BGP: 2008

Geoff Huston
APNIC

Some BGP-related questions:

The Network Operator:

- How big a router should you buy today if you want it to cope with BGP in 3 years time?
- What FIB size?
- What processing capability?
- What about if you are looking for an operational lifespan of 5 years?

More BGP Questions

The Protocol Engineer:

- Is BGP scaling or failing?
- Do we need to develop a new IDR protocol?
- How much time do we have?

And a few more as well

The Network Architecture Researcher:

- Are the Internet's concepts of names and addresses adequate?
- Are alternate models of id / loc split more friendly to routing?
- Is routing scaling an intractable problem within the confines of the current architecture?

One Approach:

guess!

Or:

you could see what answers a more disciplined approach to these questions could provide!

An approach to generating some answers

- Use real data about today's network
  - Understand the actual data about IDR and the operation of BGP
  - Use a long base line of consistent observations of BGP behaviour
  - Analyze the data carefully

BGP measurements

- There are a number of ways to “measure” BGP:
  - Assemble a large set of BGP peering sessions and record everything
    - RIPE NCC’s RIS service
    - Route Views
  - Perform carefully controlled injections of route information and observe the propagation of information
    - Beacons
    - AS Set manipulation
    - Bogon Detection and Triangulation
  - Take a single BGP perspective and perform continuous recording of a number of BGP mterics

AS131072 (or AS2.0) BGP measurement

- Successor to the AS1221 observation point
- Data collection since 1 July 2007 (since 2000 for AS1221)
- Passive data measurement technique (no advertisements or probes)
- Quagga platform, connected to AS4608 and AS4777 via eBGP
- IPv4 and IPv6 simultaneous
- Archive of all BGP updates and daily RIB dumps
- Data and reports are continuously updated and published: http://bgp.potaroo.net

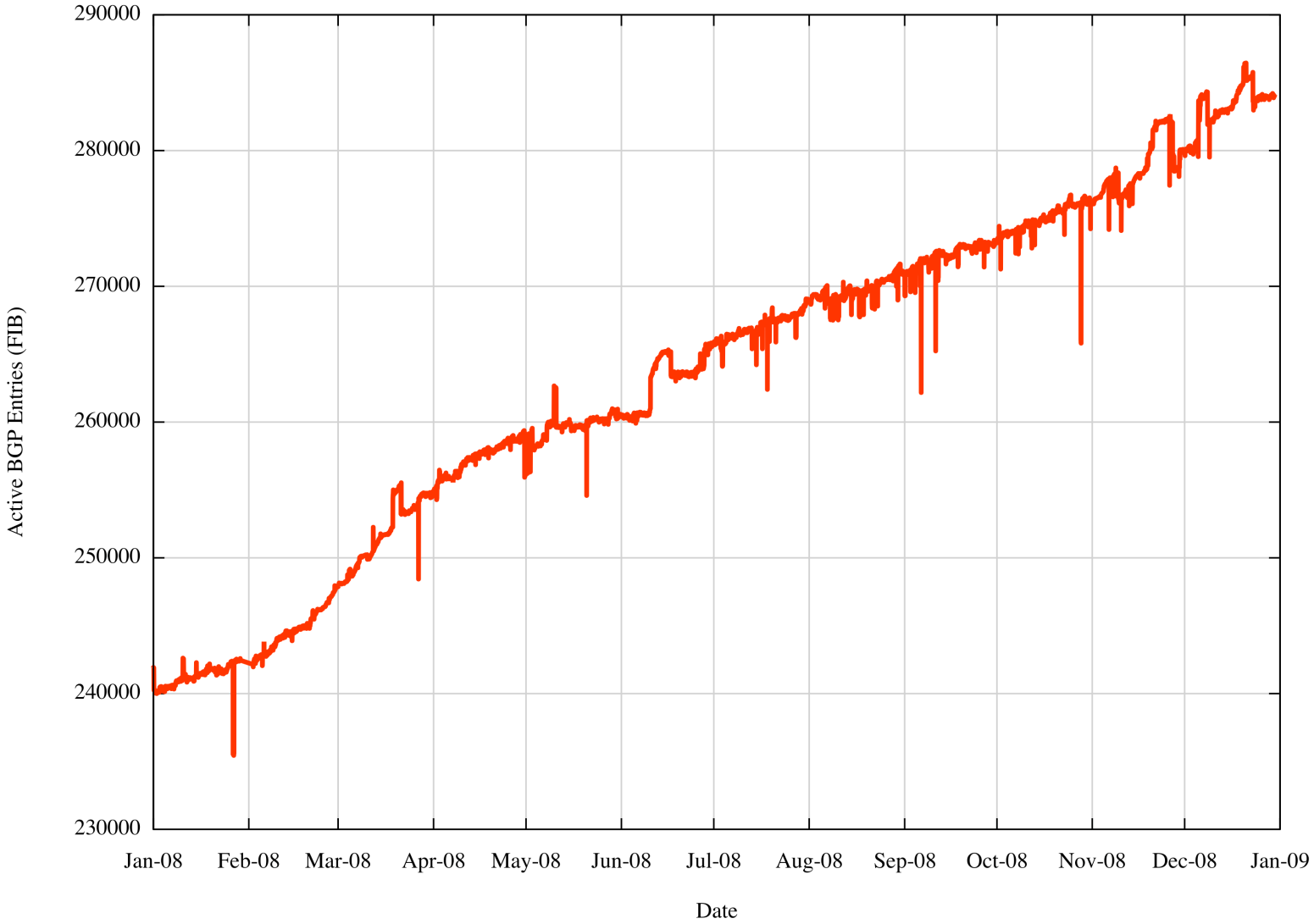
Some Caveats

- This is a measurement at the EDGE, not in the MIDDLE
- It is a single stream measurement, not an aggregated measurement
- This is a measurement of the ‘production network’ used for forwarding traffic
- There is NO iBGP traffic being measured
- This is what an eBGP customer may see in terms of load for a single eBGP feed

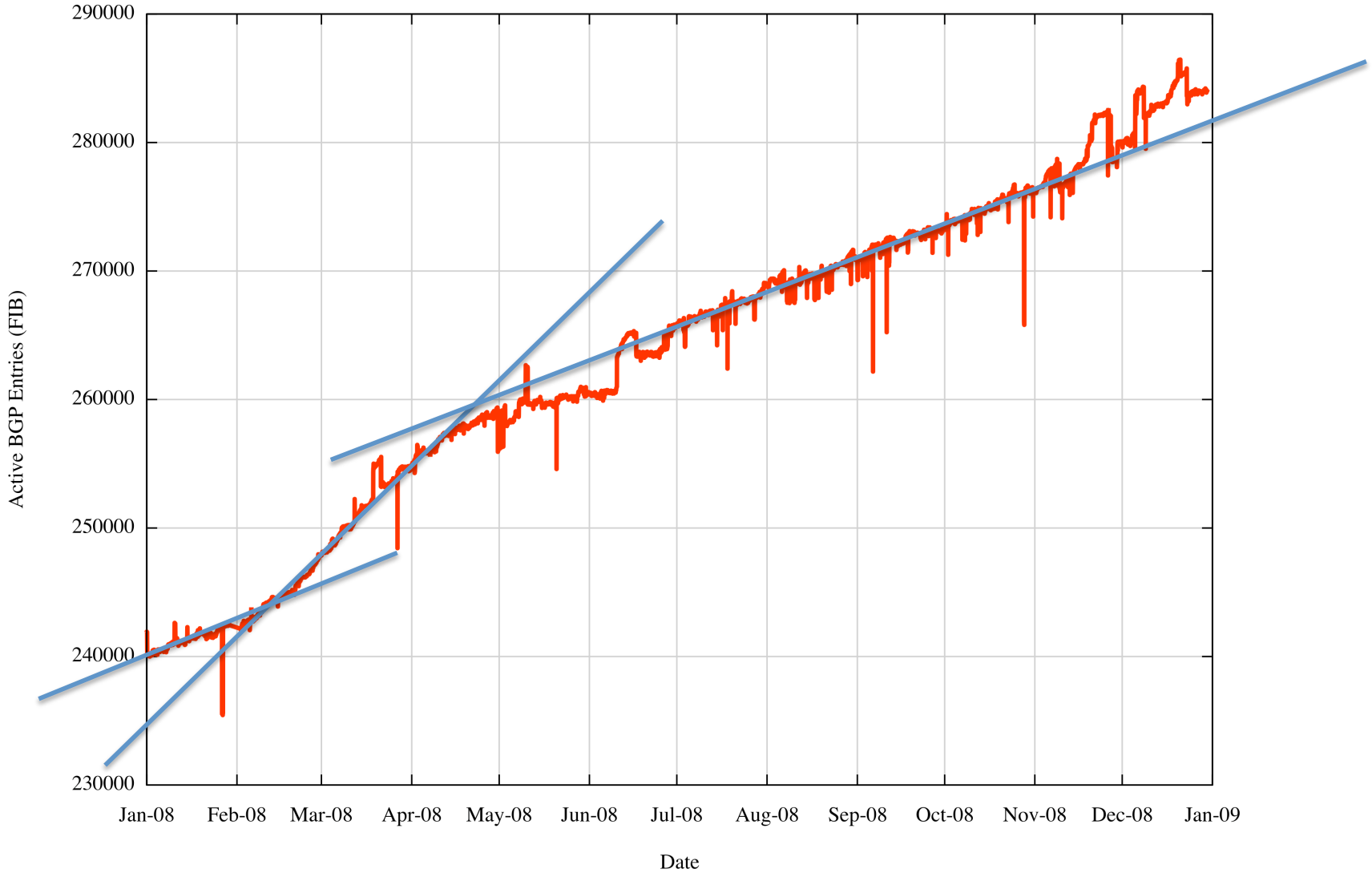
BGP: 2008

- Analysis of BGP data collected at AS131072 from 1 January 2008 to 31 December 2008

