seeing a uniform distribution of updates across all ASs and all Prefixes?
- Or is this a skewed heavy tail distribution where a small number of prefixes contribute to most of the BGP updates?

Prefix Stats

- Number of unique prefixes announced: 289,558
- Prefix Updates: 70,761,786
- Stable prefixes: 12,640
- Updated prefixes (year end): 162,039
- Withdrawn prefixes: 127,519

Distribution of Updates by Prefix

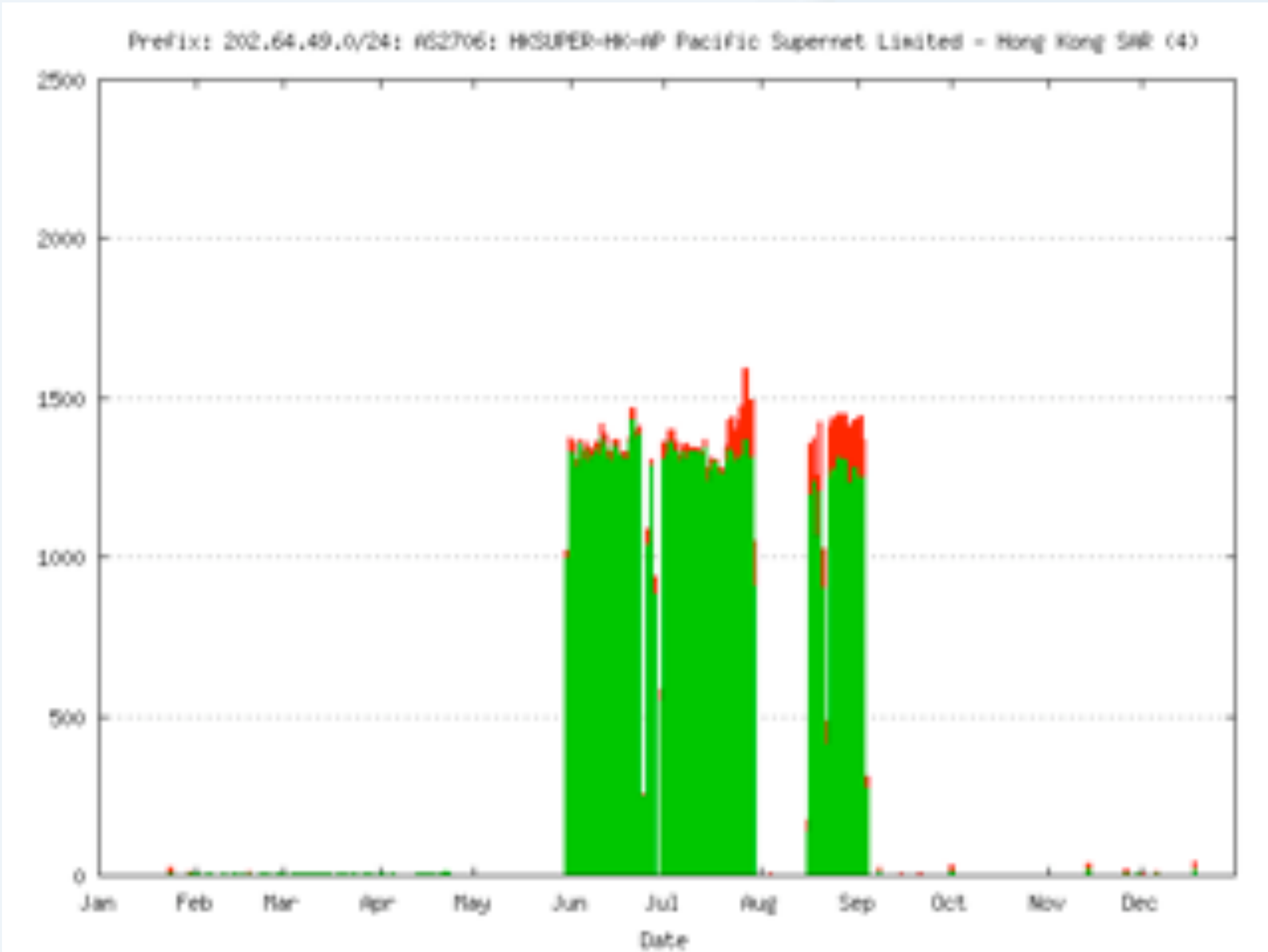


Active Prefixes

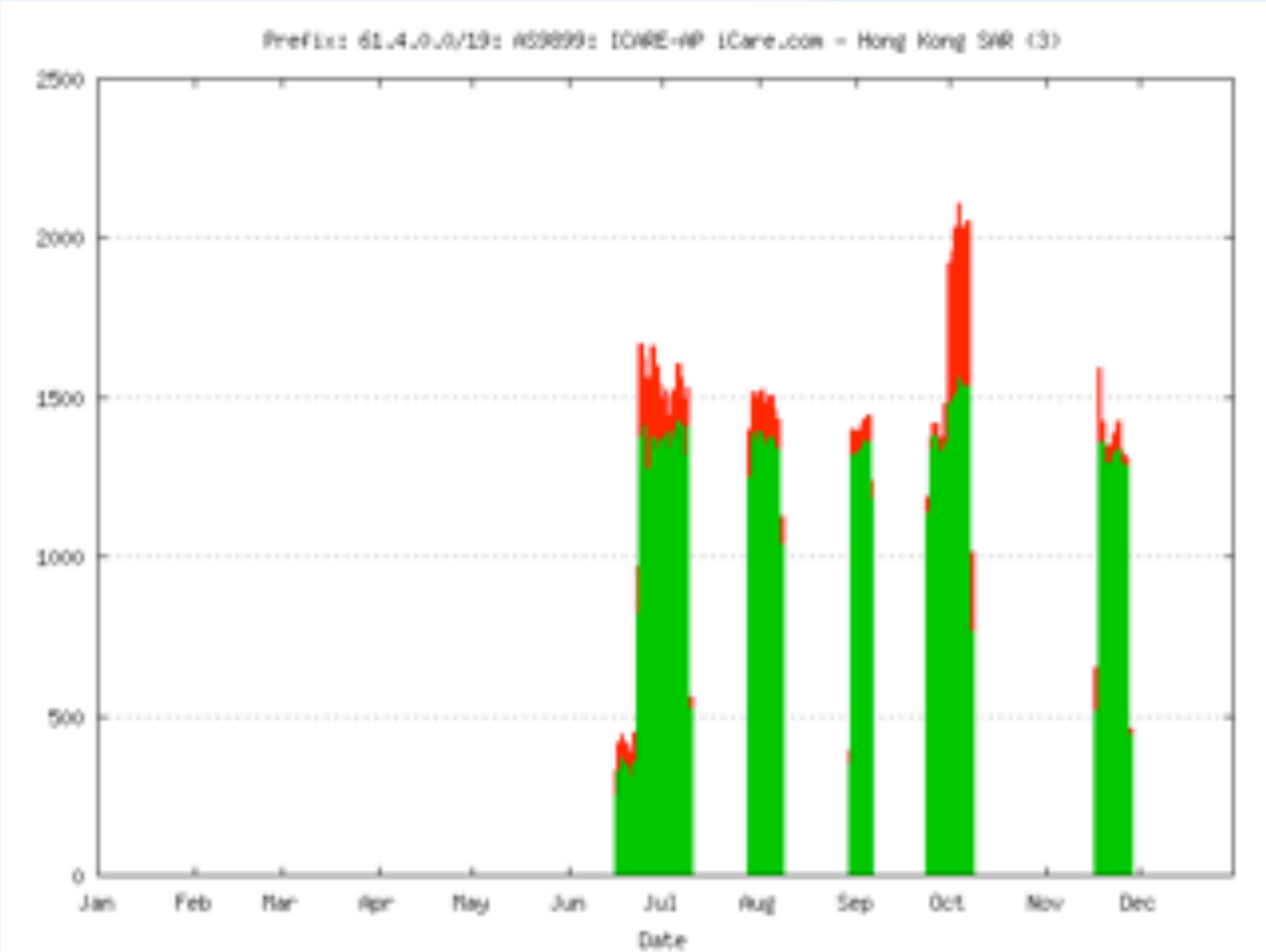
Top 10 Prefixes

	Prefix	Updates	Flaps	Re-Homes
1.	202.64.49.0/24	198,370	96,330	918
2.	61.4.0.0/19	177,132	83,277	55
3.	202.64.40.0/24	160,127	78,494	1,321
4.	81.212.149.0/24	158,205	61,455	20,031
5.	81.213.47.0/24	138,526	60,885	12,059
6.	209.140.24.0/24	132,676	42,200	0
7.	207.27.155.0/24	103,709	42,292	0
8.	81.212.197.0/24	99,077	37,441	15,248
9.	66.150.140.0/23	84,956	11,109	5,963
10.	207.168.184.0/24	74,679	34,519	0