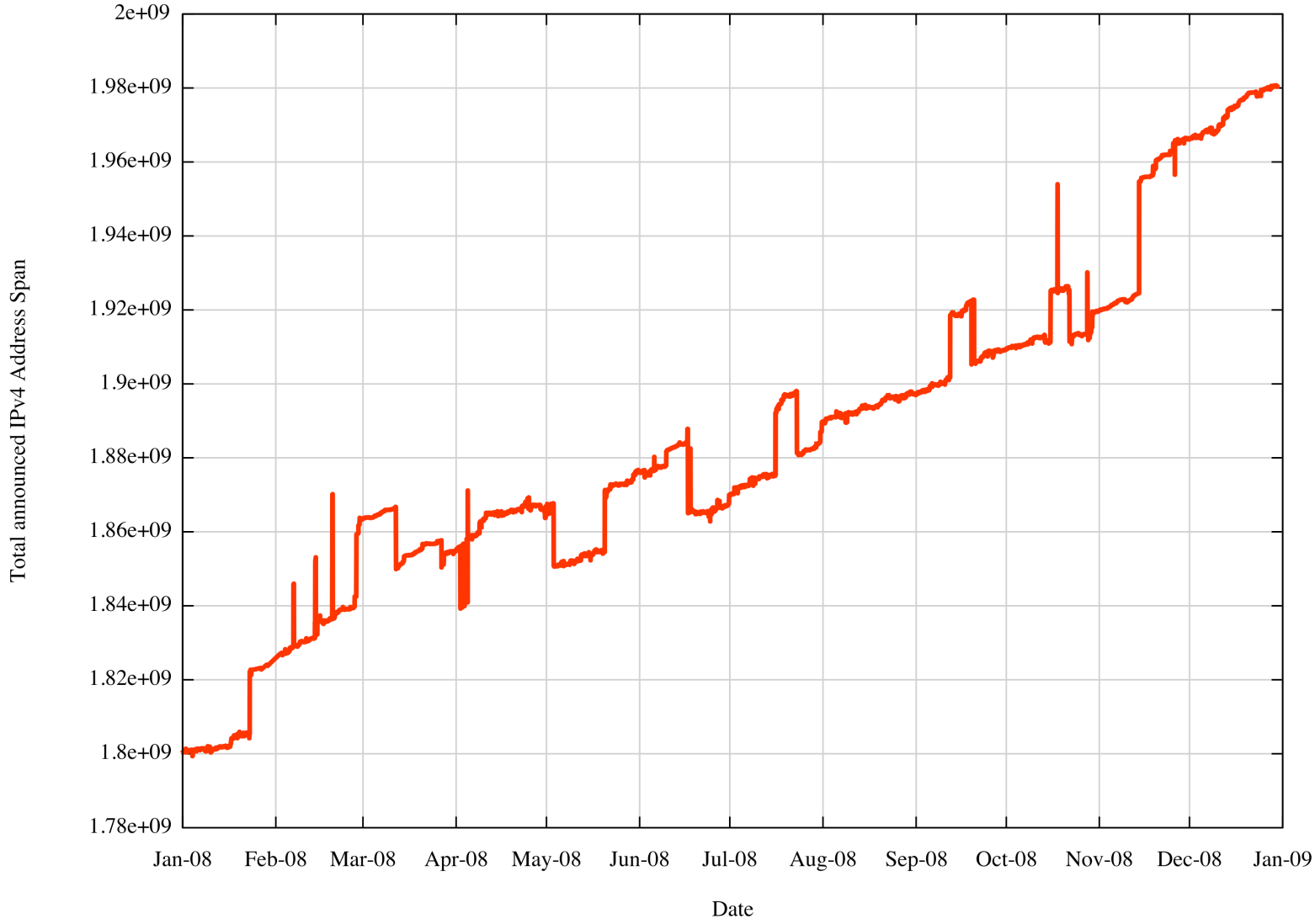
IPv4 BGP Prefix Count



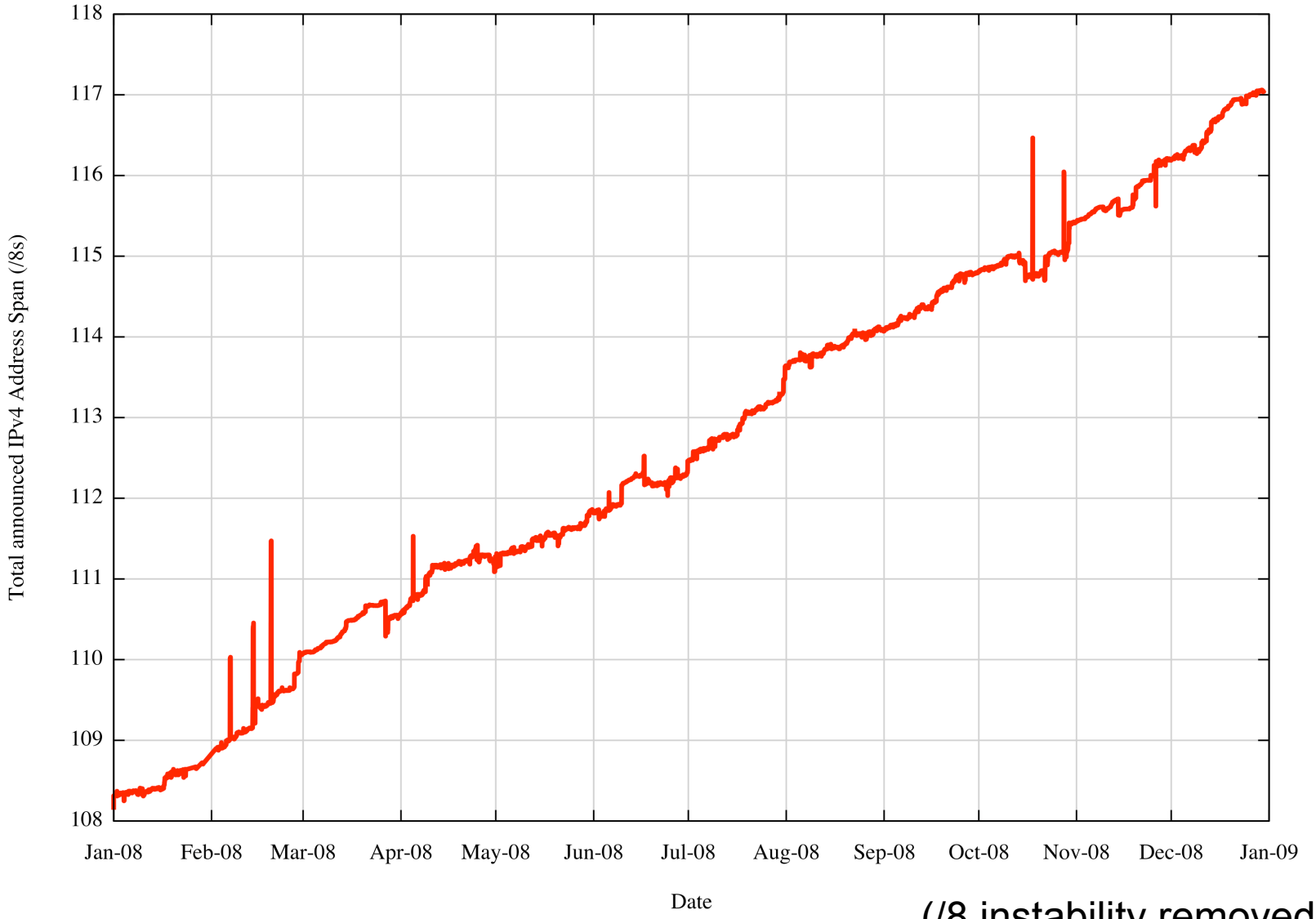
IPv4 BGP Prefix Count



IPv4 Routed Address Span

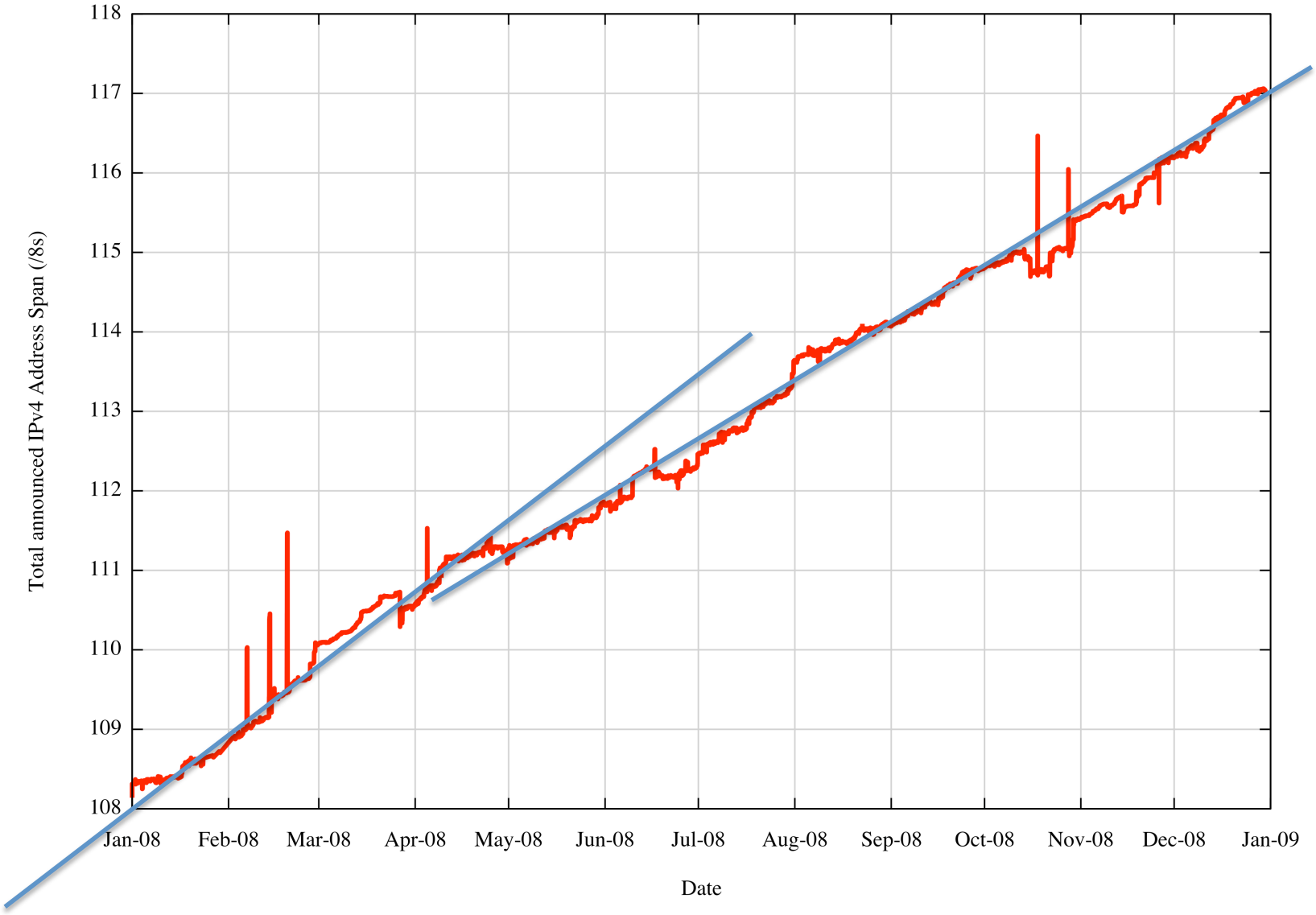


IPv4 Routed Address Span

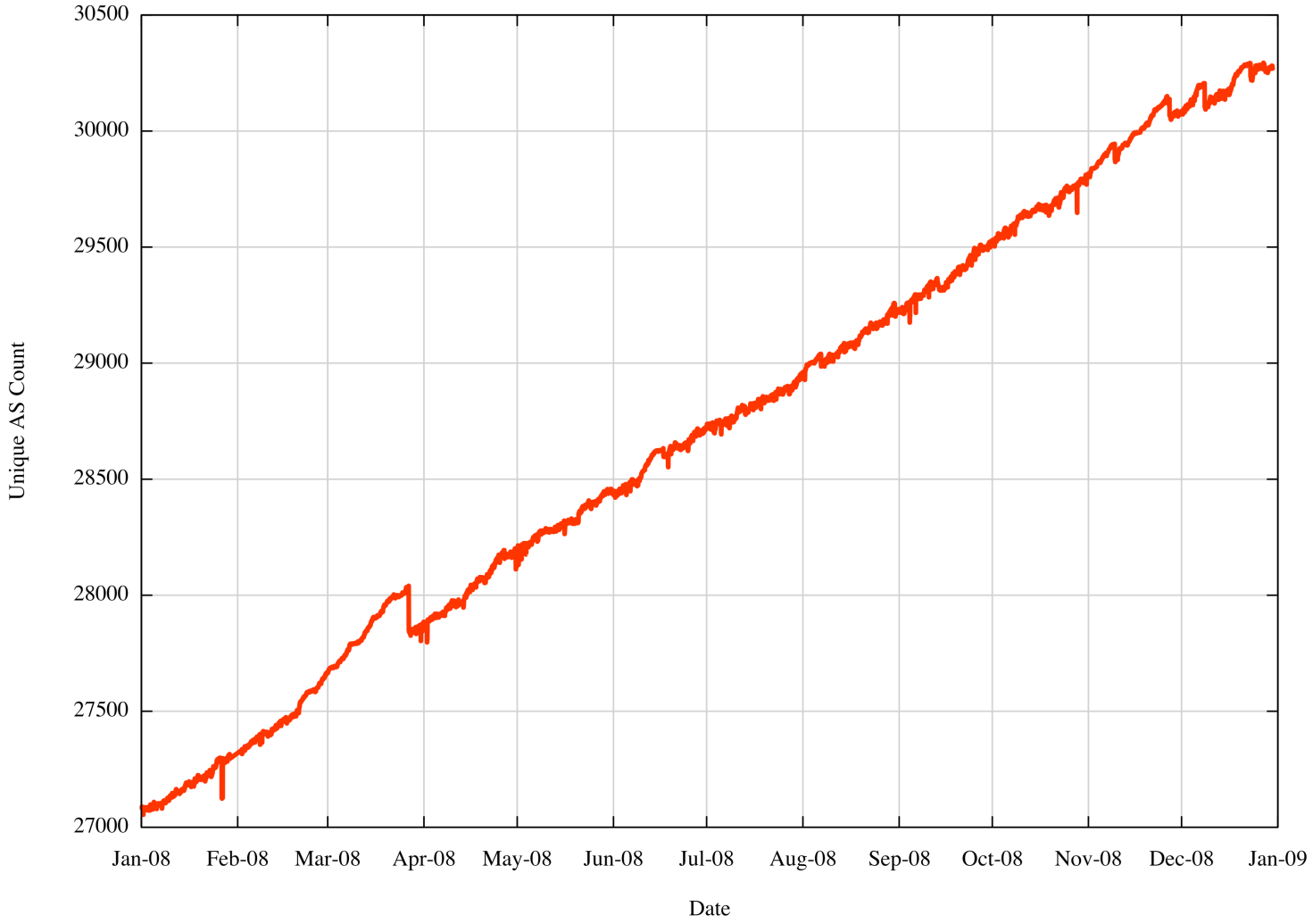


(/8 instability removed)

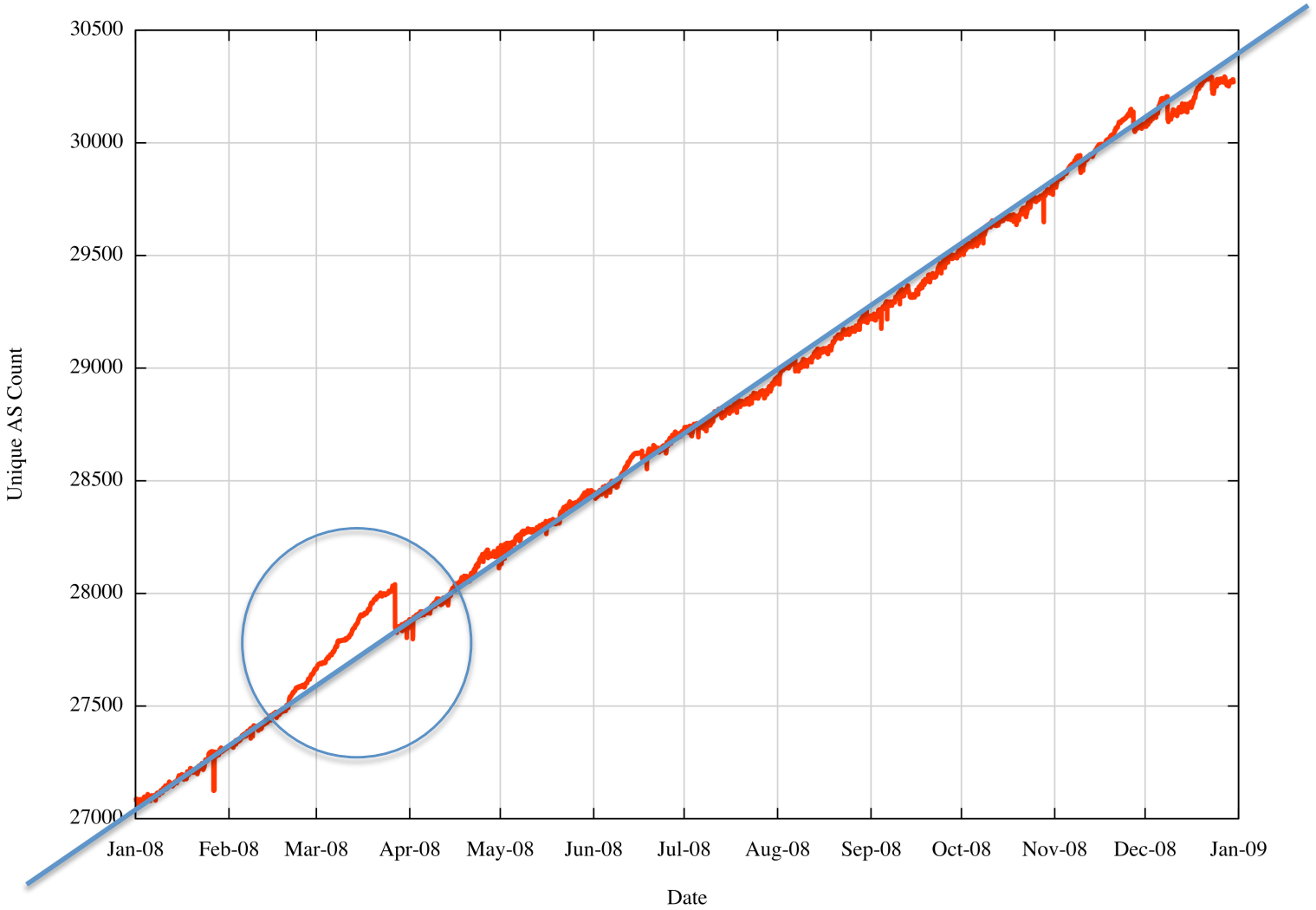
IPv4 Routed Address Span



IPv4 Routed AS Count



IPv4 Routed AS Count



IPv4 Vital Statistics for 2008

	Jan-08	Dec-08	
Prefix Count	245,000	286,000	**17%**
Roots	118,000	133,000	13%
More Specifics	127,000	152,000	20%
Address Span	106.39	118.44	**11%**
AS Count	27,000	30,300	**11%**
Transit	3,600	4,100	14%
Stub	23,400	26,200	11%

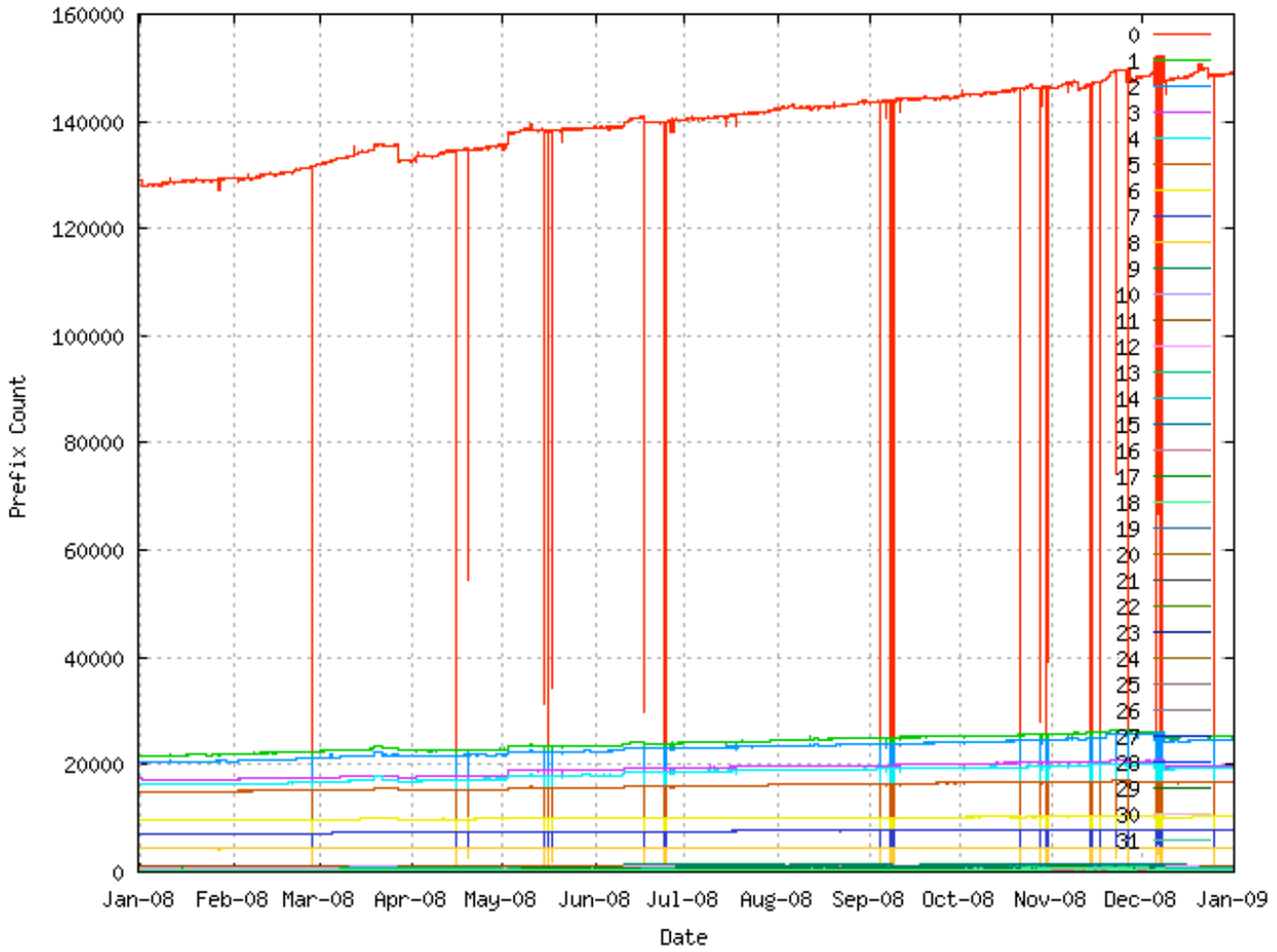
Some Observations

- Growth in IPv4 deployment slowed considerably as of the end of April 2008
  - Is this a possible consequence of the financial crash of 2008?
- Fragmentation of the IPv4 routing space continues to grow at a faster pace than underlying growth of the network itself

IPv4 prefix distribution

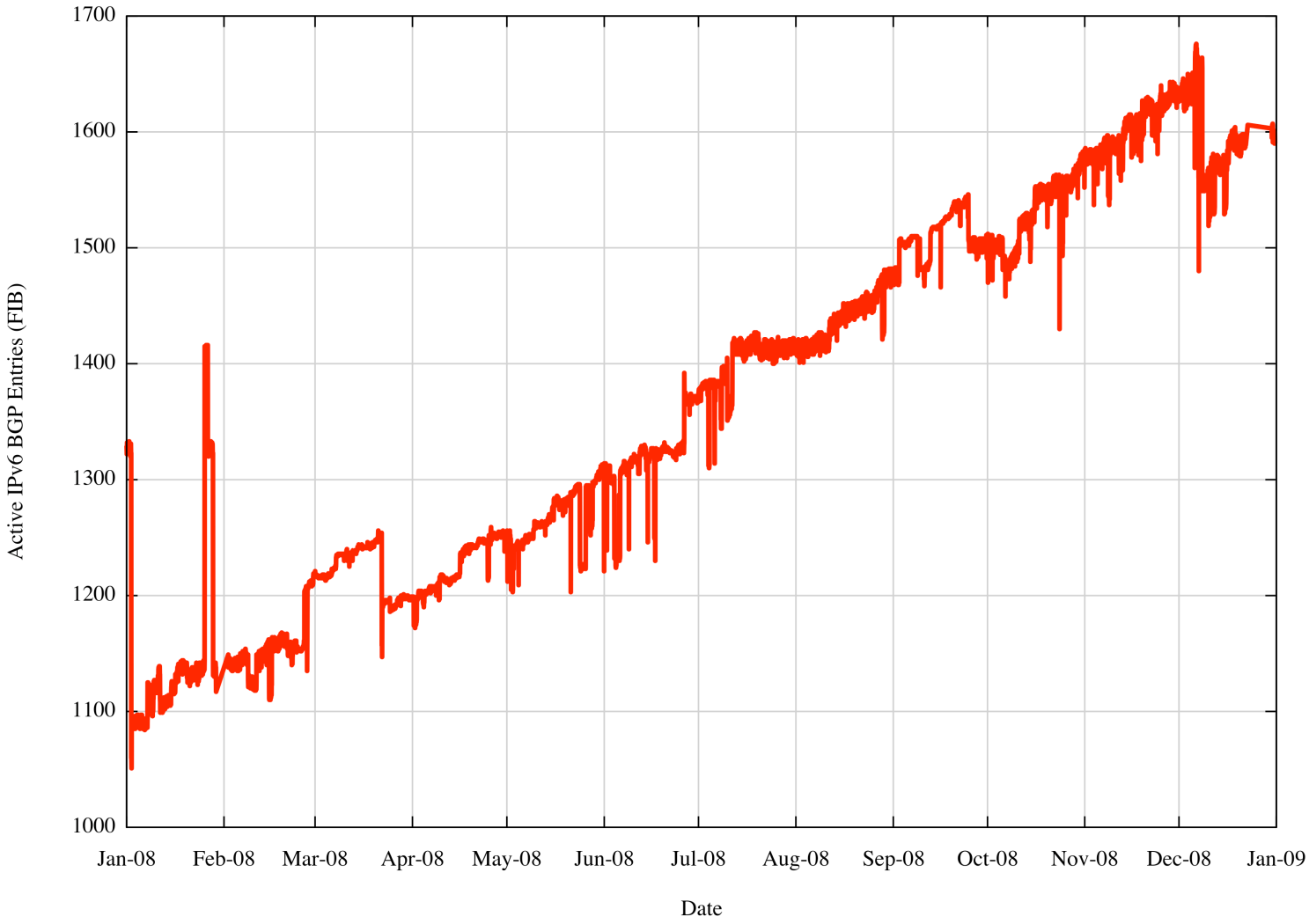
- Its all about /24's

IPv4 prefix distribution

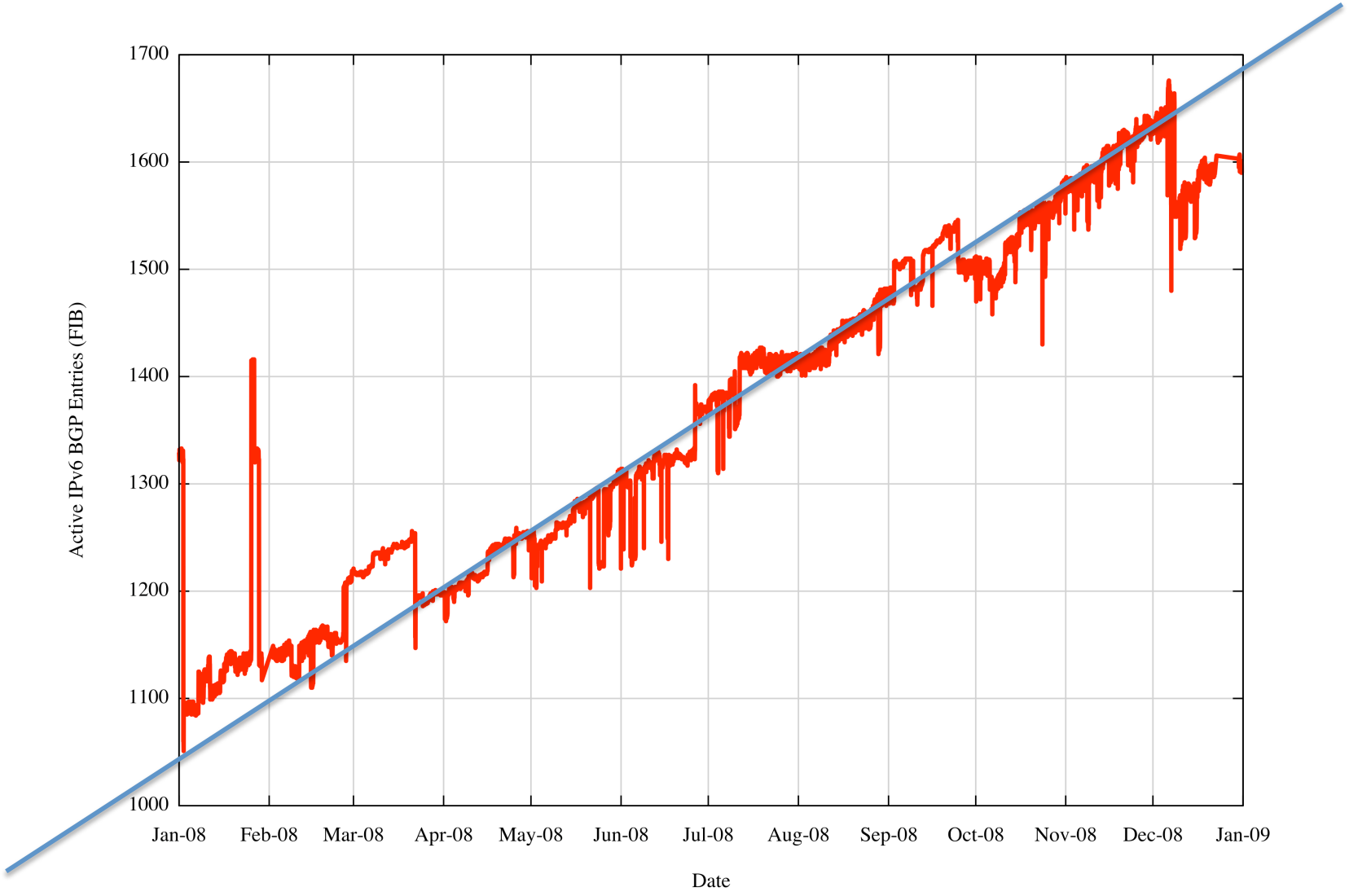


IPv6 BGP in 2008

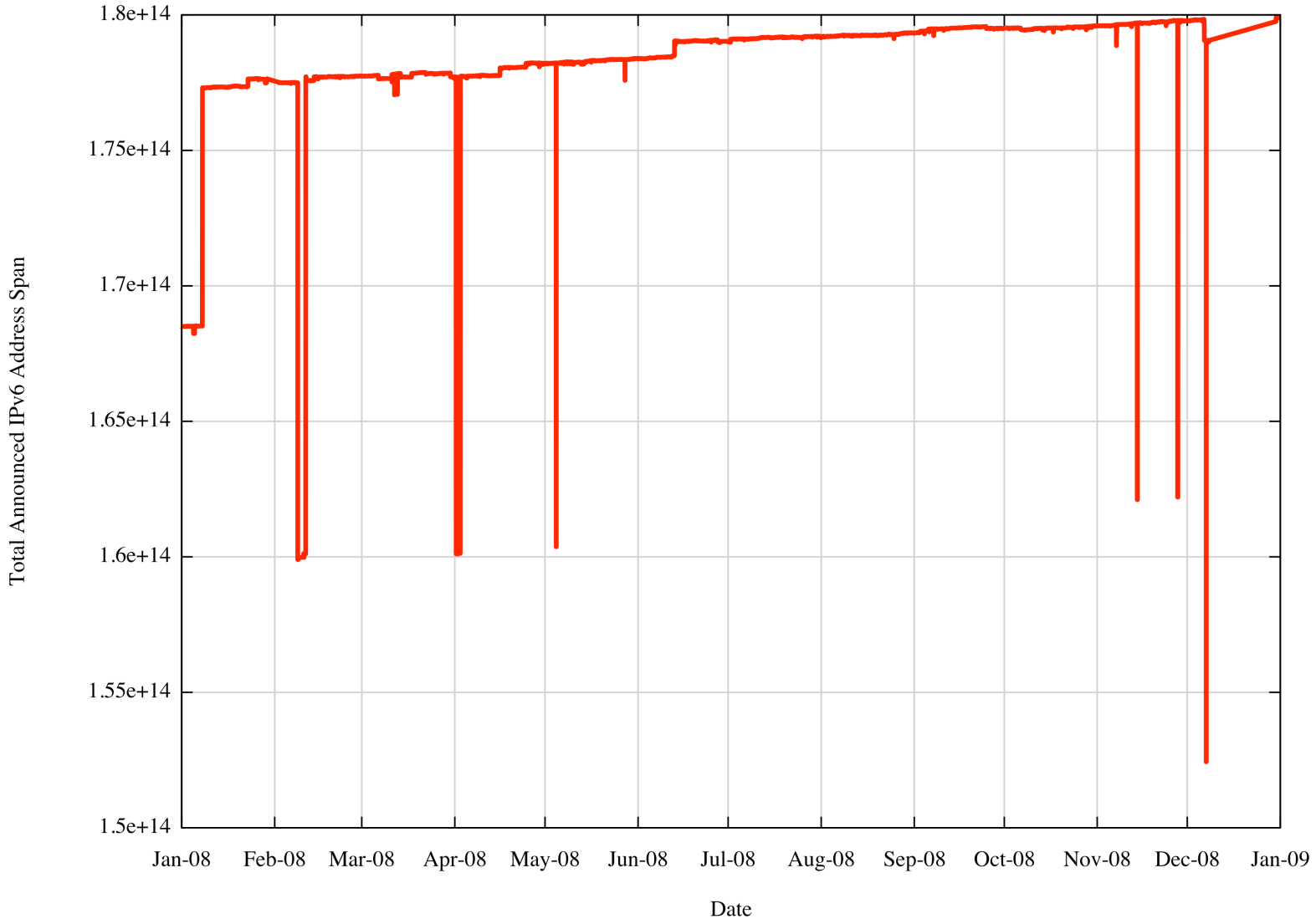
IPv6 BGP Prefix Count



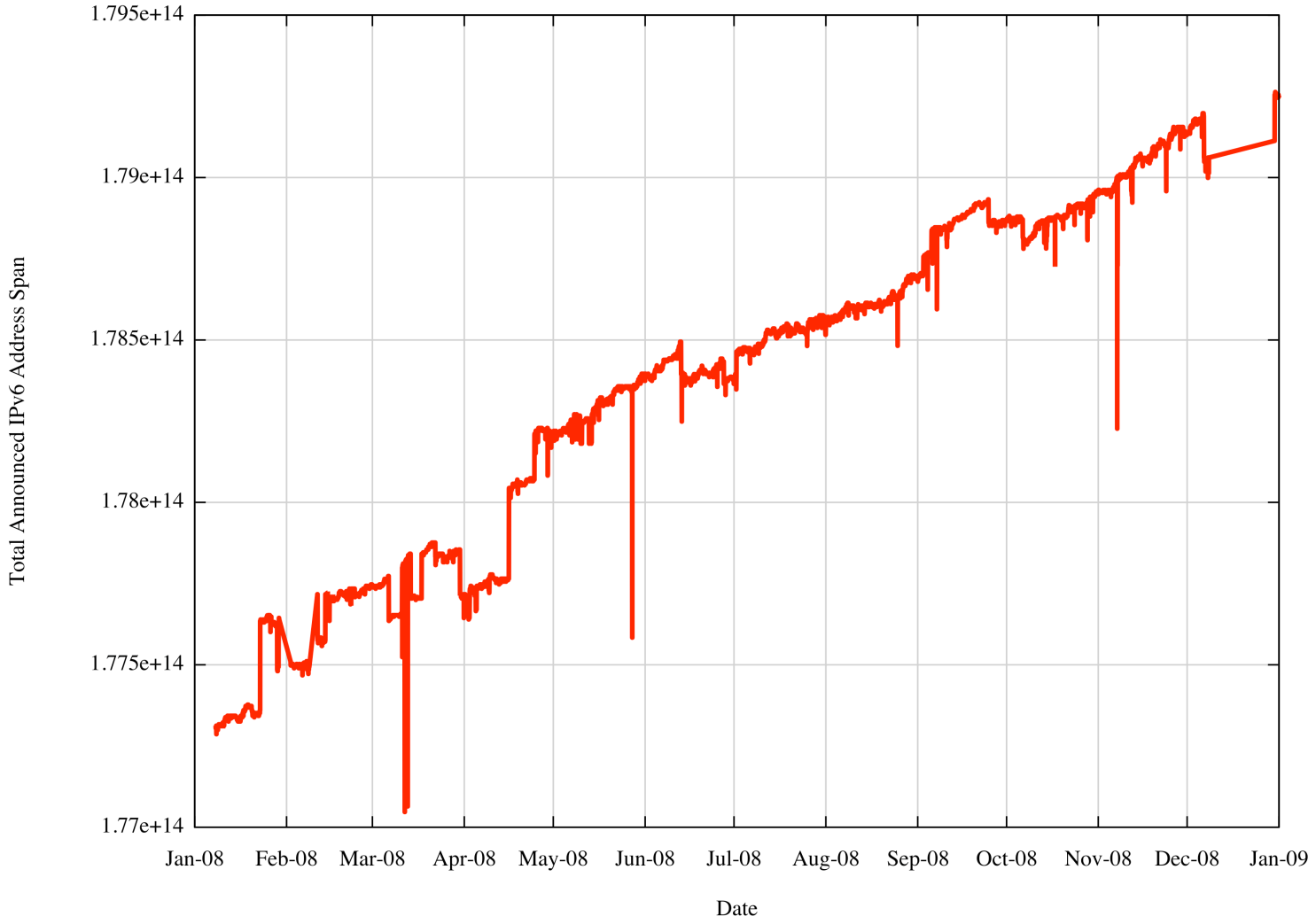
IPv6 BGP Prefix Count



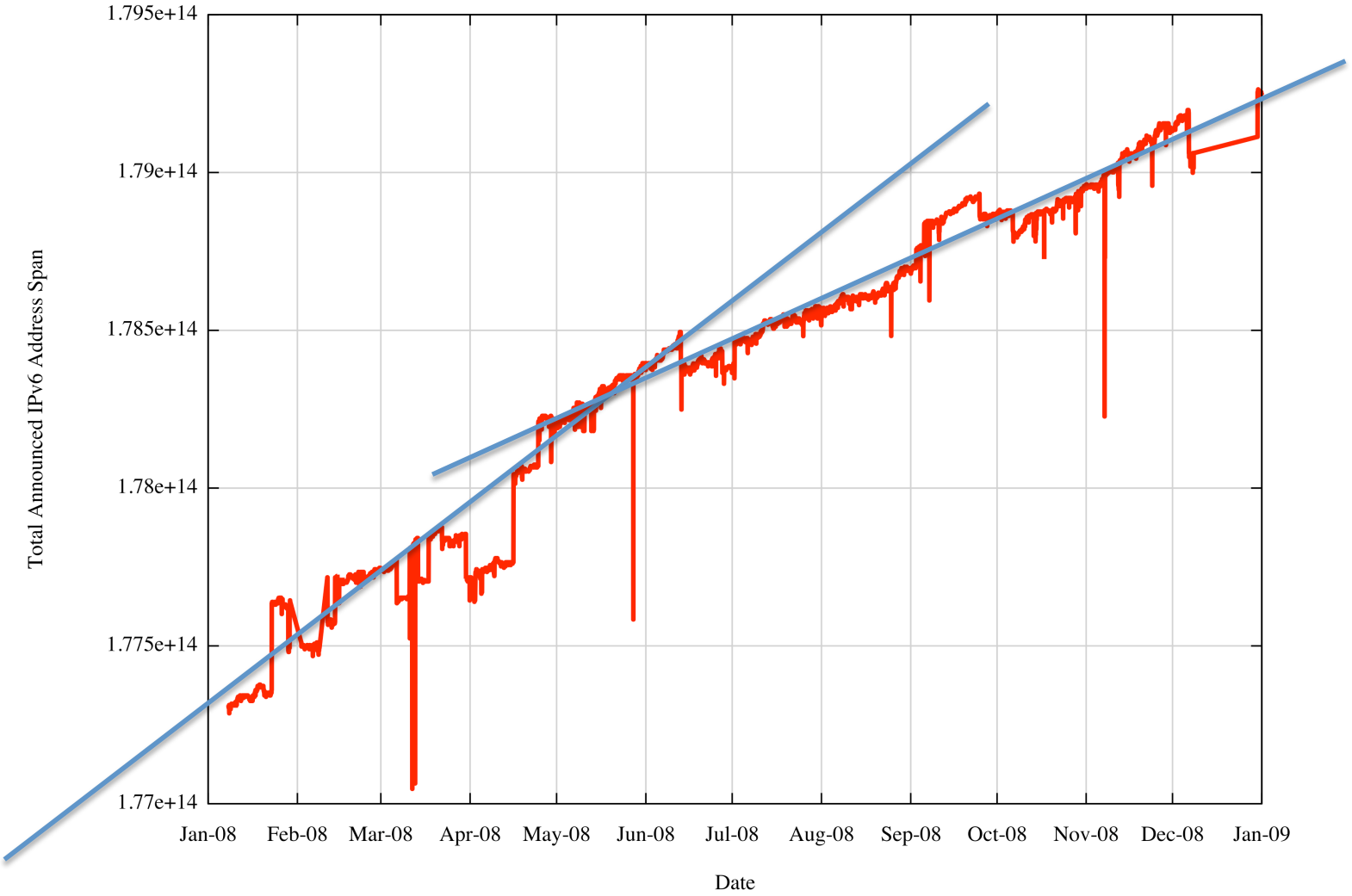
IPv6 Routed Address Span



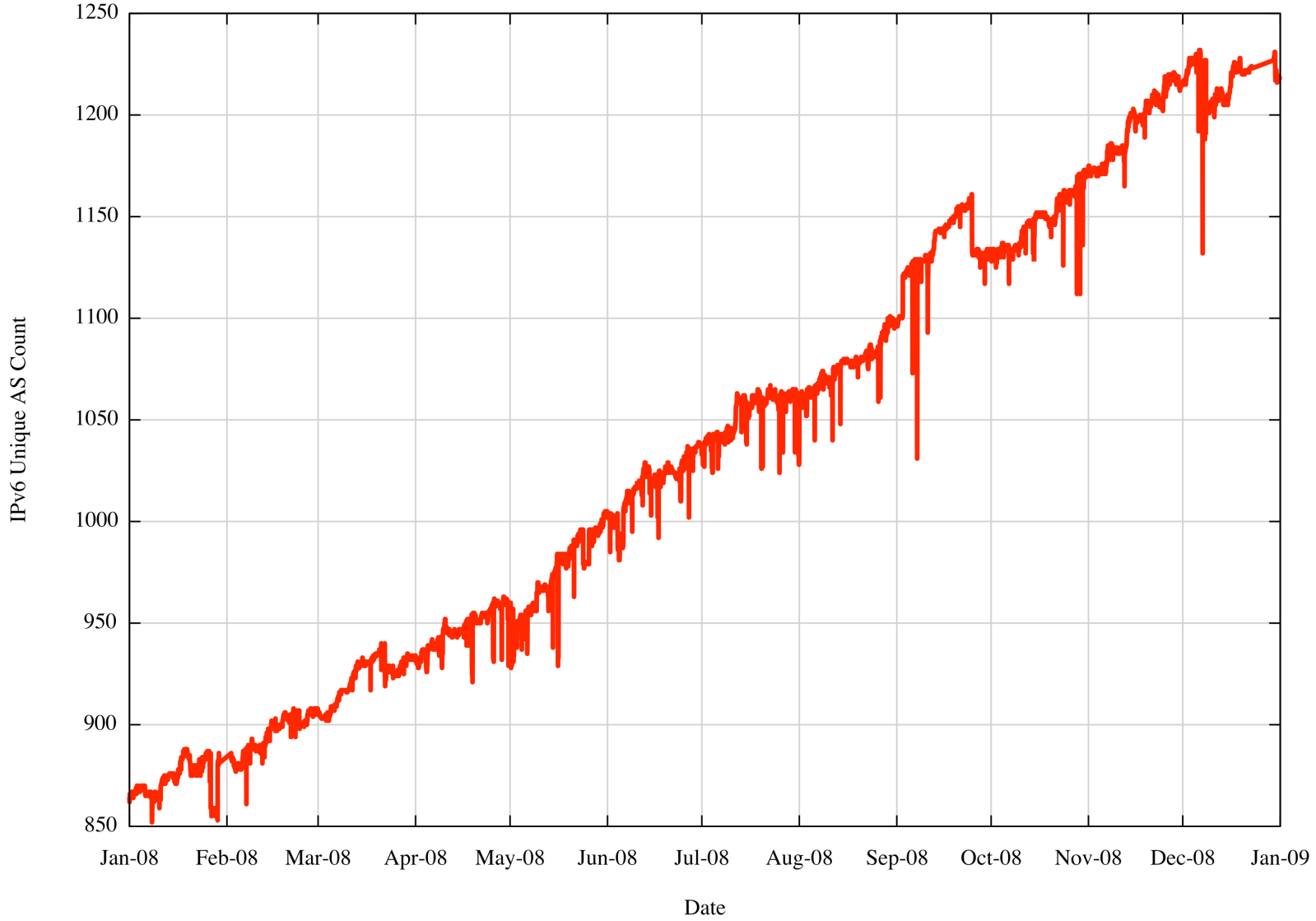
IPv6 Routed Address Span



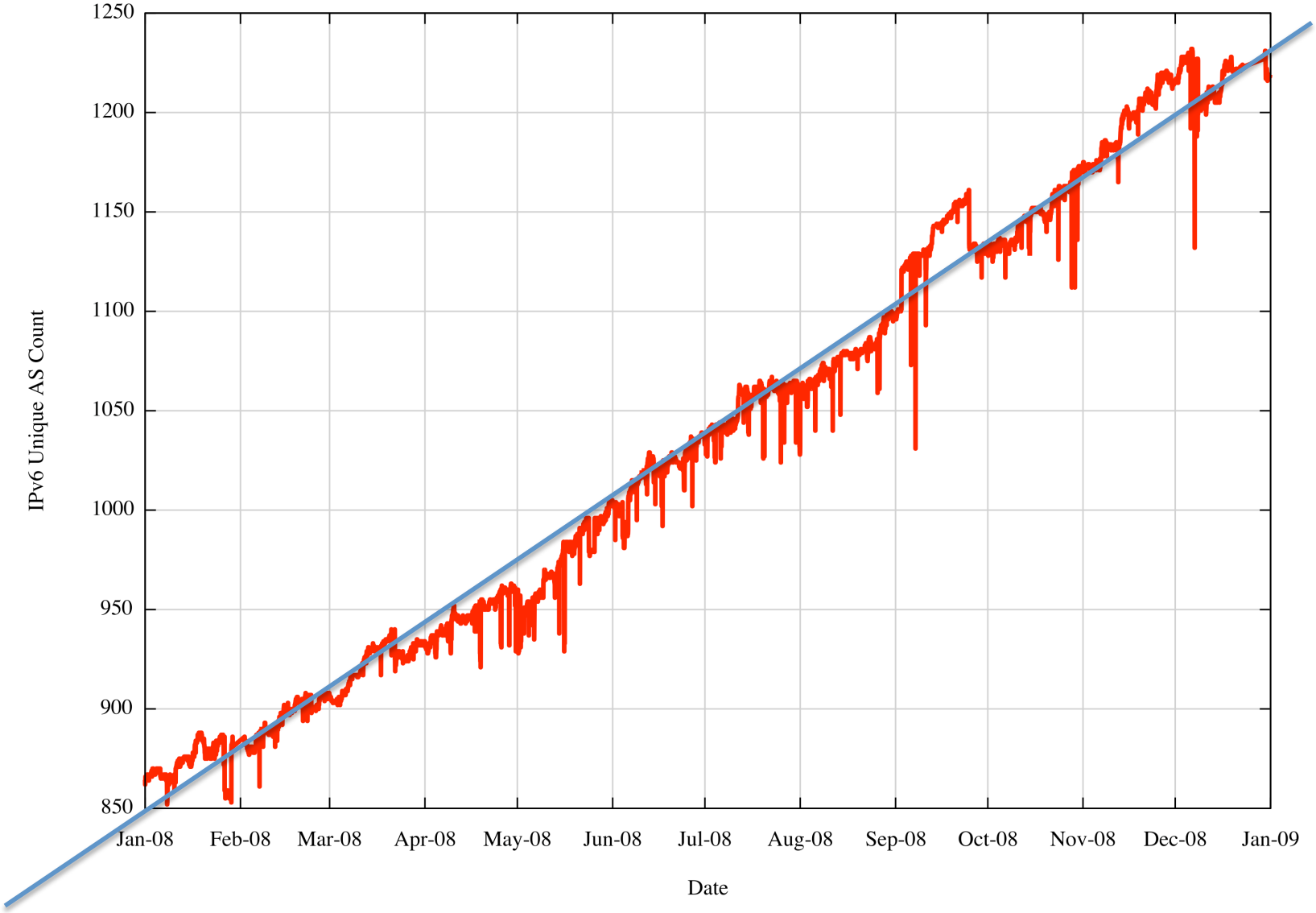