1 - 202.64.49.0/24

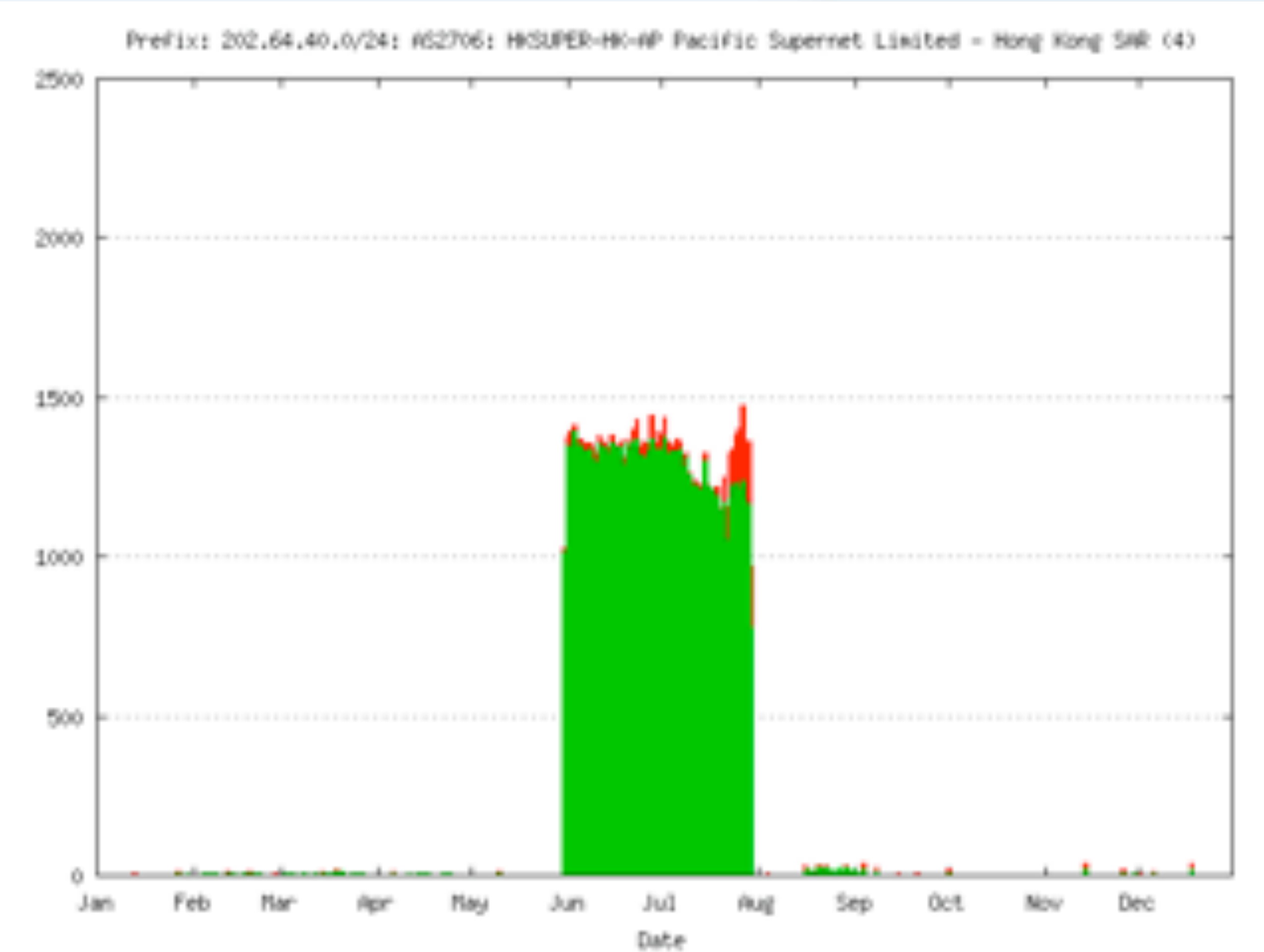


2 - 61.4.0.0/19

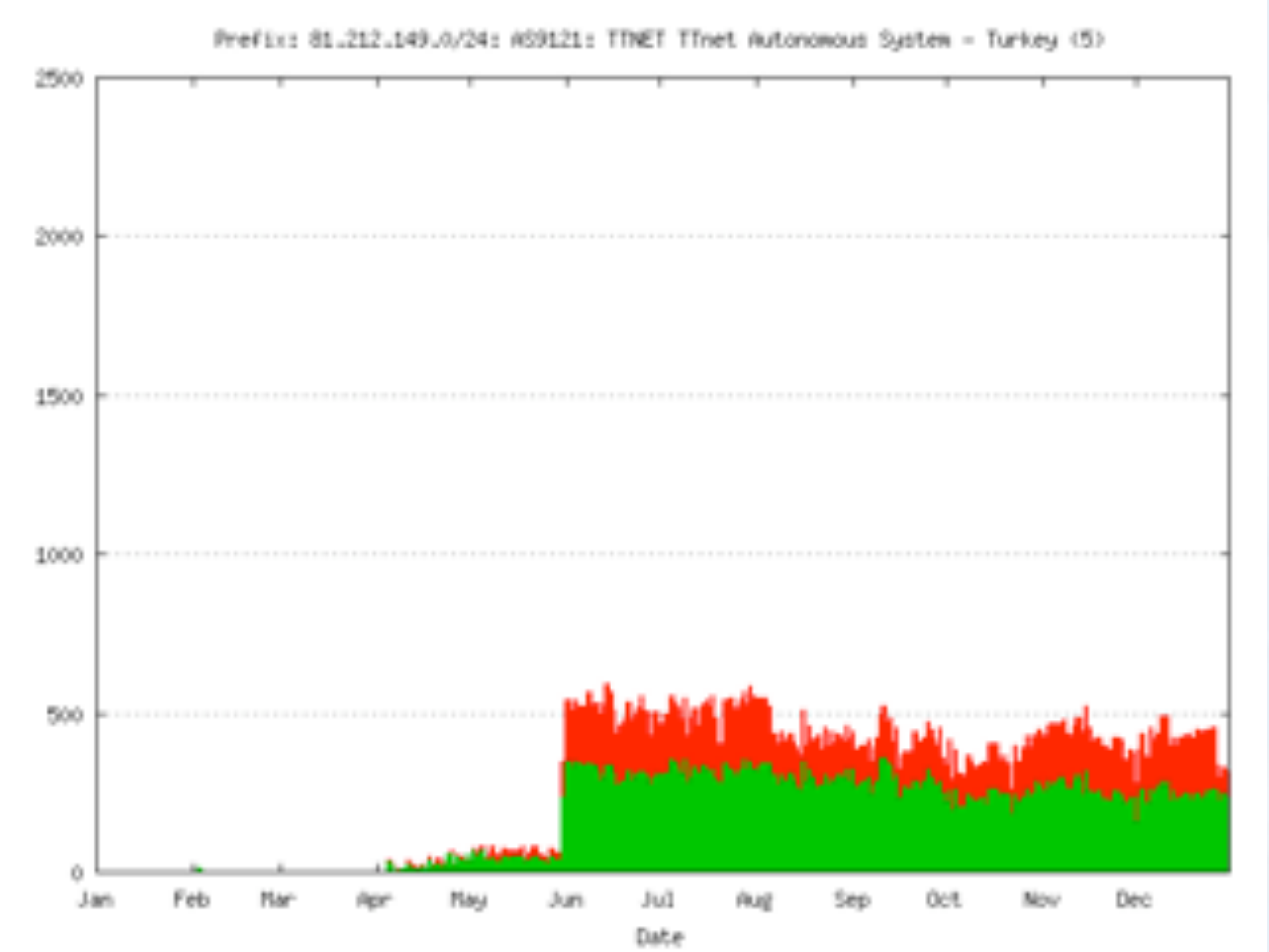


Prefix: 61.4.0.0/19: AS9899: ICARE-AP iCare.com - Hong Kong SAR (3)
2500
2000
1500
1000
500
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Date