IPv6 Routed Address Span



IPv6 Routed AS Count



IPv6 Routed AS Count



IPv6 Vital Statistics for 2008

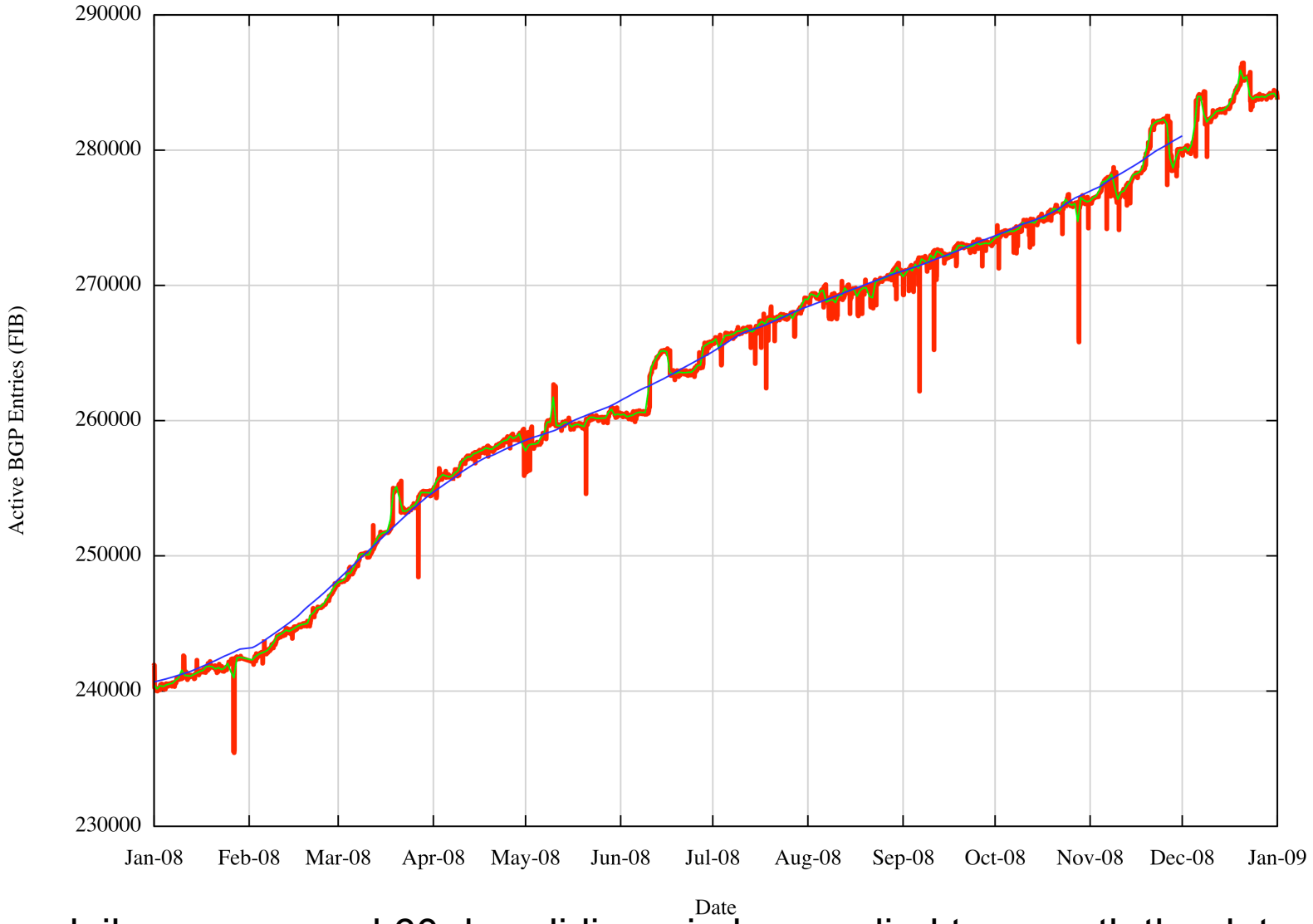
	Jan-08	Dec-08	
Prefix Count	1,050	1,600	**52%**
Roots	840	1,300	55%
More Specifics	210	300	43%
Address Span	/16.67	/16.65	**1%**
AS Count	860	1,230	**43%**
Transit	240	310	29%
Stub	620	920	48%

Trends and Projections

BGP Table Size Projection

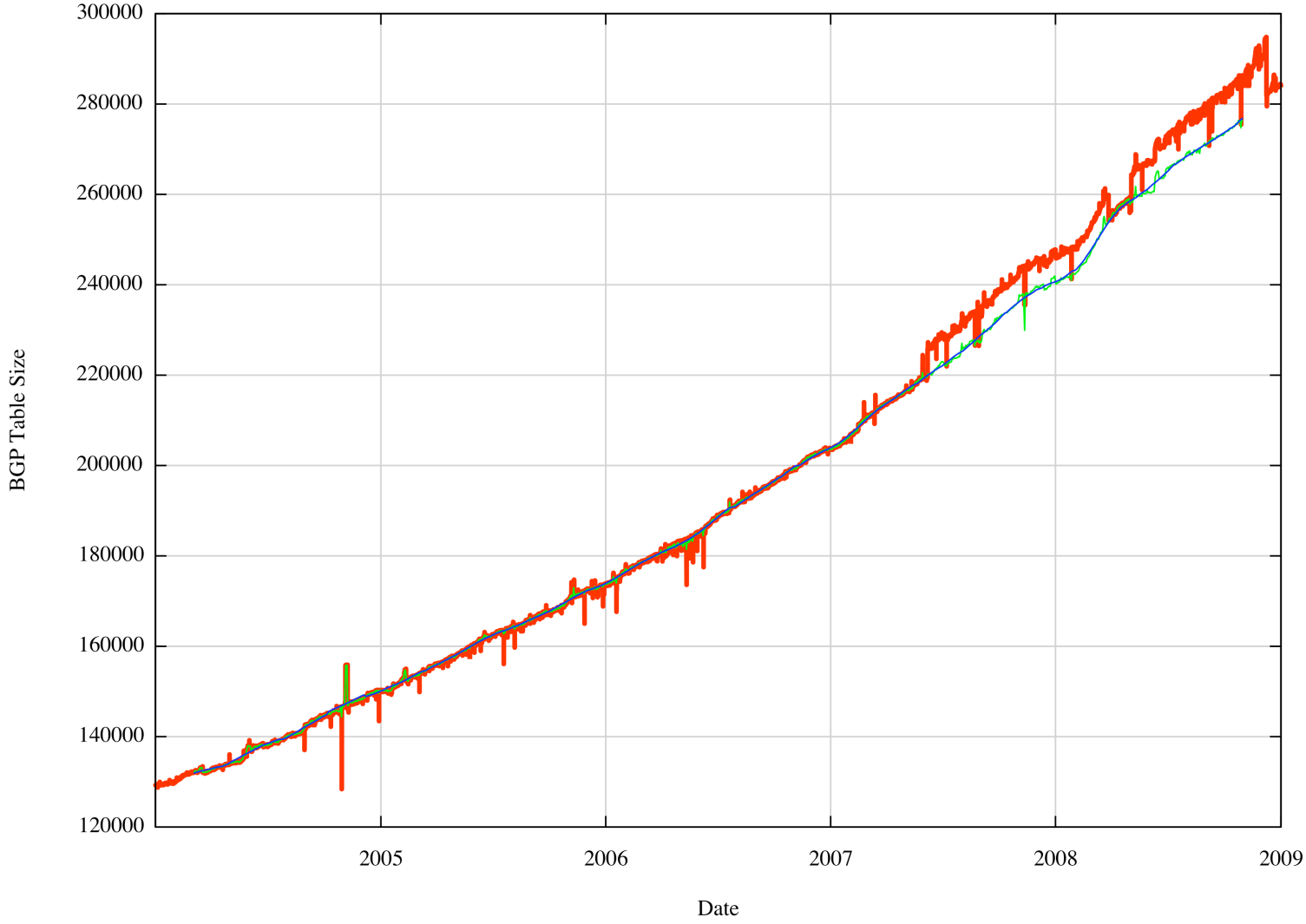
- 4 year projection of the IPv4 routing table size