3 - 202.64.40.0/24



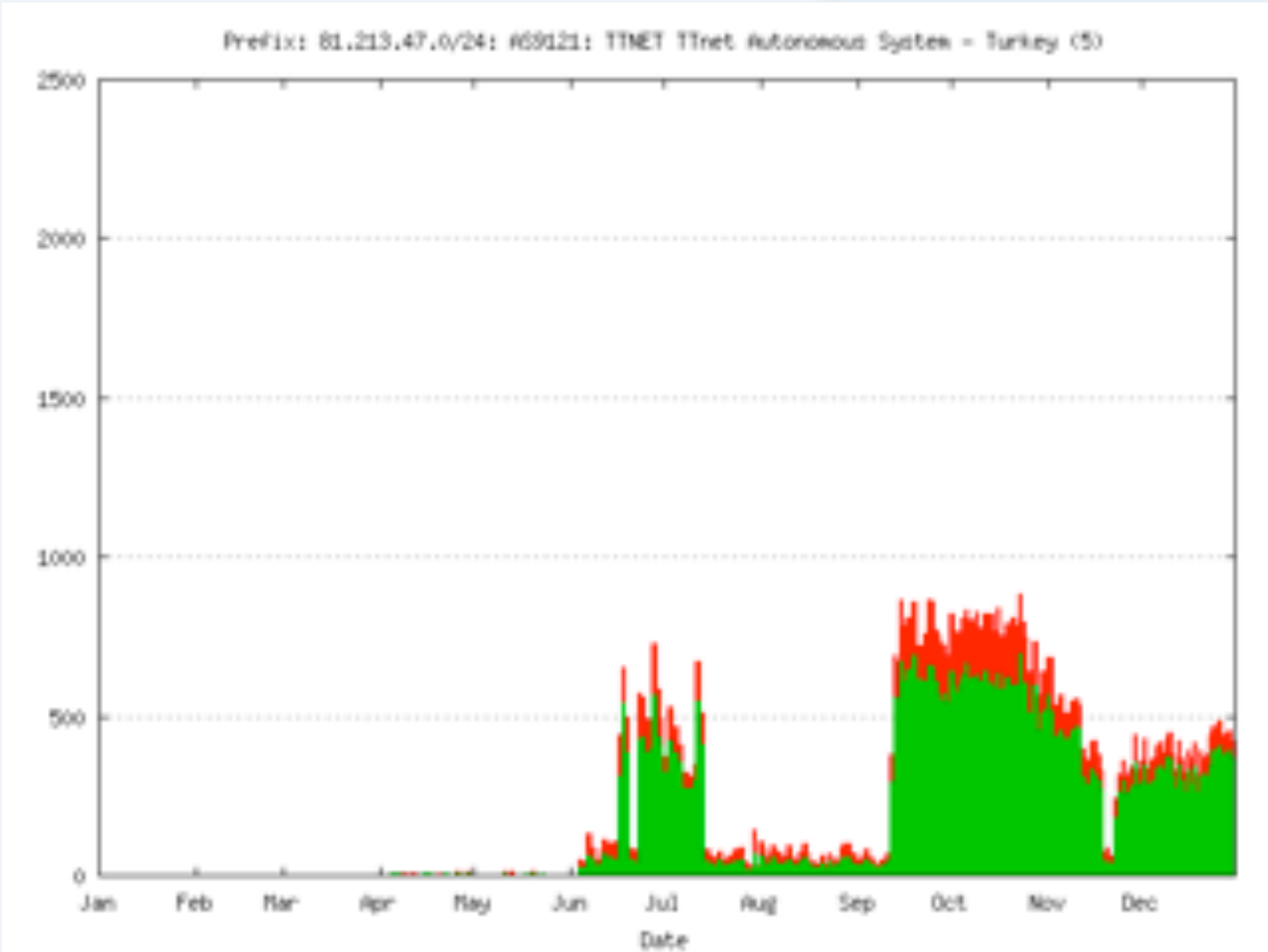
4 - 81.212.149.0/24



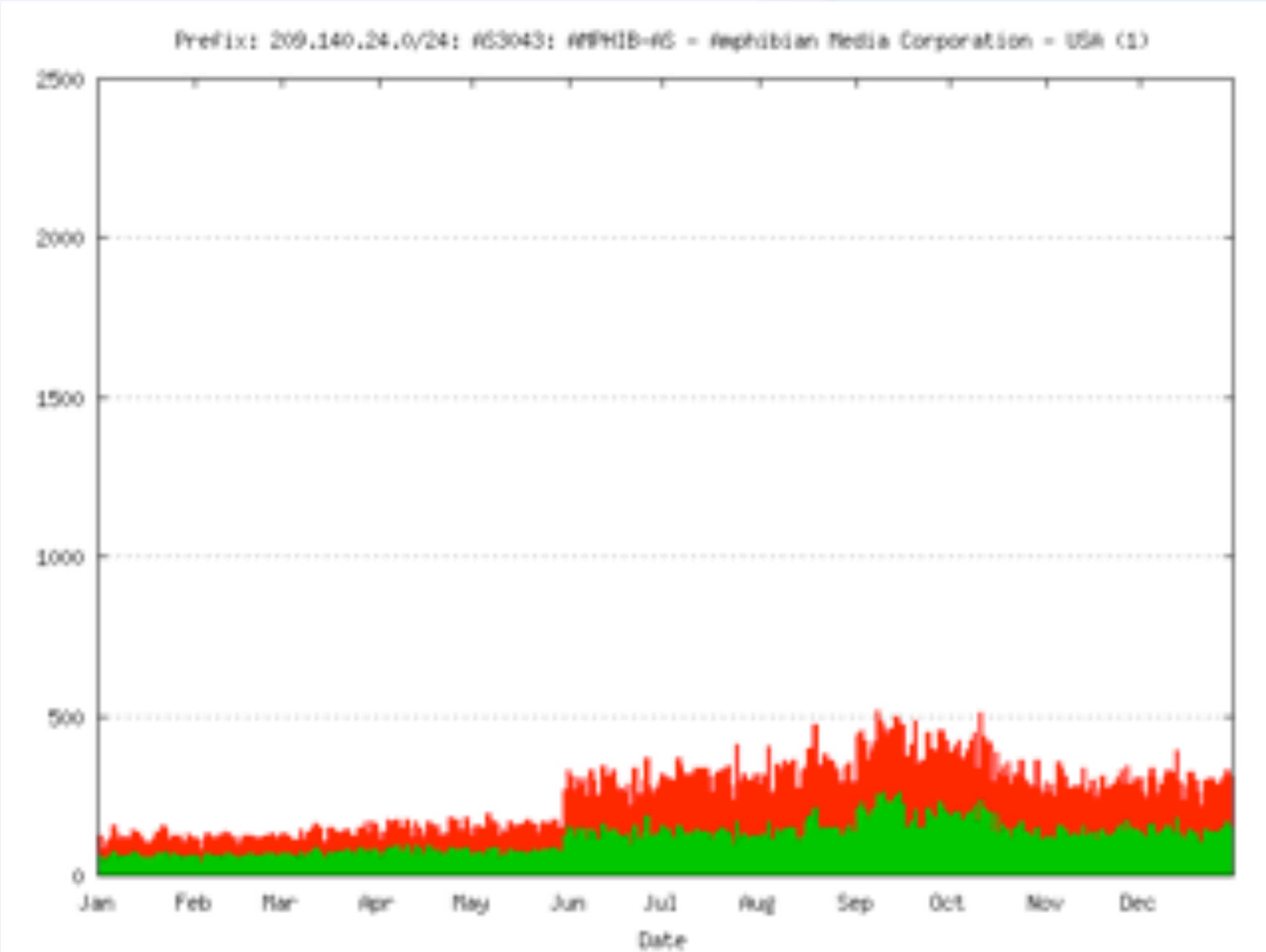
Asia Pacific Network Information Centre

APNIC

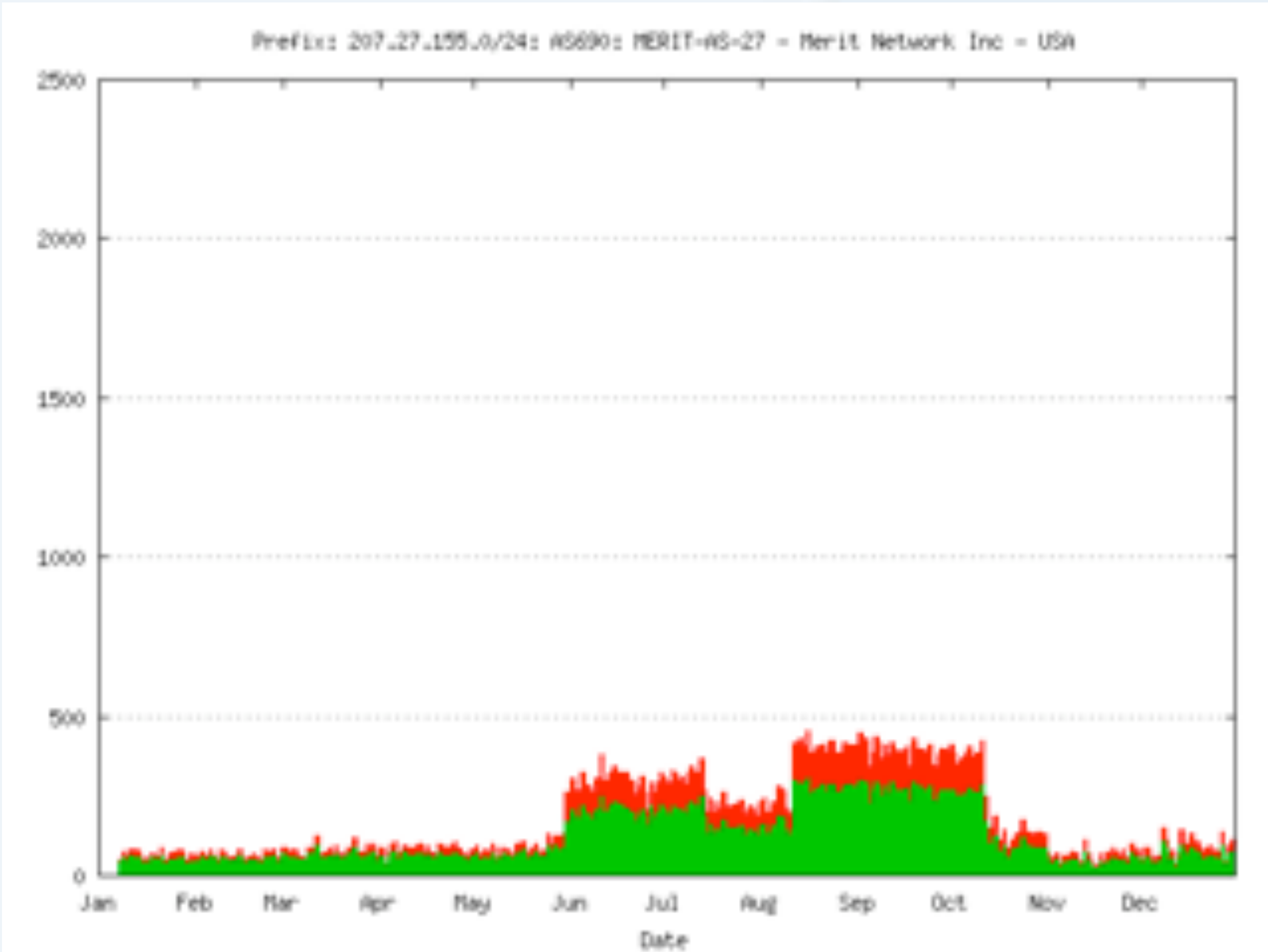
5 - 81.213.47.0/24



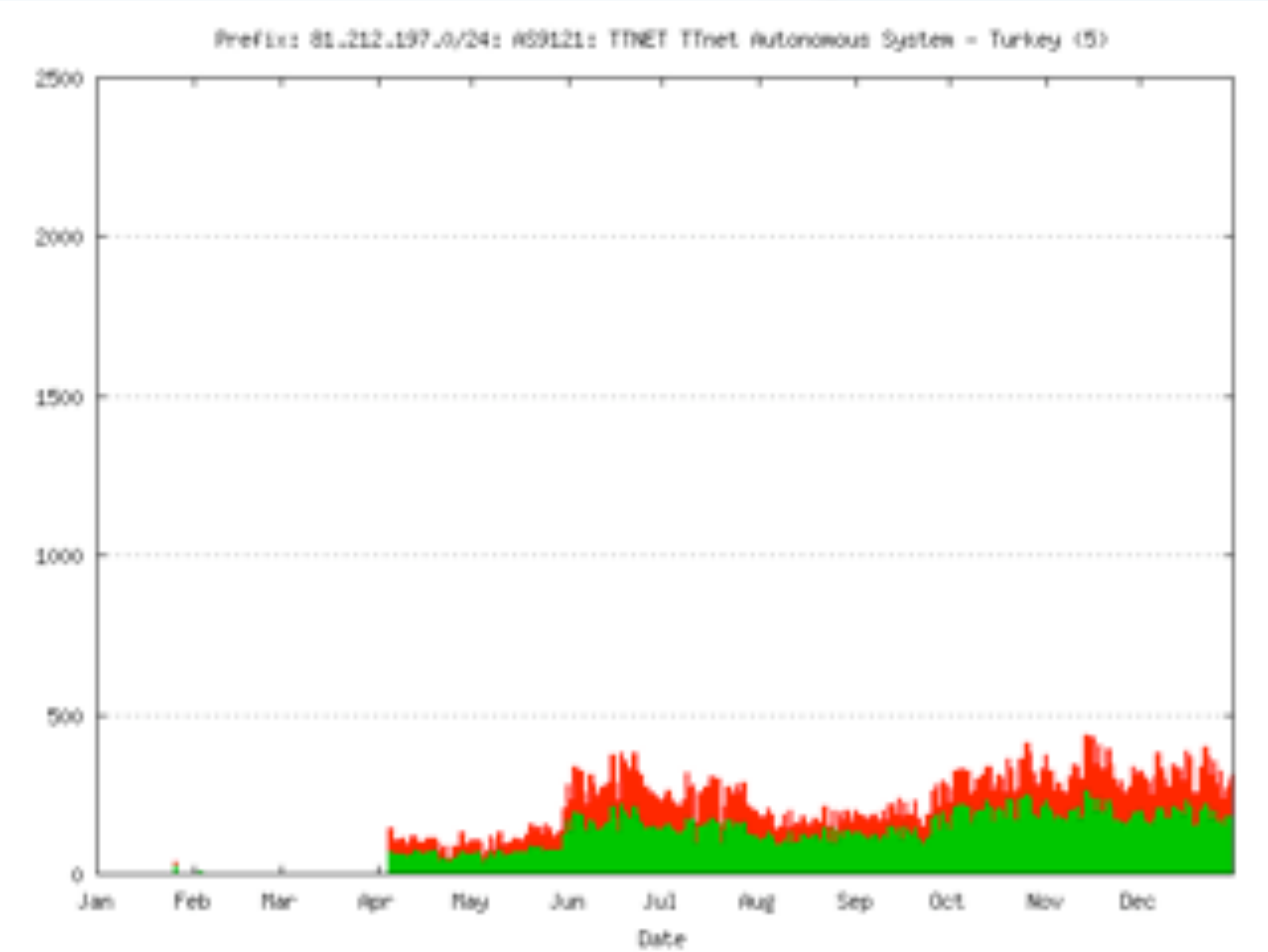
6 - 209.140.24.0/24



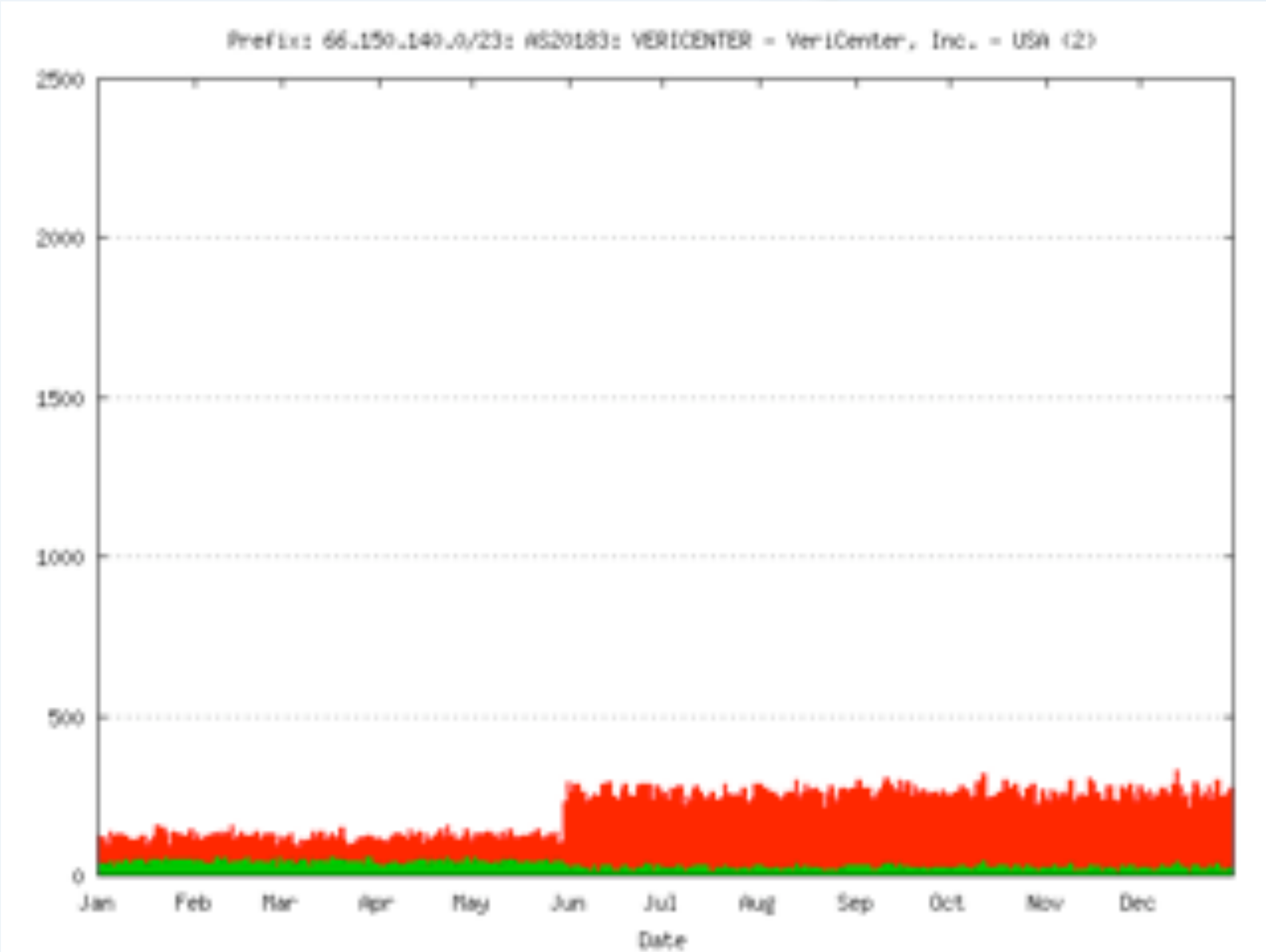