IPv4 Table Size - 12 months

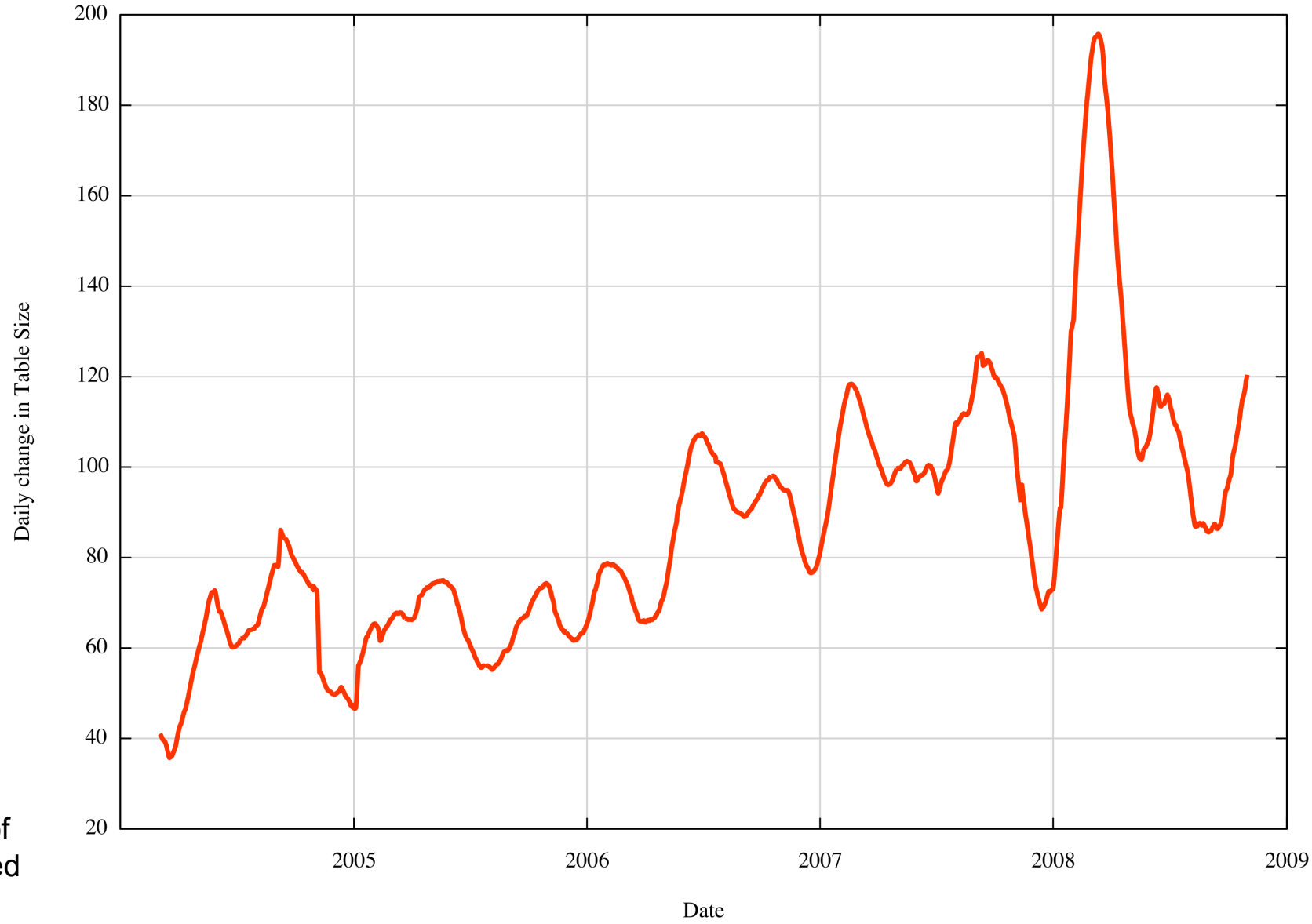


daily average and 60 day sliding window applied to smooth the data

IPv4 Table Size - 60 months

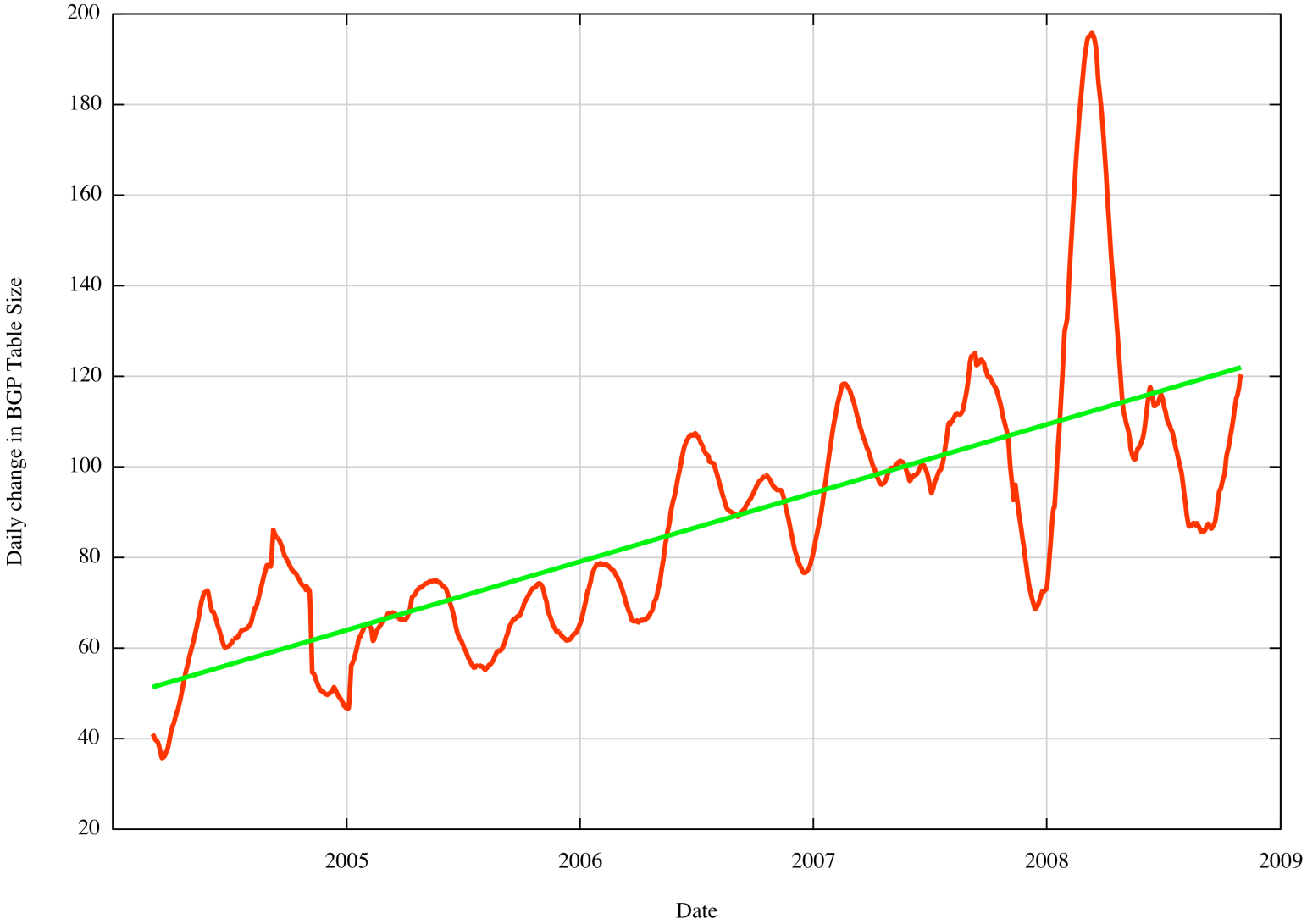


Daily Growth Rates

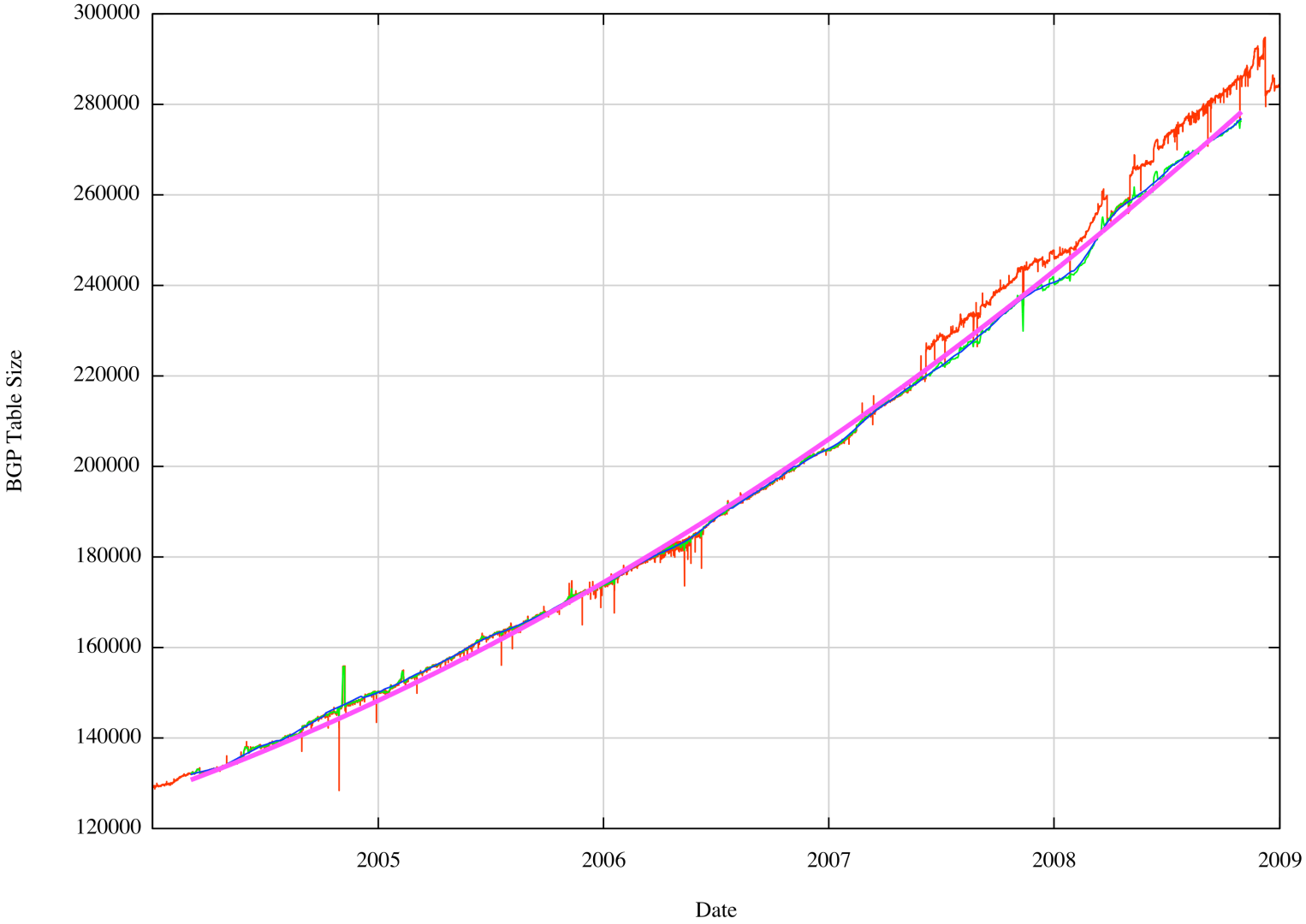


First order differential of the smoothed data

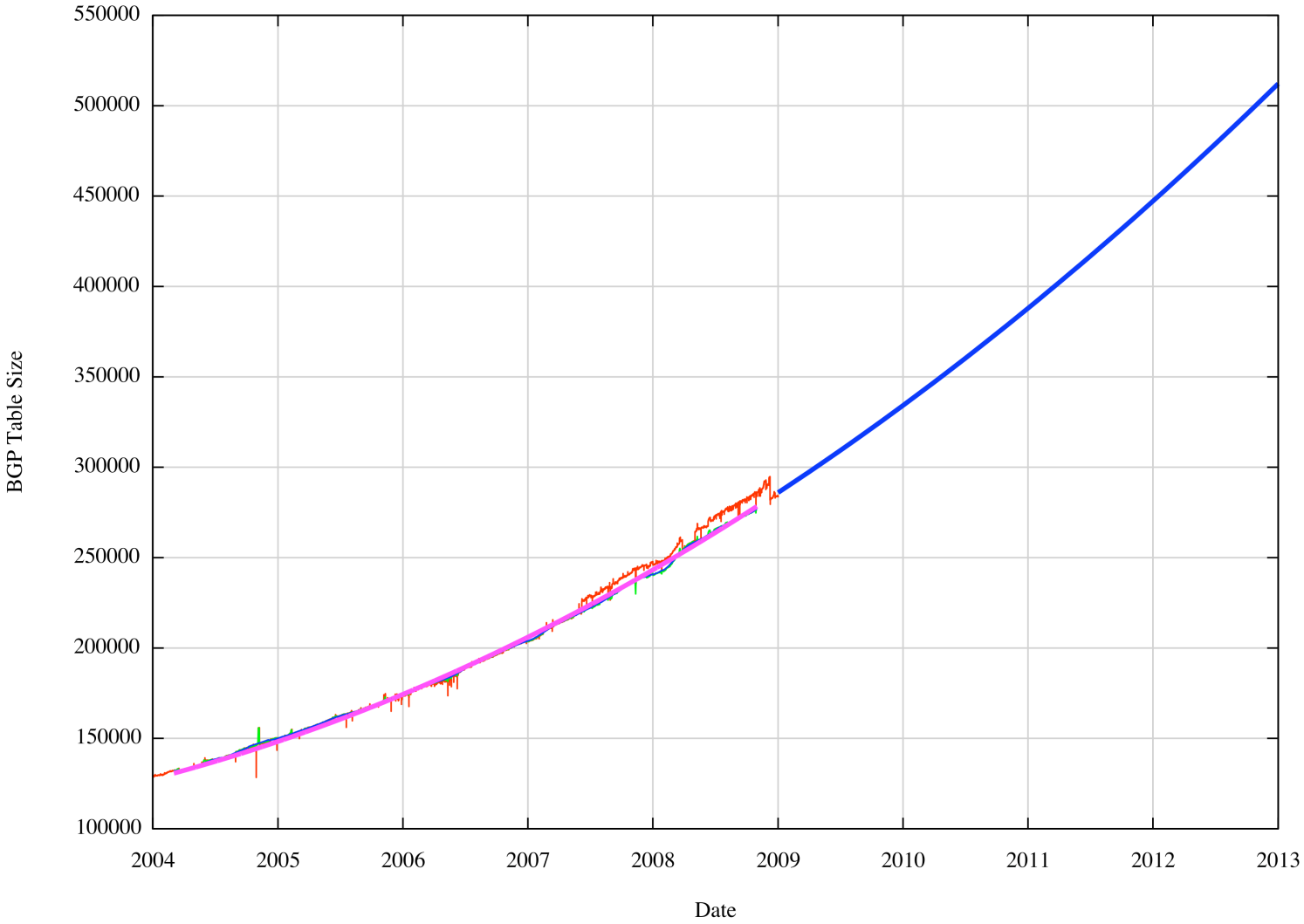
Daily Growth Rates



IPv4 Table Size
Quadratic Growth Model



IPv4 Table Size Quadratic Growth Model - Projection



BGP Table Size Predictions

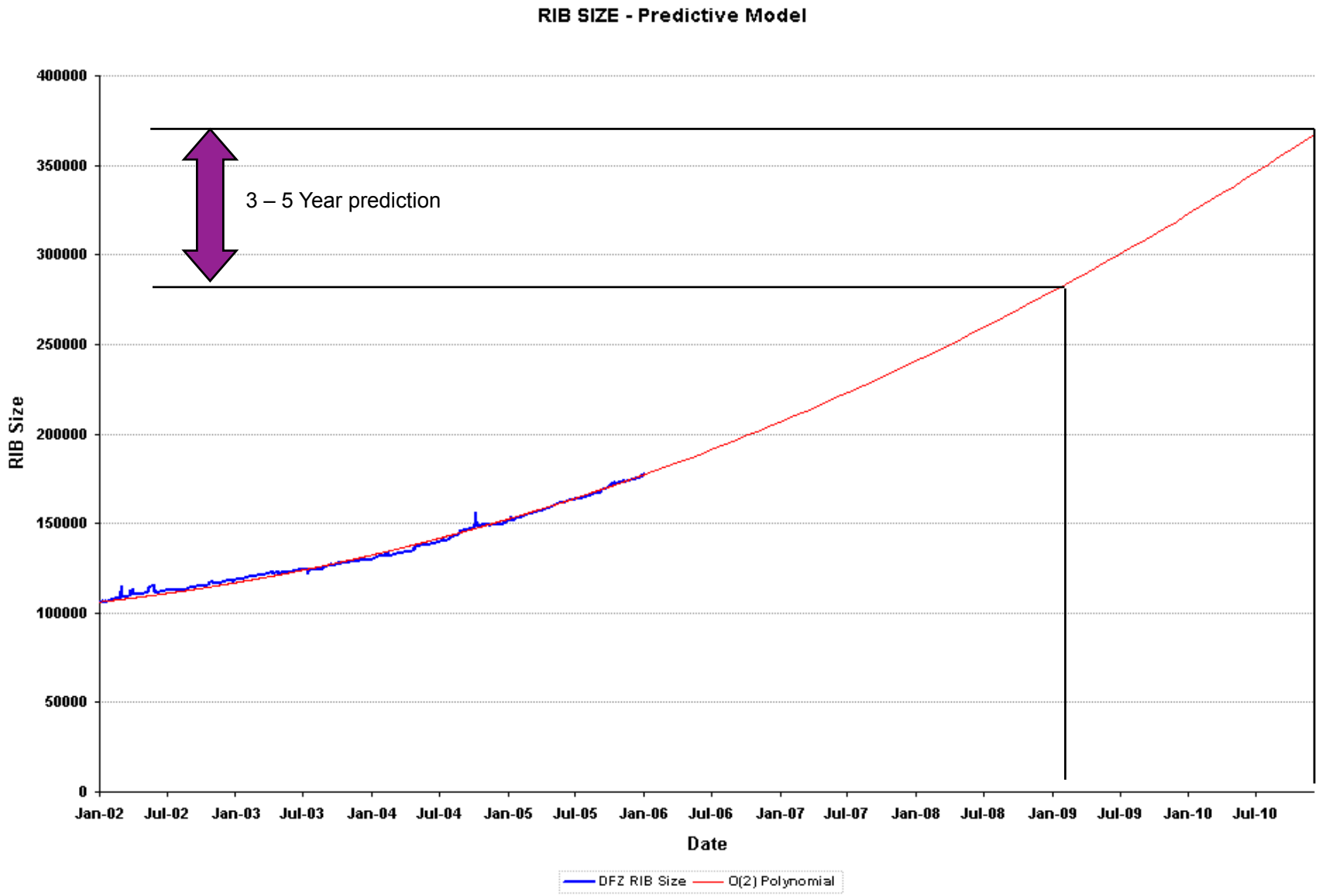
now	285,000 entries
12 months	335,000 entries
24 months	388,000 entries
36 months*	447,000 entries
48 months*	512,000 entries

* These numbers are dubious due to IPv4 address exhaustion pressures. It is possible that the number will be larger than the values predicted by this model.

Back in 2006 ….

- This modelling work of the IDR table size was performed at the end of 2005 to generate a 3 and 5 year projection

2006 prediction



BGP Table Size Predictions

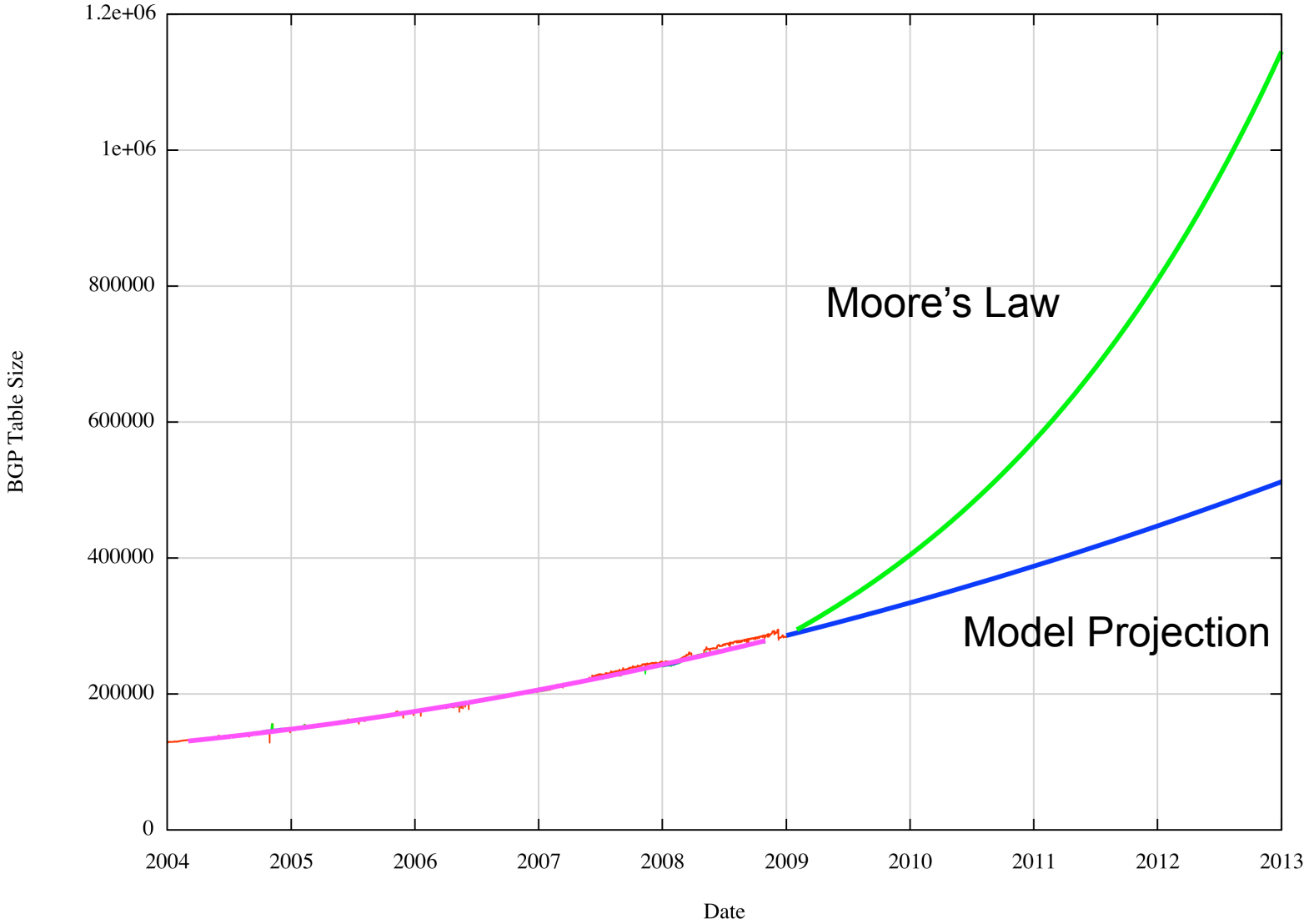
now	285,000 entries (2006: 275,000)
12 months	335,000 entries
24 months	388,000 entries (2006: 370,000)
36 months*	447,000 entries
48 months*	512,000 entries

* These numbers are dubious due to IPv4 address exhaustion pressures. It is possible that the number will be larger than the values predicted by this model.

Moore's Law

- As long as growth rates stay within the general parameters of Moore's Law the unit cost of the routing function should not escalate
  - use a function of doubling every two years as the application of Moore's Law
  - assuming that Moore's law continues to hold

Projections against Moore's Law



BGP Table Size

- 3 years = 500,000 IPv4 entries + ? IPv6 entries
  - plan for probably no less than 600,000 entries
- 5 years = ? IPv4 entries + ? IPv6 entries
  - plan for probably no less than 800,000 entries

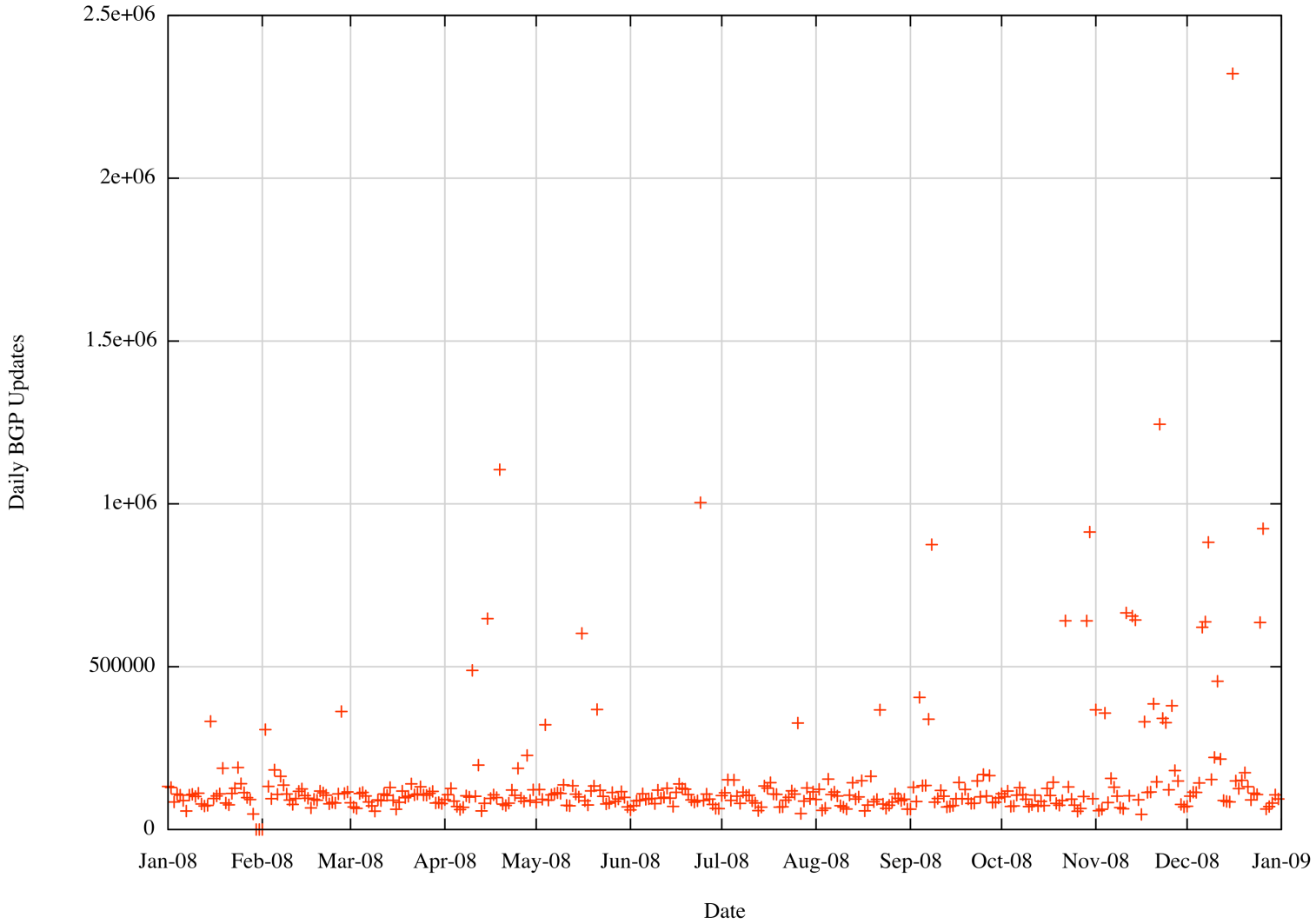
BGP Metrics

- Is it the size of the RIB or the level of dynamic update that is the concern here?
- Lets look at update trends in BGP…

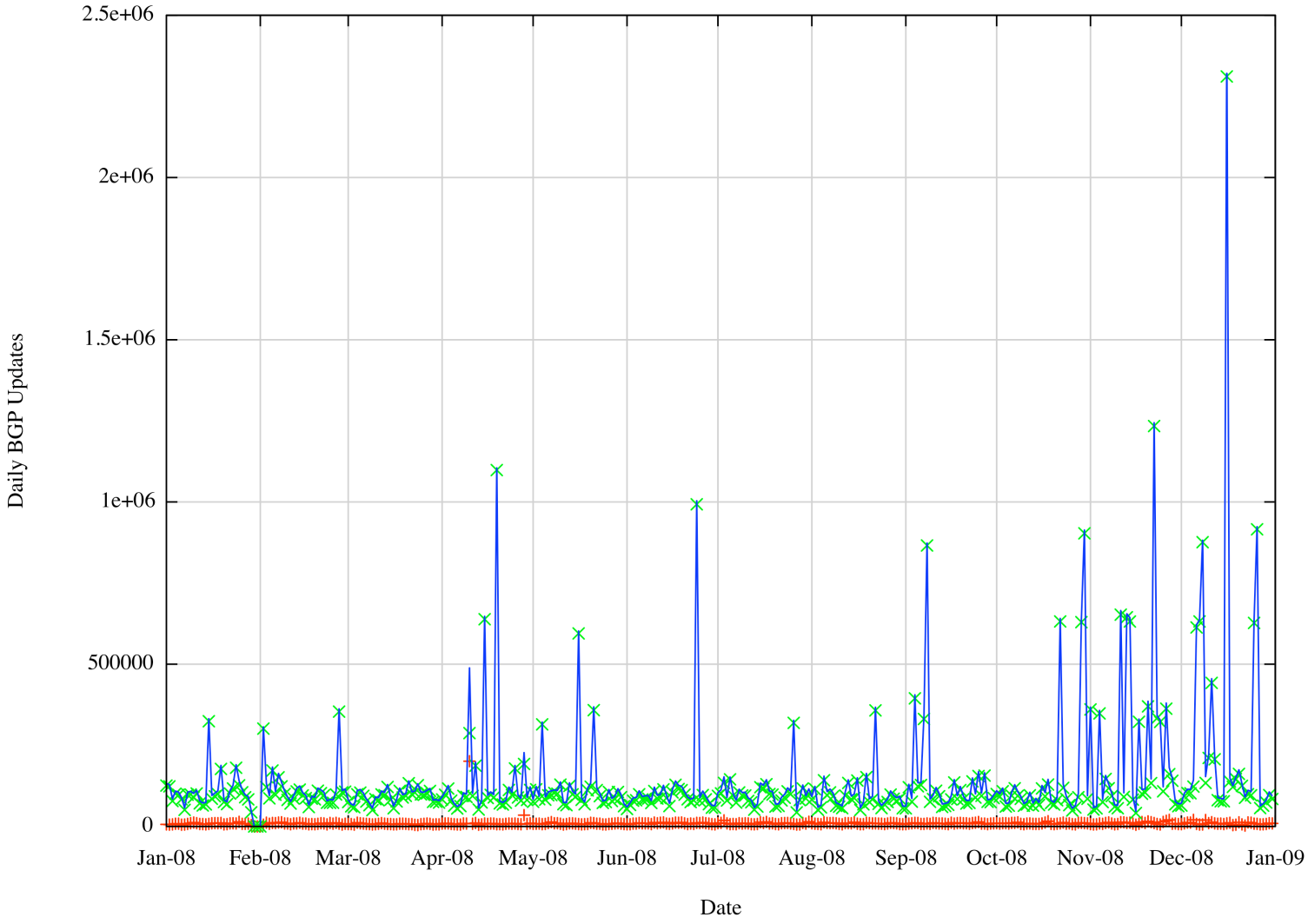
BGP Update Behaviour

- What trends are visible in the number of BGP updates?