7 - 207.27.155.0/24



8 - 81.212.197.0/24

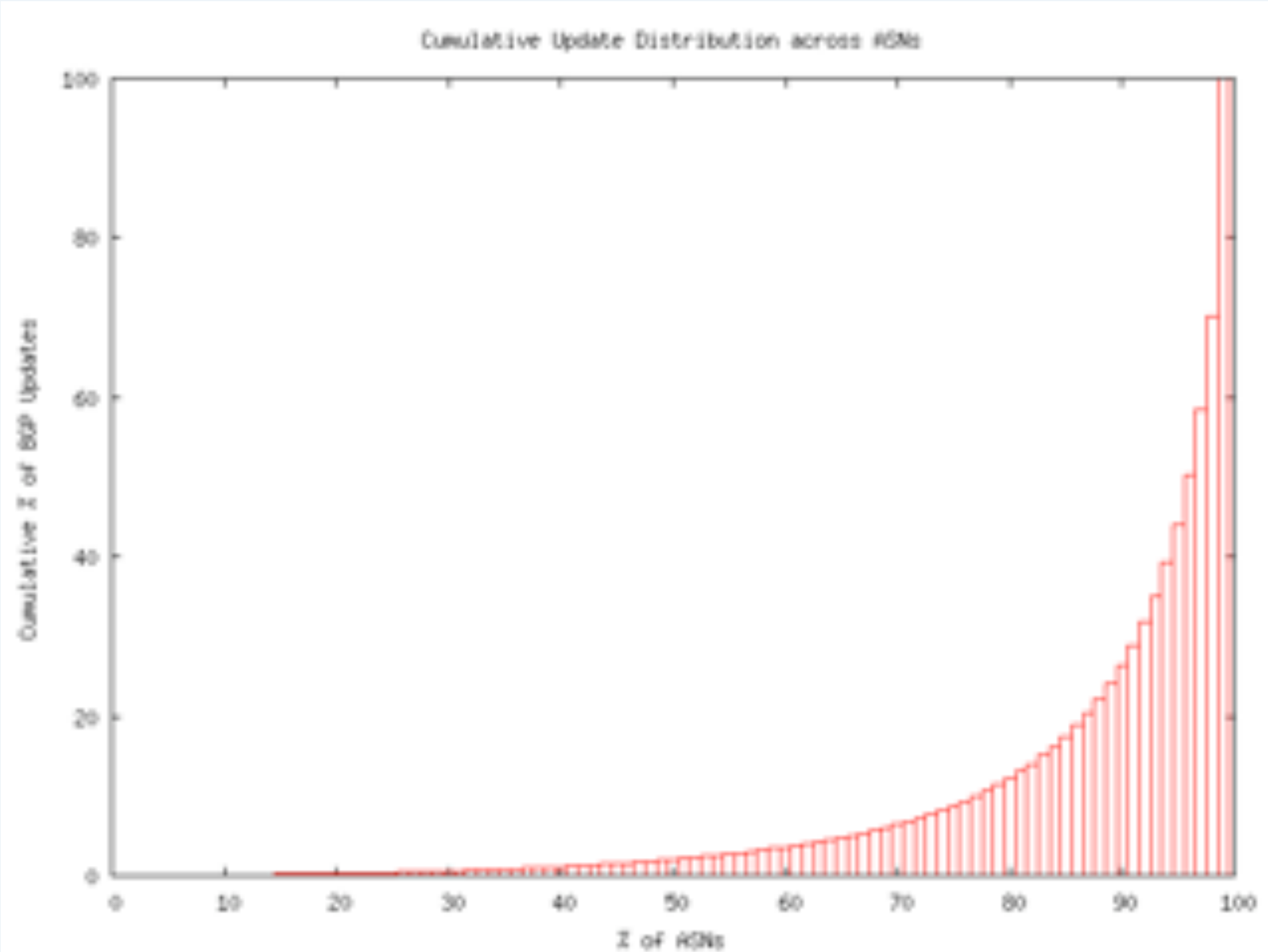


9 - 66.150.140.0/23

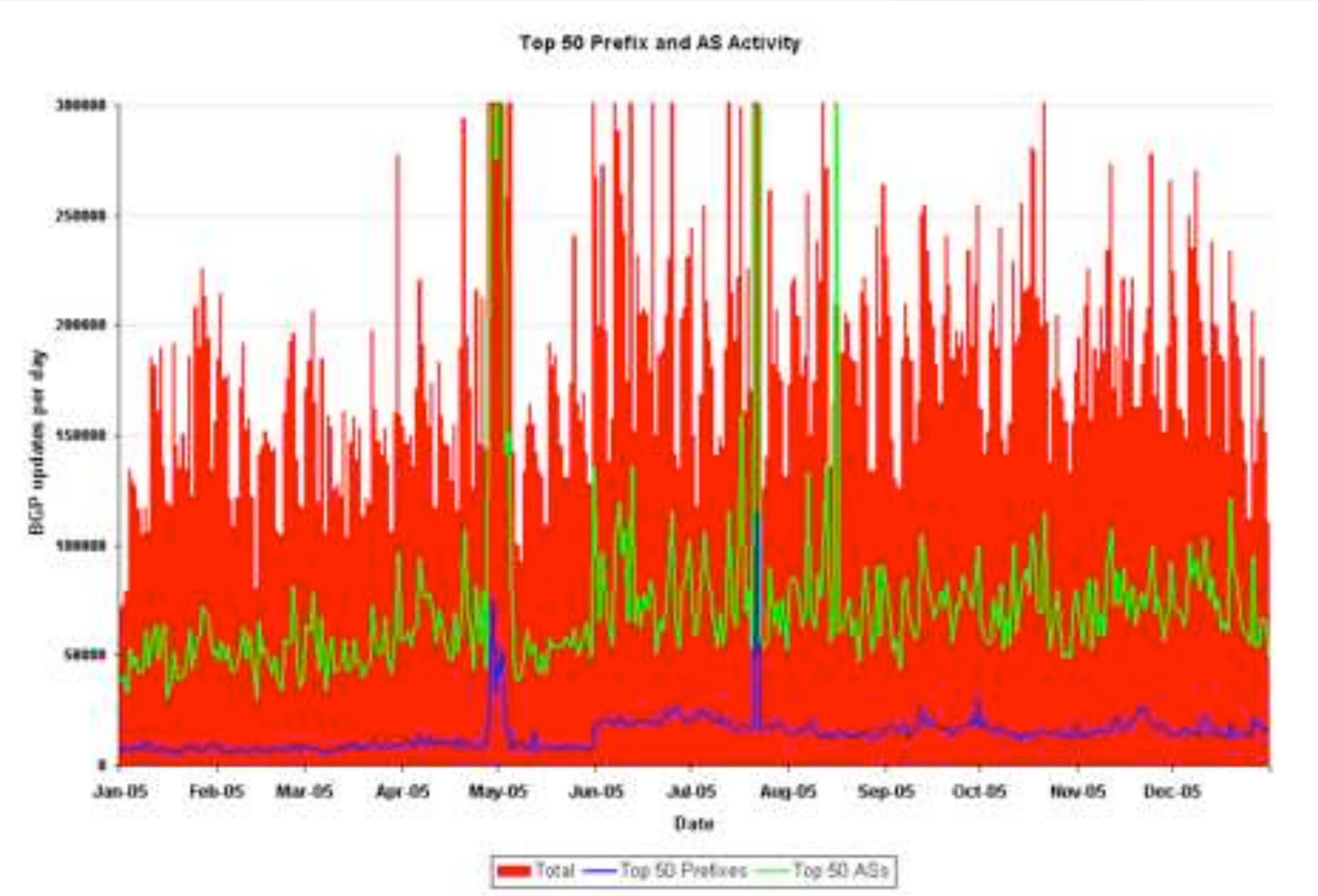


10 - 207.168.184.0/24

Distribution of Updates by AS



Distribution of Updates

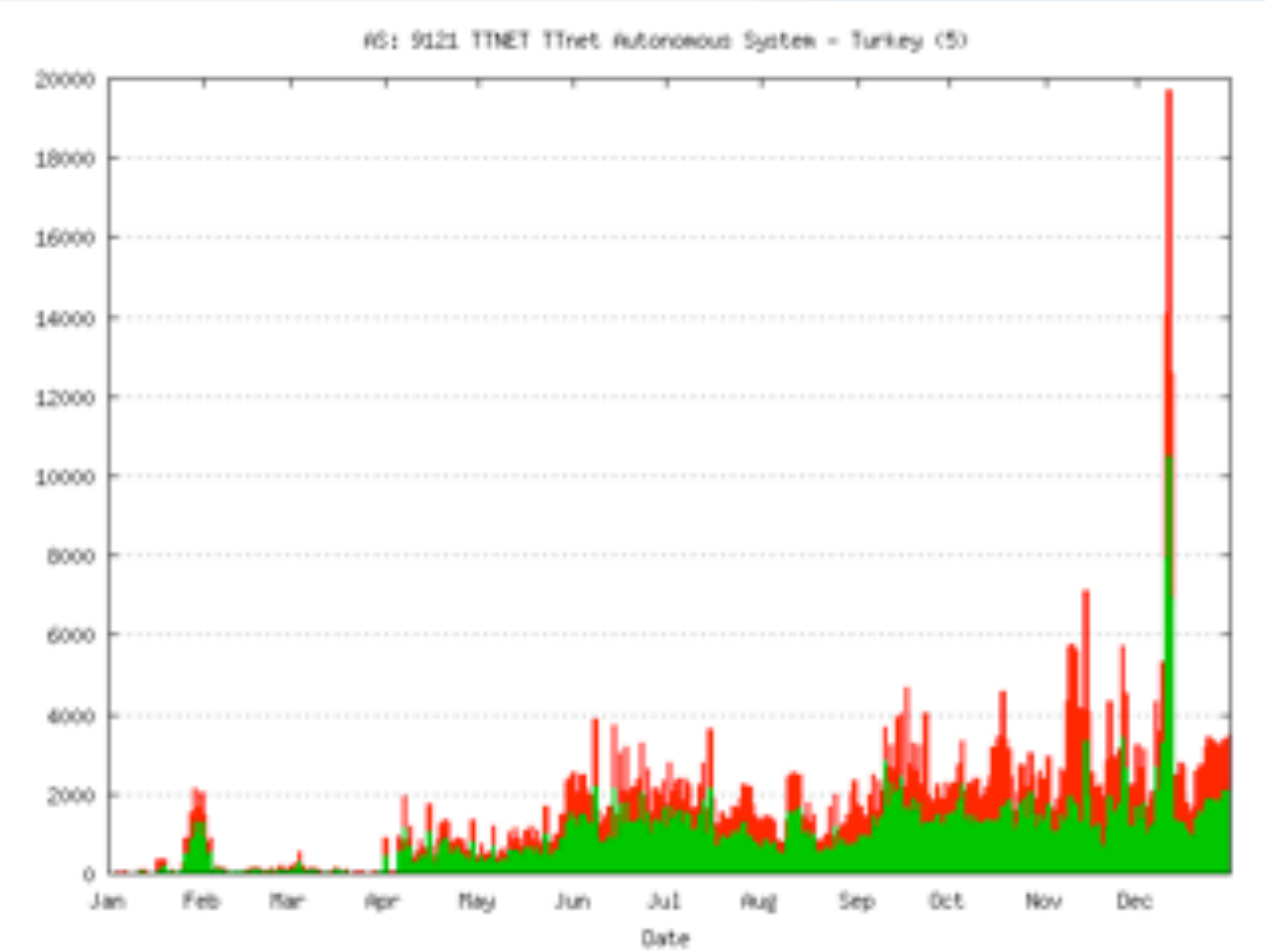


Active ASNs

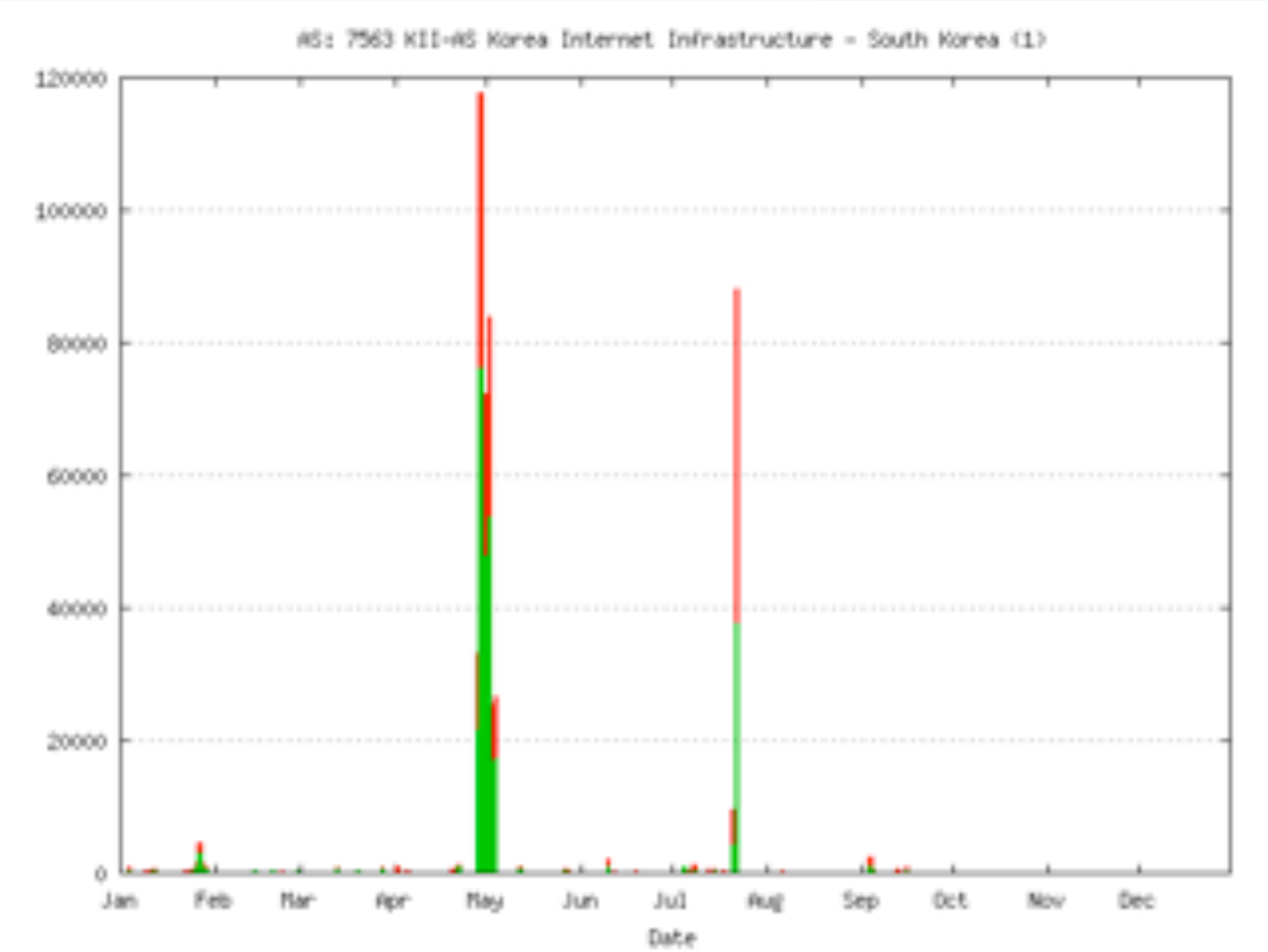
Top 10 ASns

	AS	Updates	Flaps	Re-Homes
1.	9121	970,782	349,241	206802
2.	7563	869,665	326,707	5
3.	702	605,090	232,876	144523
4.	17557	576,974	178,044	175275
5.	17974	569,806	198,948	310
6.	7545	562,879	200,425	8931
7.	721	498,297	175,623	35866
8.	2706	418,542	196,136	16945
9.	9950	411,617	148,725	6
10.	17832	393,052	143,018	0

1 – AS 9121

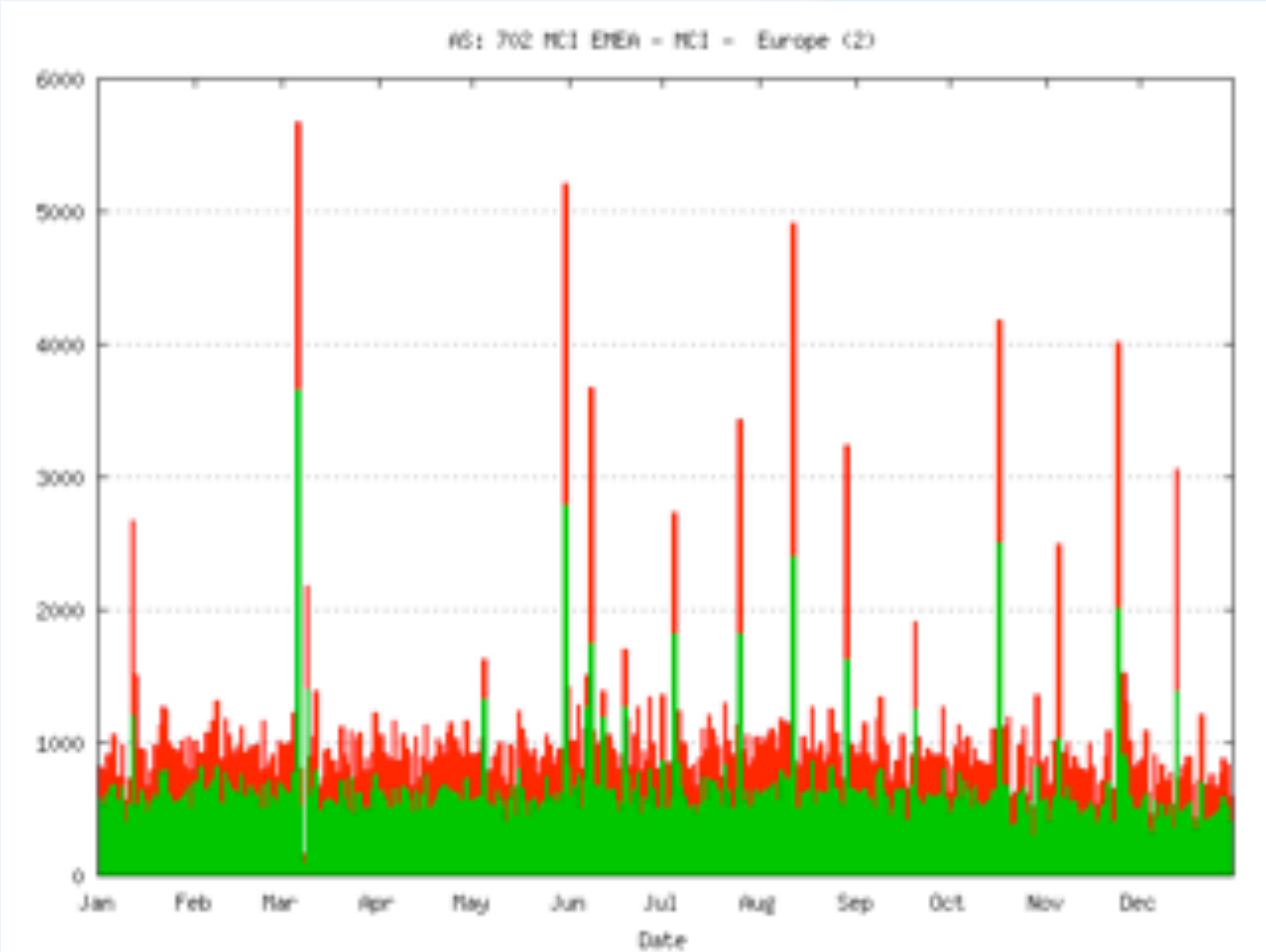

AS: 9121 TTNET TTnet Autonomous System - Turkey (5)
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Date

2 – AS 7563



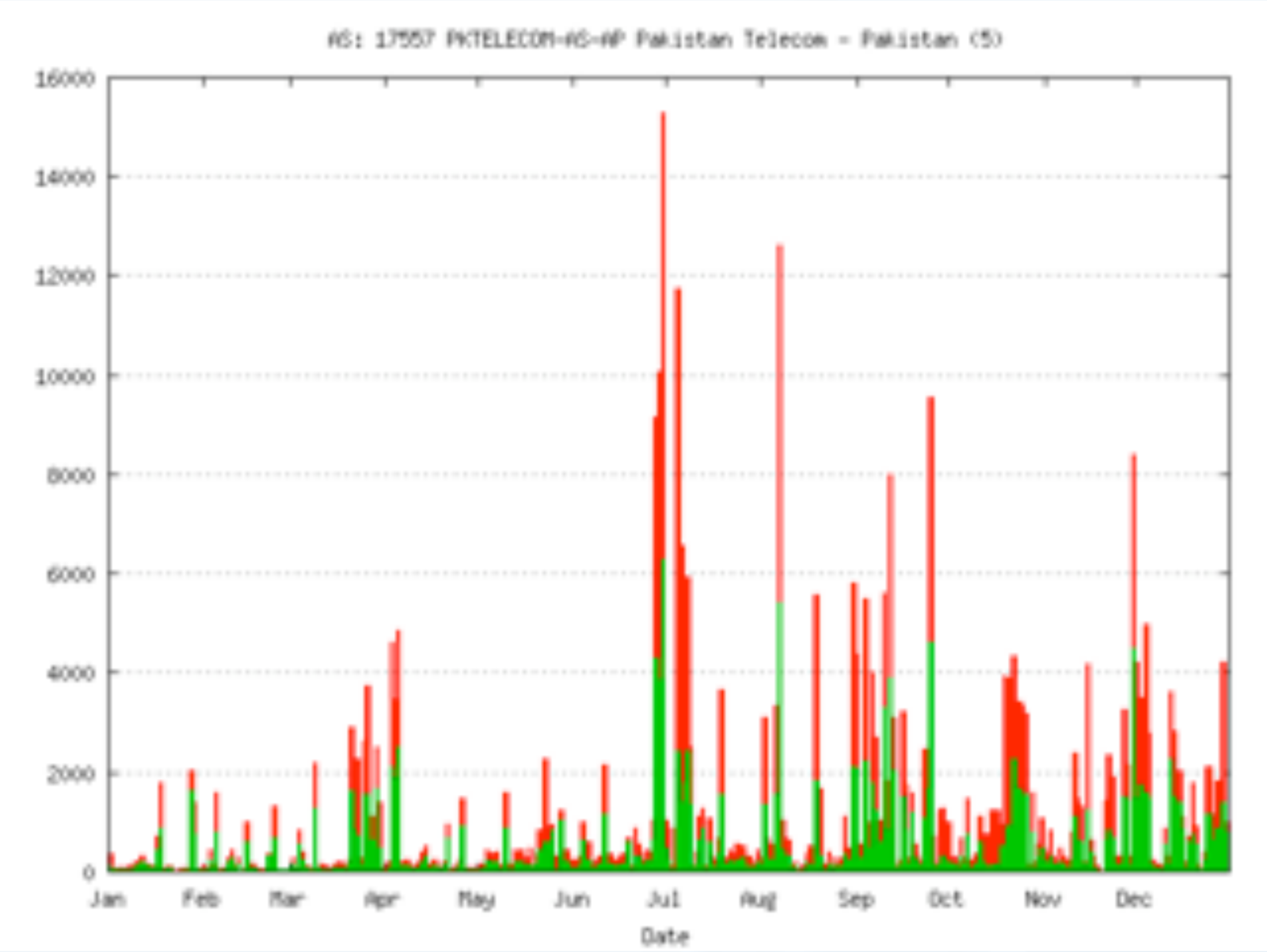
AS: 7563 KII-AS Korea Internet Infrastructure - South Korea (1)
120000
100000
80000
60000
40000
20000
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Date

3 – AS 702

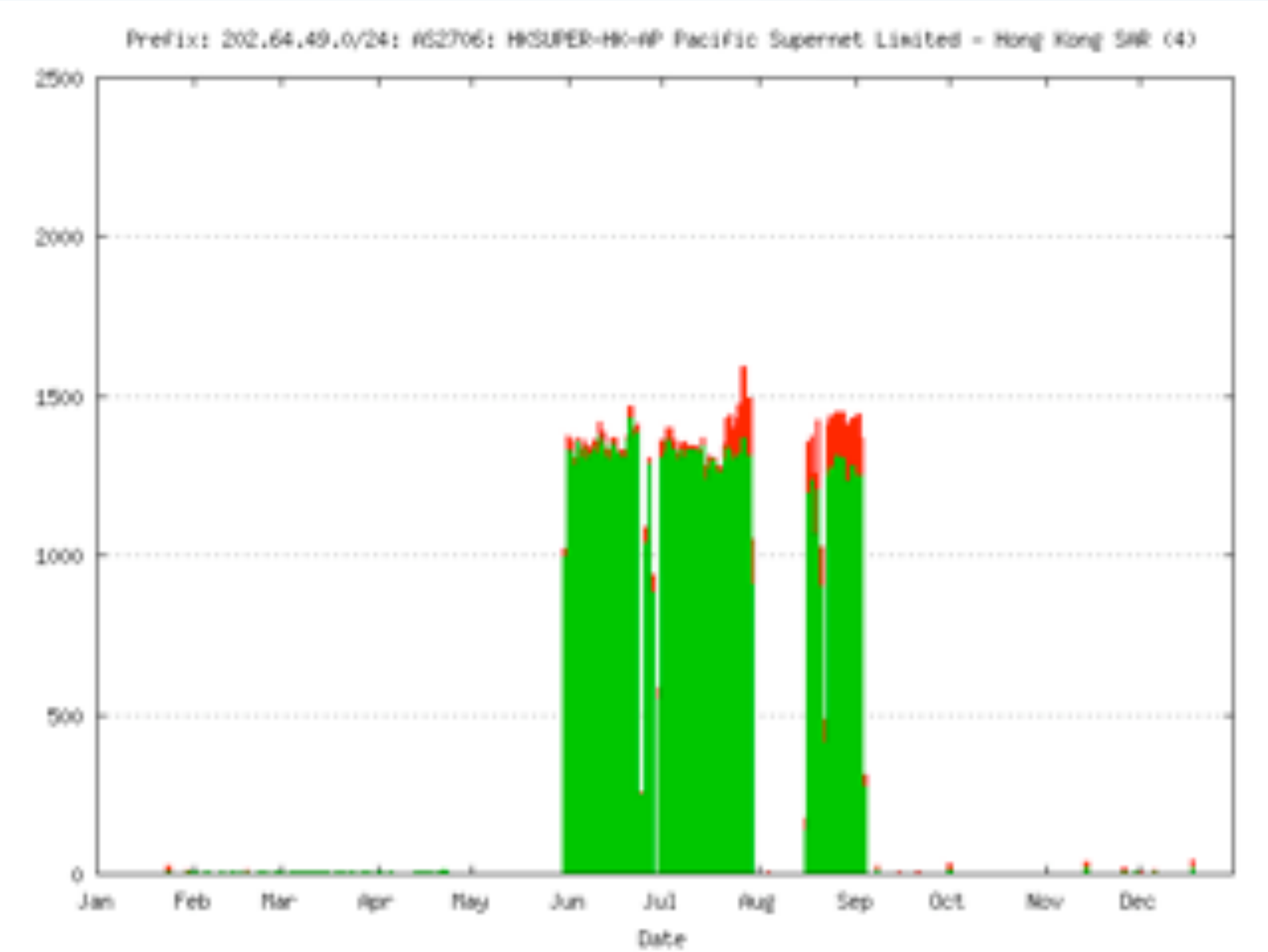


AS: 702 MCI EMEA - MCI - Europe (2)
6000
5000
4000
3000
2000
1000
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Date