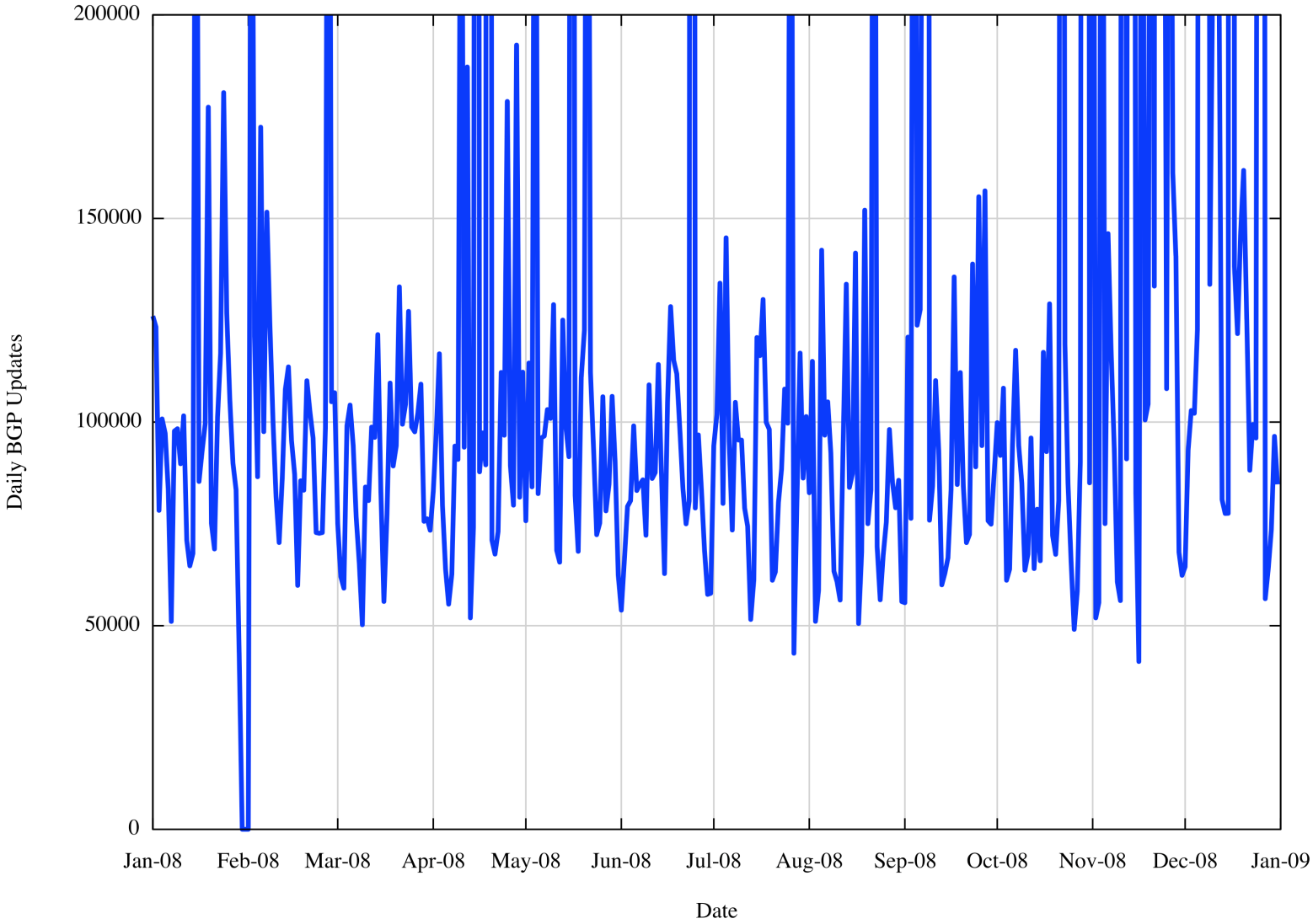
IPv4 BGP Update Behaviour



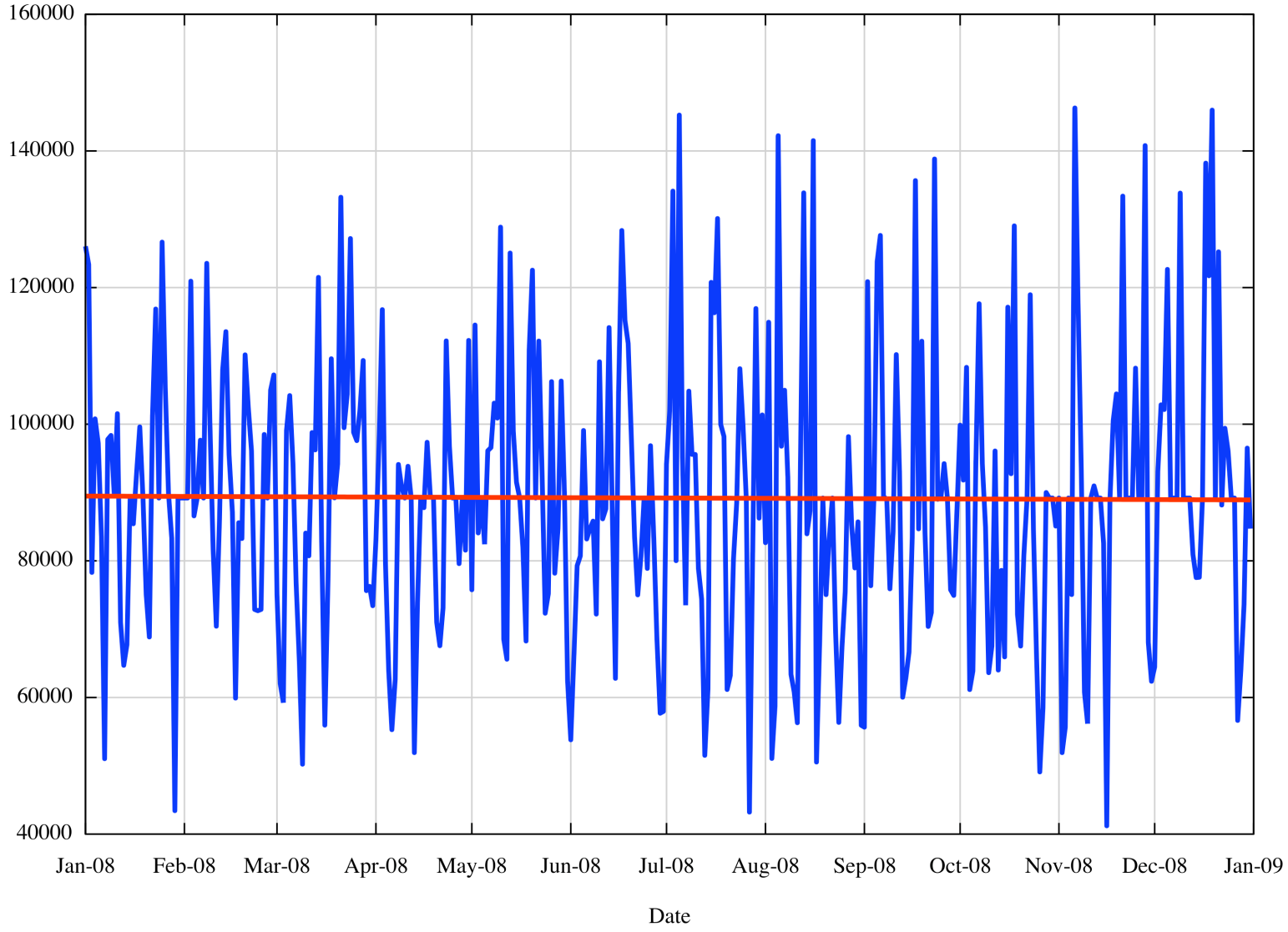
BGP Upates - Announce and Withdrawals



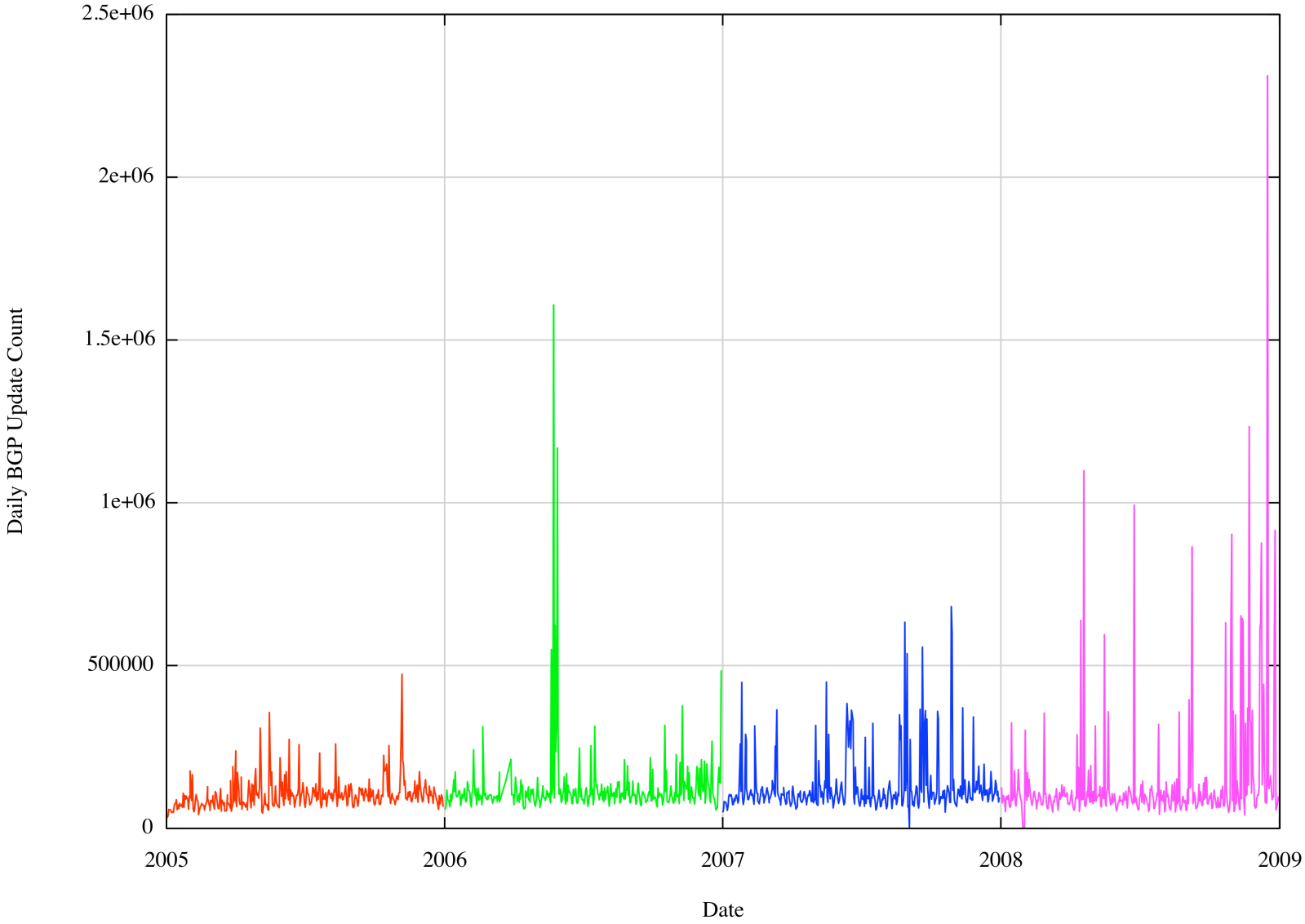
Daily Updates



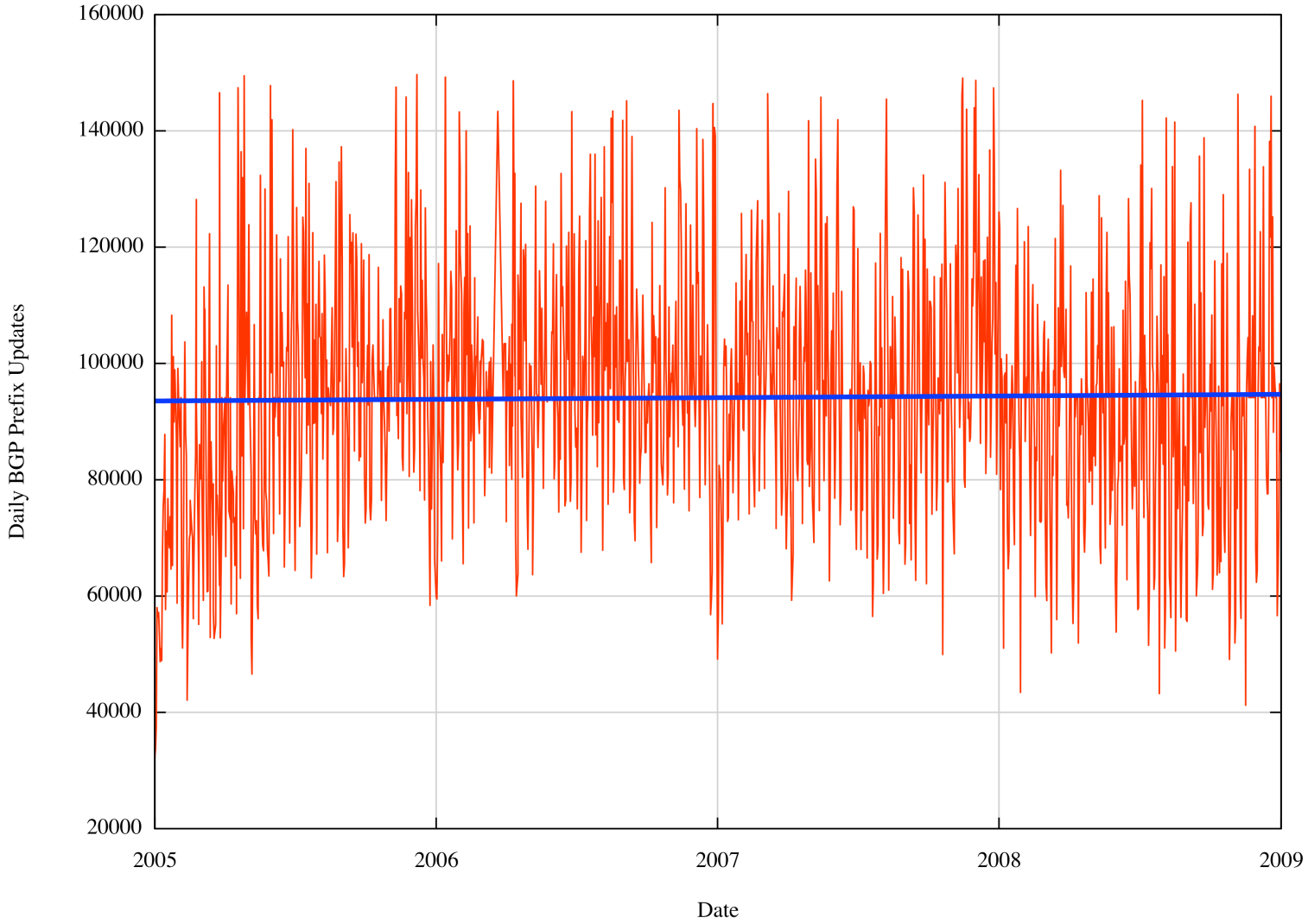
Best Fit to Updates



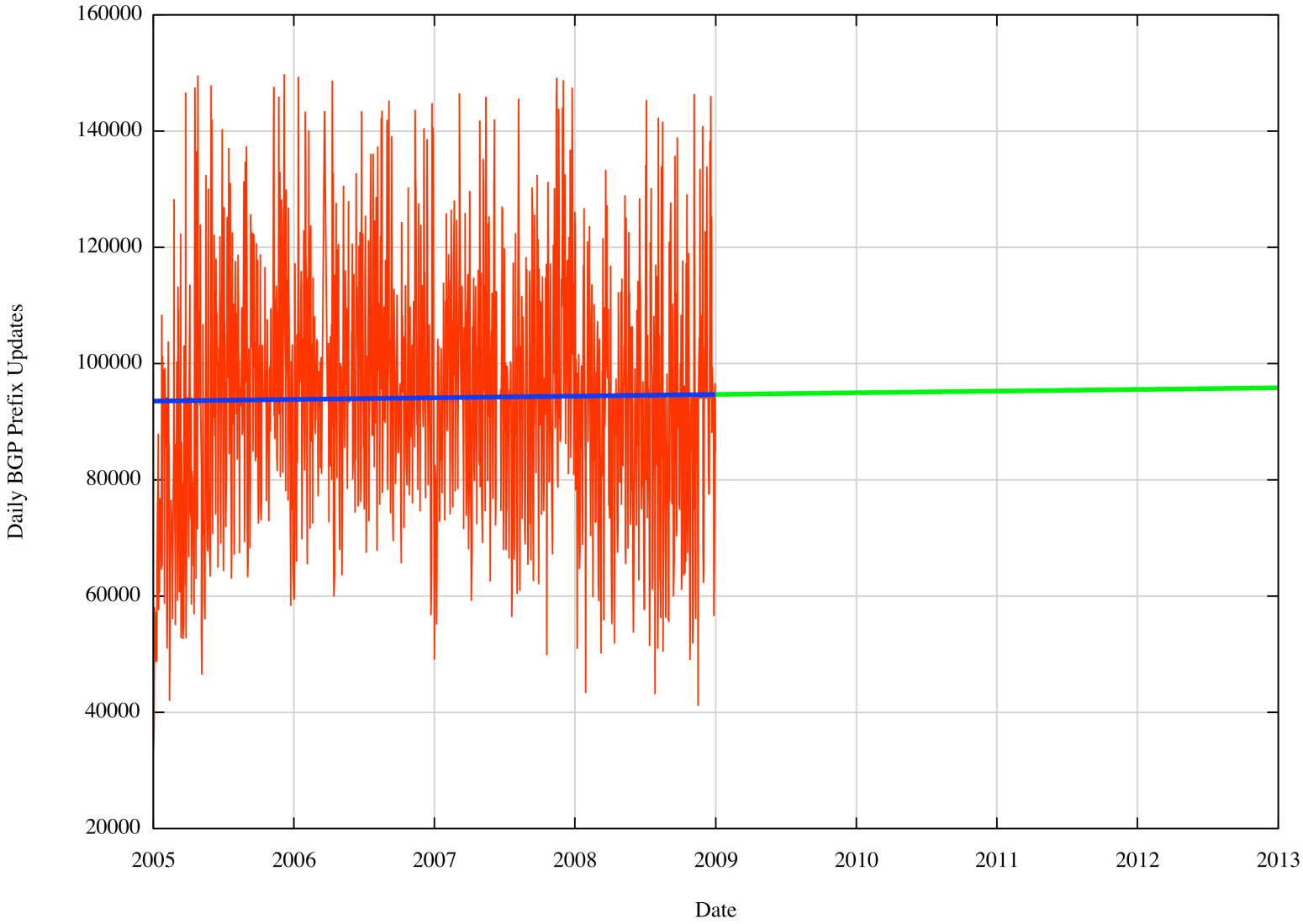
BGP Updates - Extended Data Set



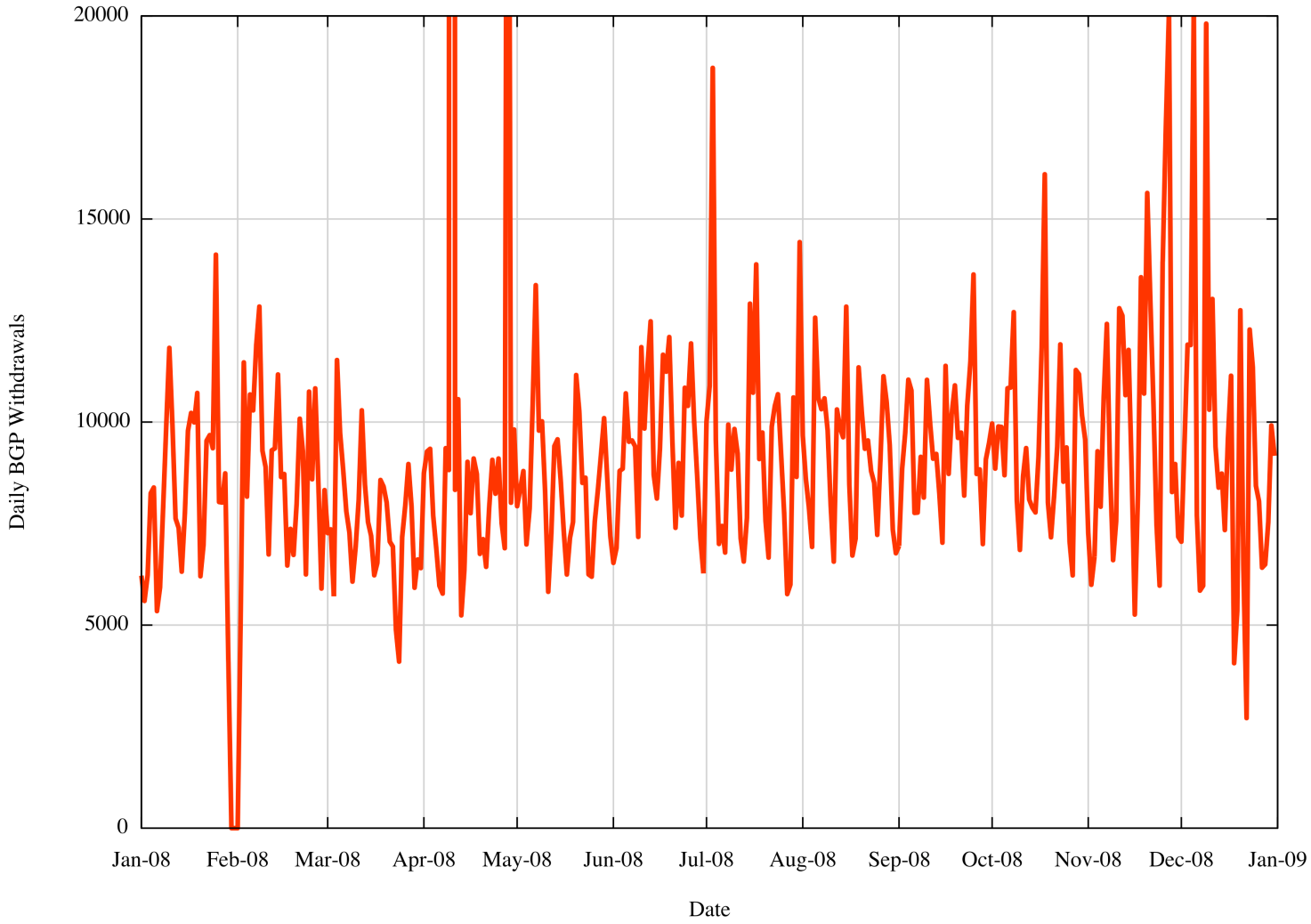
BGP Updates - Extended Data Set