4 – AS 17557

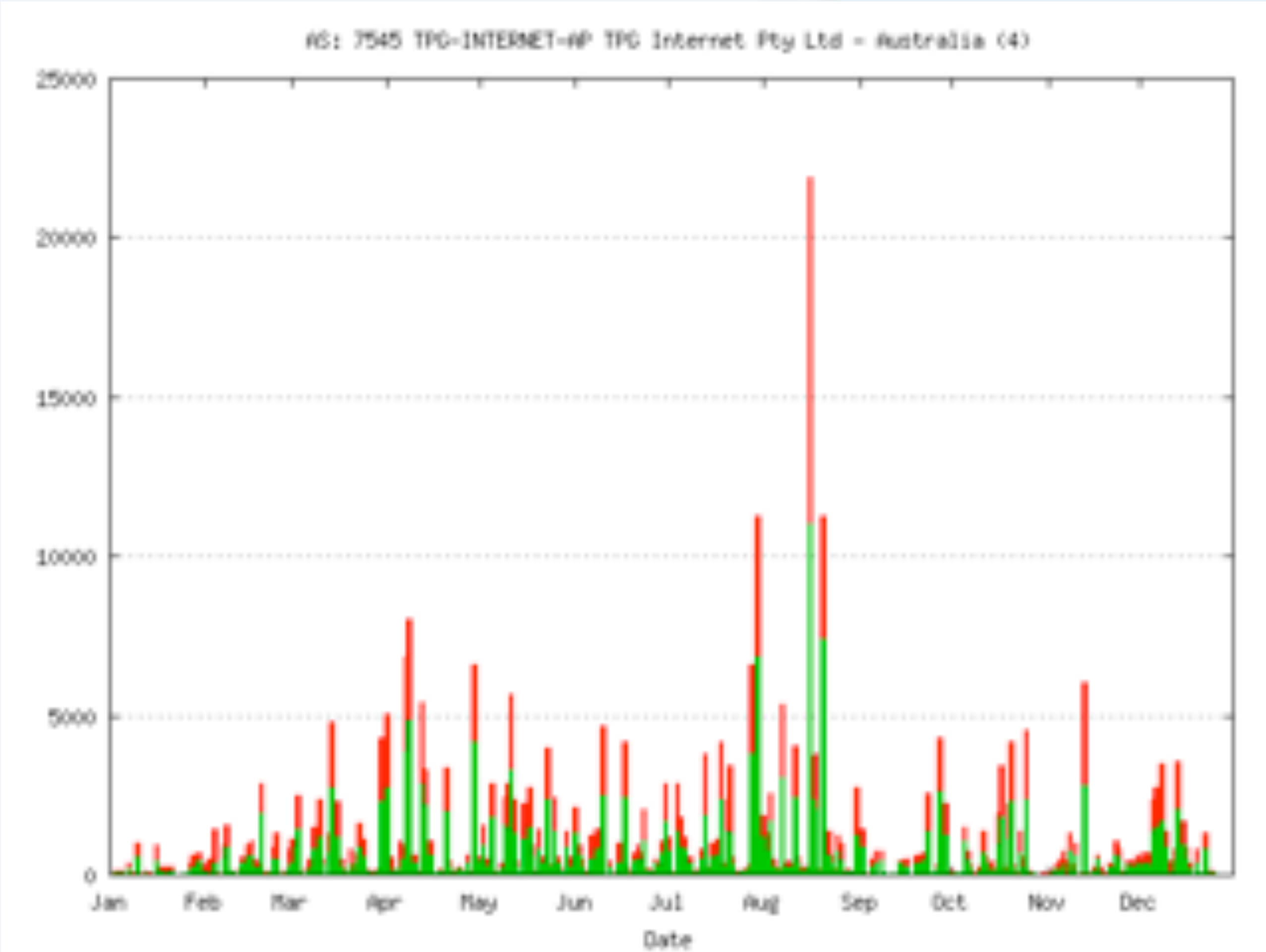

AS: 17557 PKTELECOM-AS-AP Pakistan Telecom - Pakistan (5)
16000
14000
12000
10000
8000
6000
4000
2000
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Date

5 – AS17974

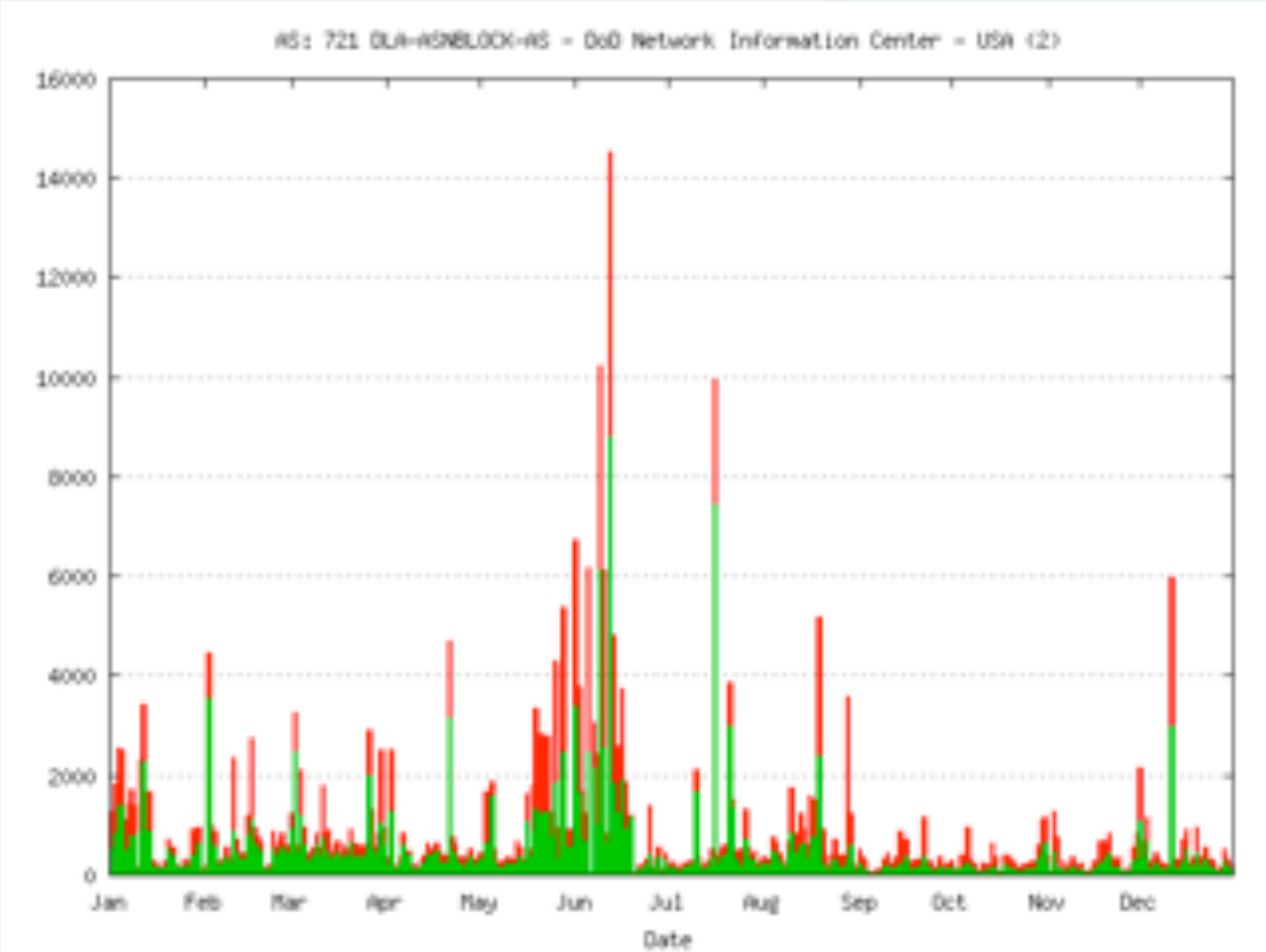


Prefix: 202.64.49.0/24: AS2706: HKSUPER-HK-AP Pacific Supernet Limited - Hong Kong SAR (4)
2500
2000
1500
1000
500
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Date

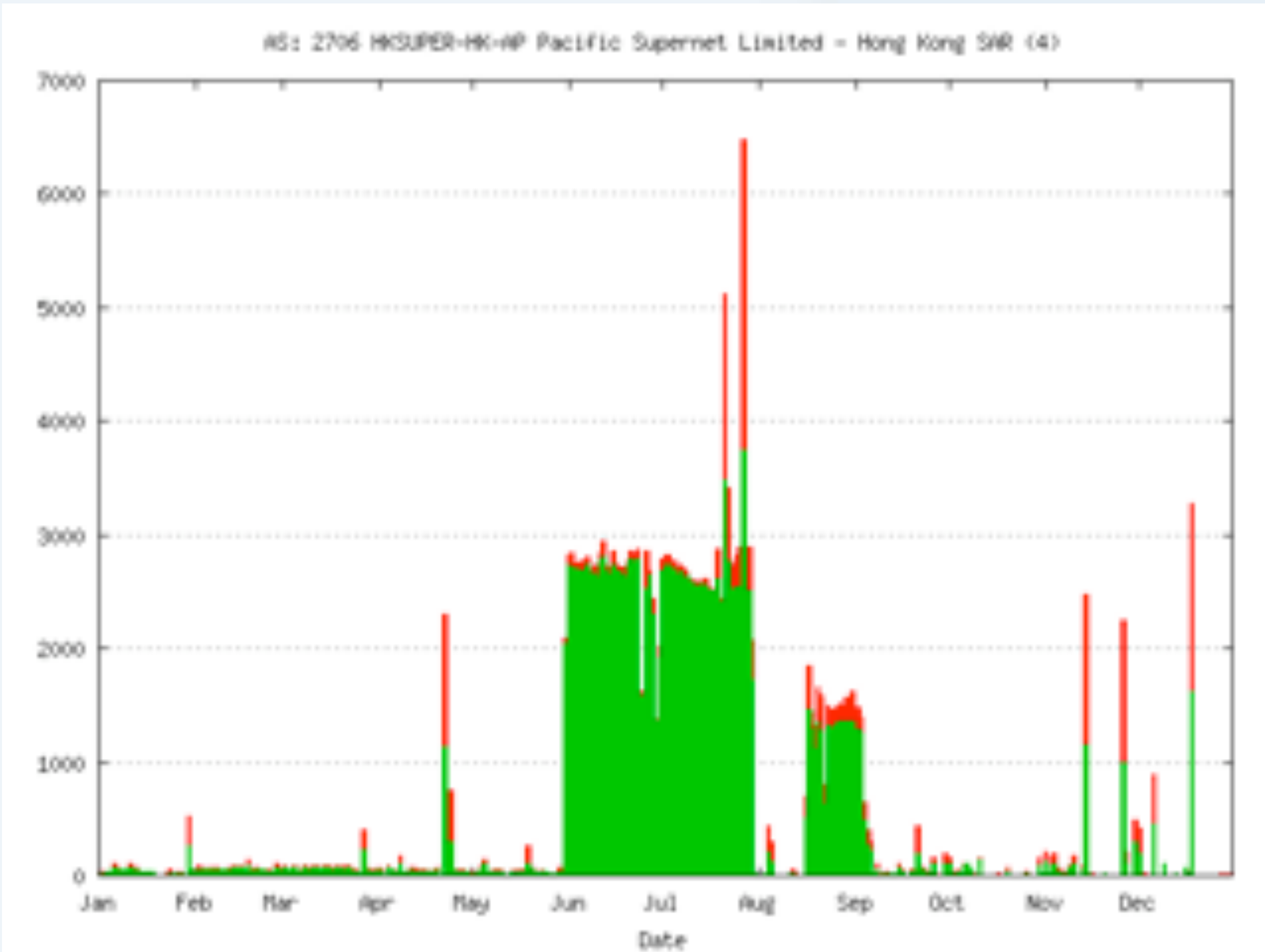
6 – AS 7545


AS: 7545 TPG-INTERNET-AP TPG Internet Pty Ltd - Australia (4)
25000
20000
15000
10000
5000
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Date

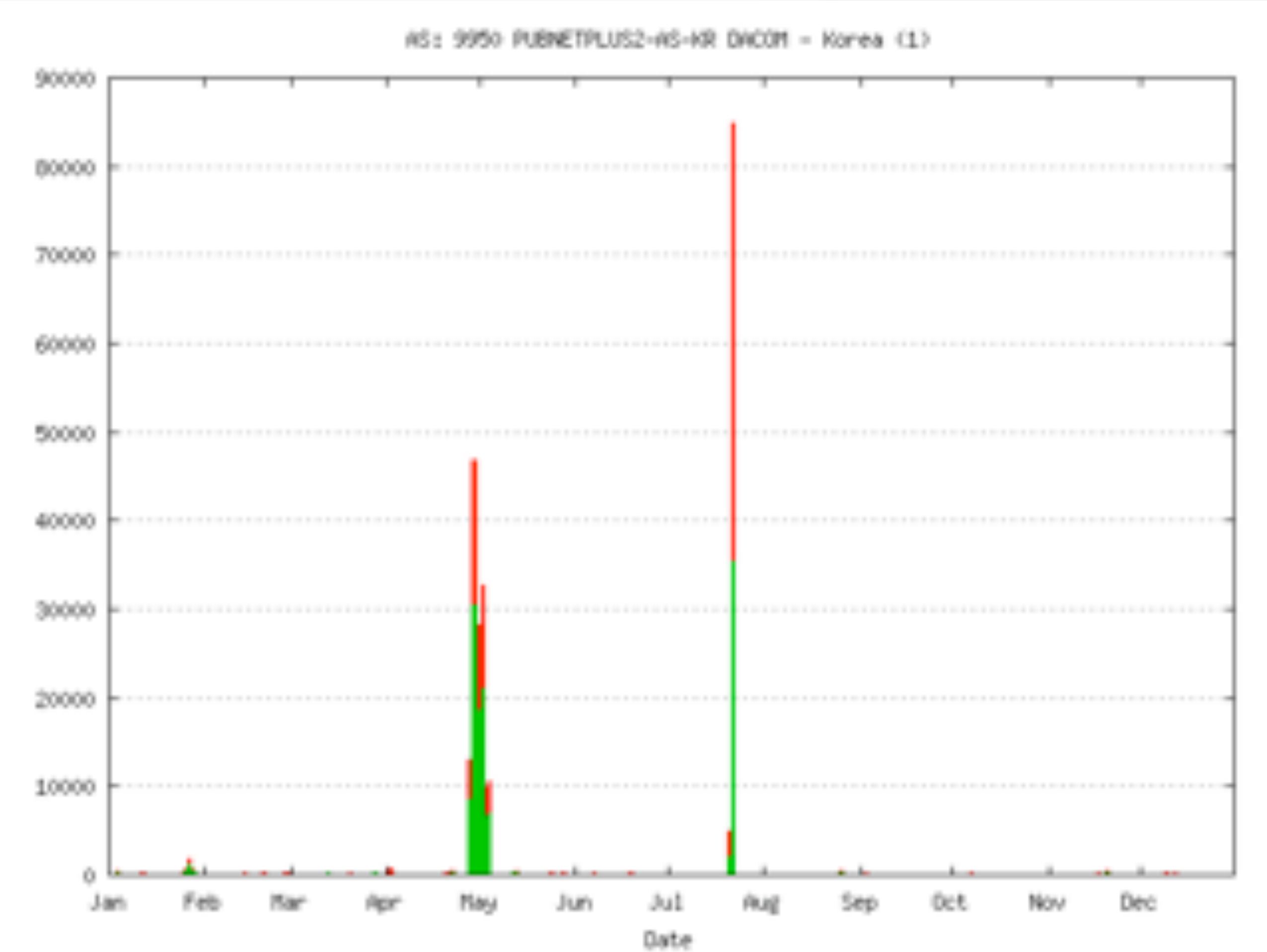
7 – AS721


AS: 721 DLA-ASNBLOCK-AS - DoD Network Information Center - USA (2)
16000
14000
12000
10000
8000
6000
4000
2000
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Date

8 – AS2706

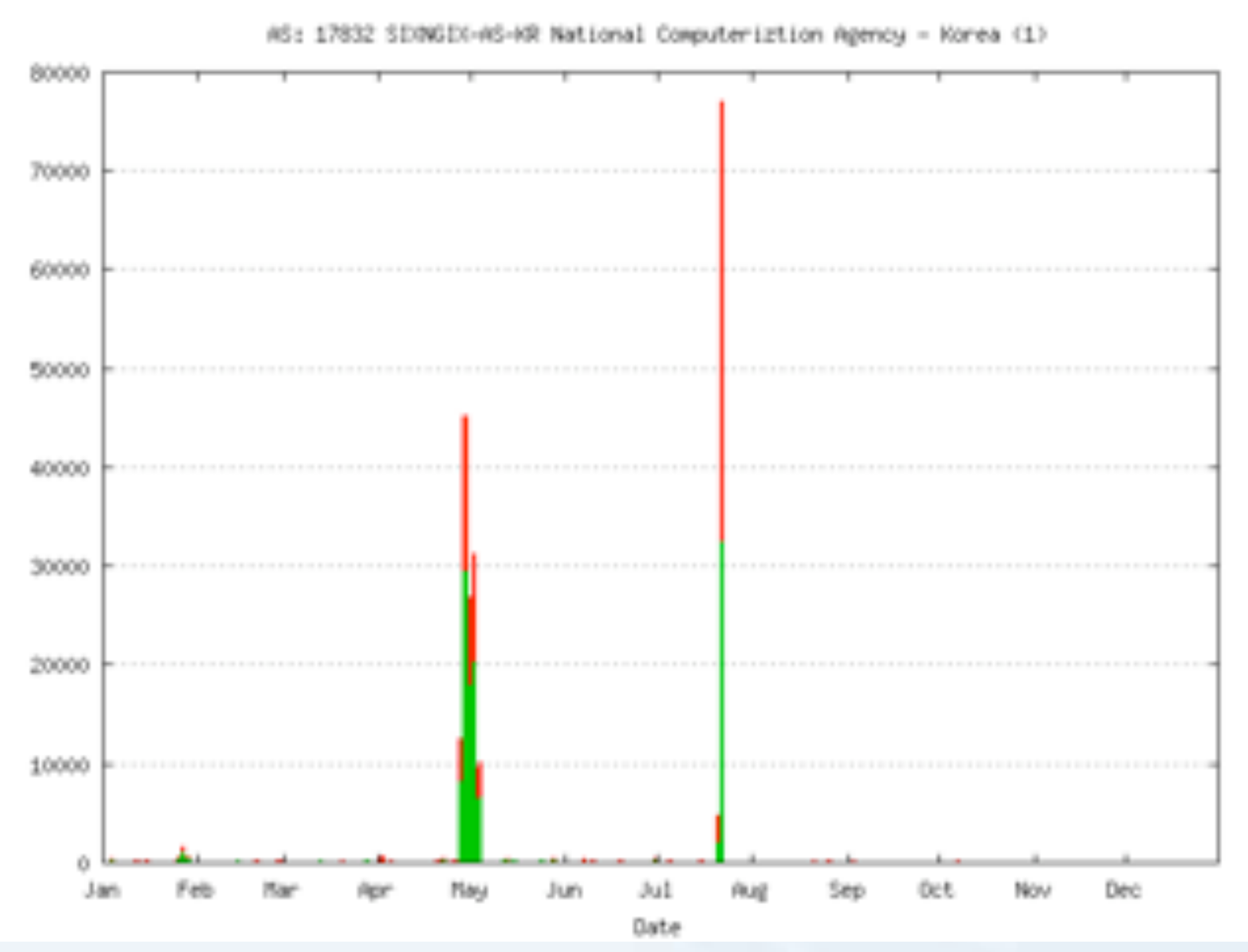

AS: 2706 HKSUPER-HK-AP Pacific Supernet Limited - Hong Kong SAR (4)
7000
6000
5000
4000
3000
2000
1000
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Date

9 – AS9950



AS: 9950 PUBNETPLUS2-AS-KR DACOM - Korea (1)
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Date

10 – AS17832



AS: 17832 SIXNGIX-AS-KR National Computeriztion Agency - Korea (1)
80000
70000
60000
50000
40000
30000
20000
10000
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Date

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

The Uncertainty Factor

- Given that the overwhelming majority of updates are being generated by a very small number of sources, the level of uncertainty in extrapolation of trend models of BGP update rates is extremely high
- This implies that the predictions of router capabilities in a 3 – 5 year interval is also extremely uncertain

Thank You