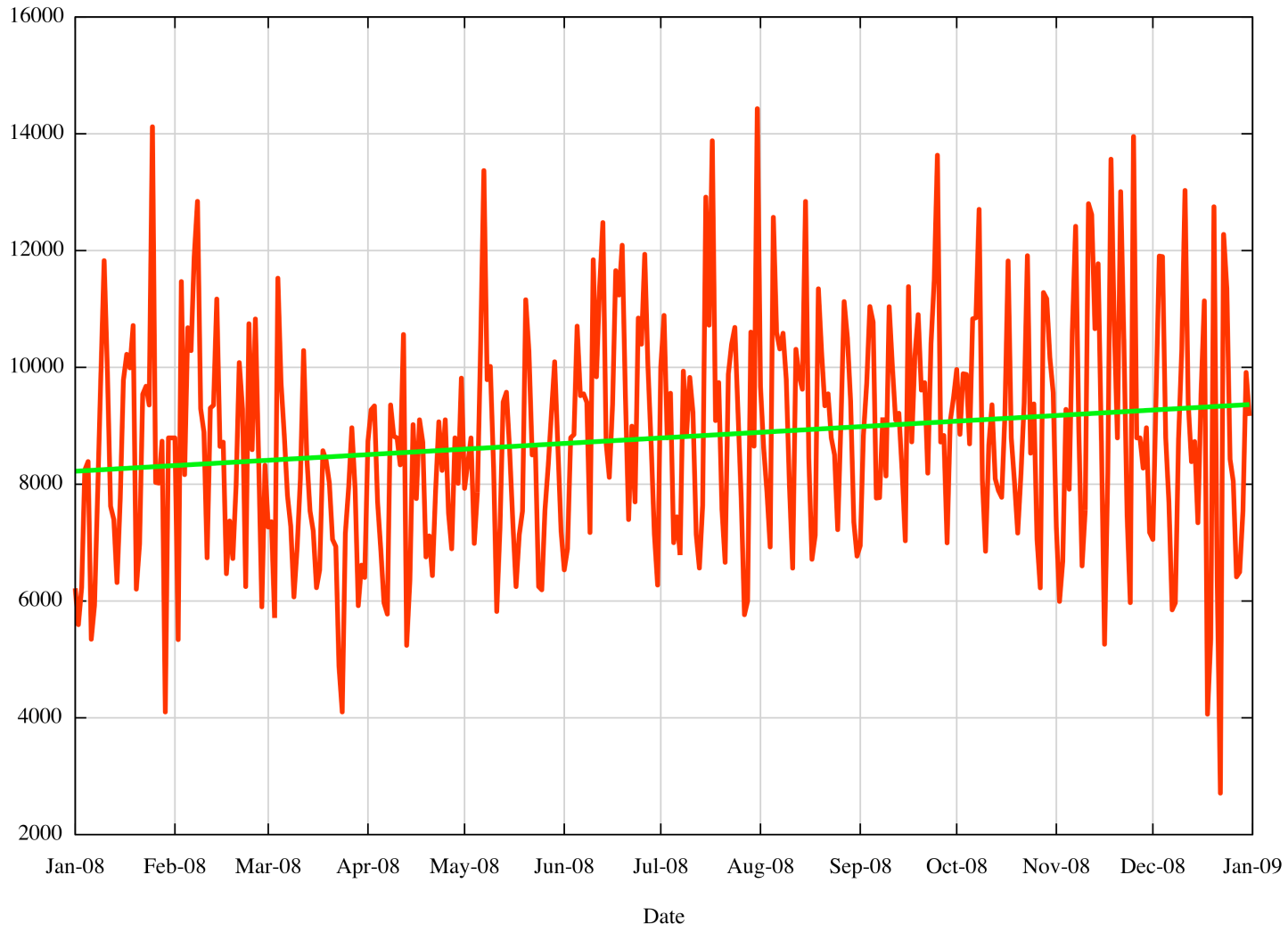
BGP Update Projection



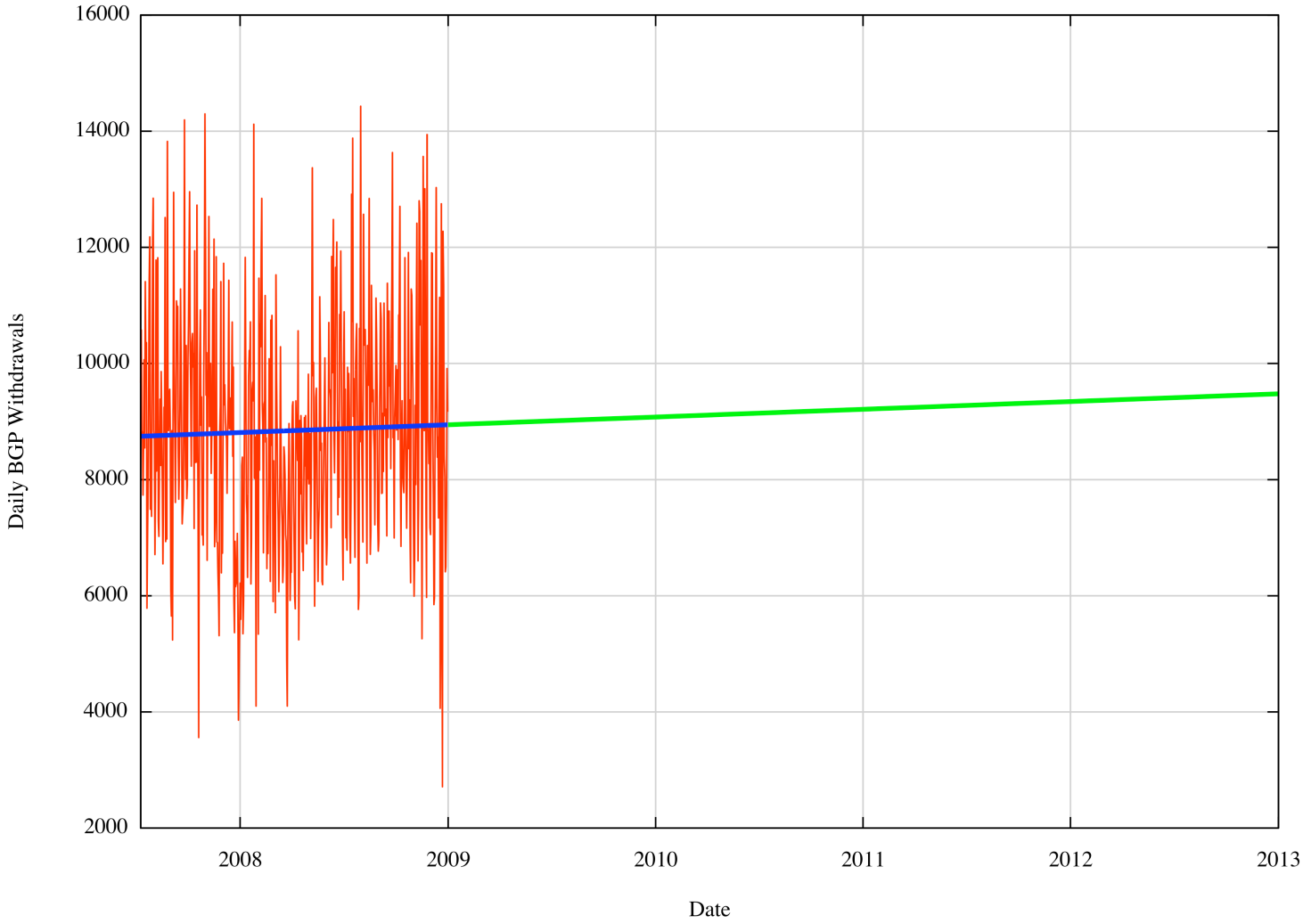
Daily Withdrawals



Best Fit to WDLs



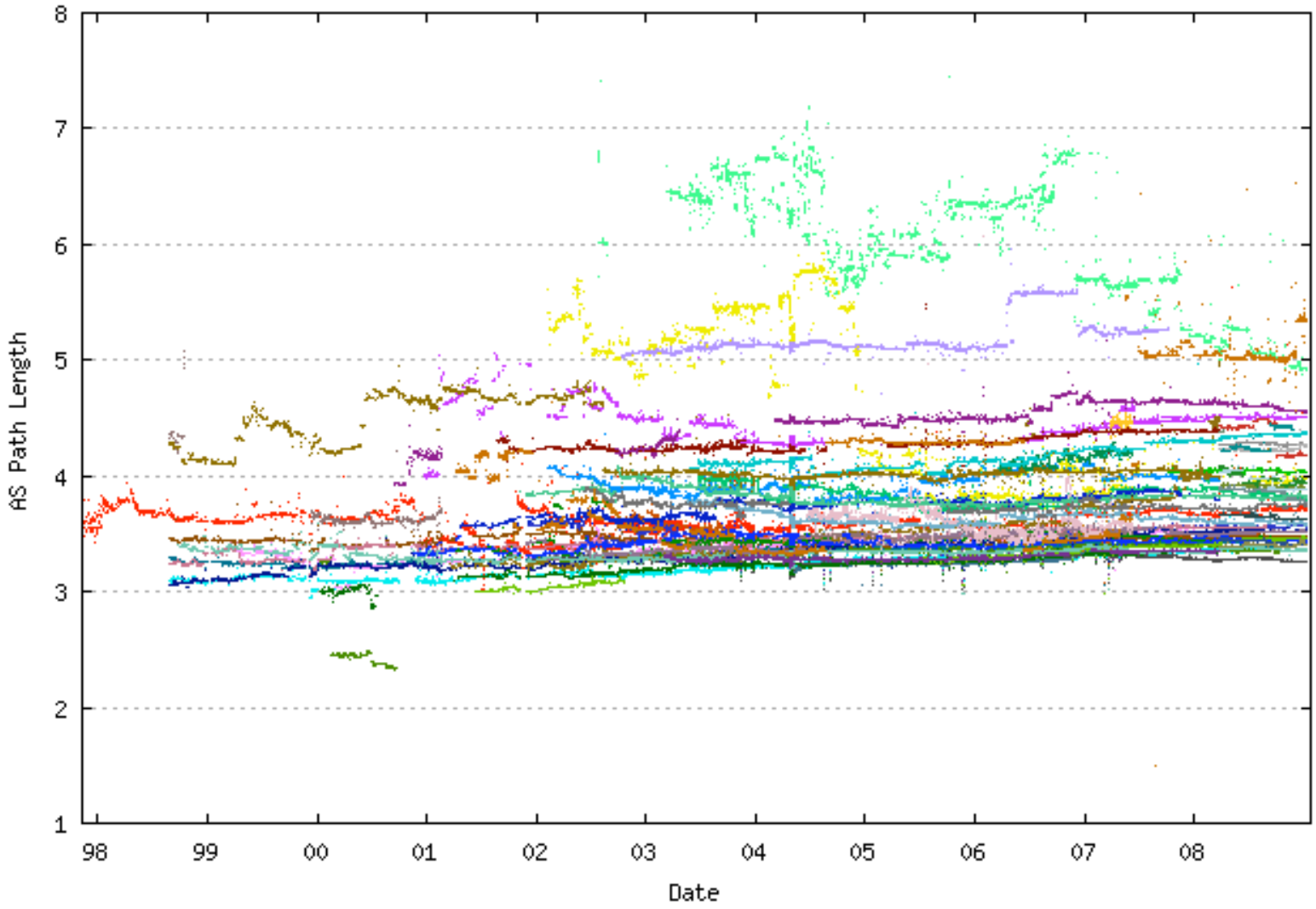
BGP Withdrawal Projection



Why is this?

- Why are the levels of growth in BGP updates not proportional to the size of the routing table?
  - growth rates of BGP updates appear to be far smaller than the growth rate of the routing space itself
- What is going on?

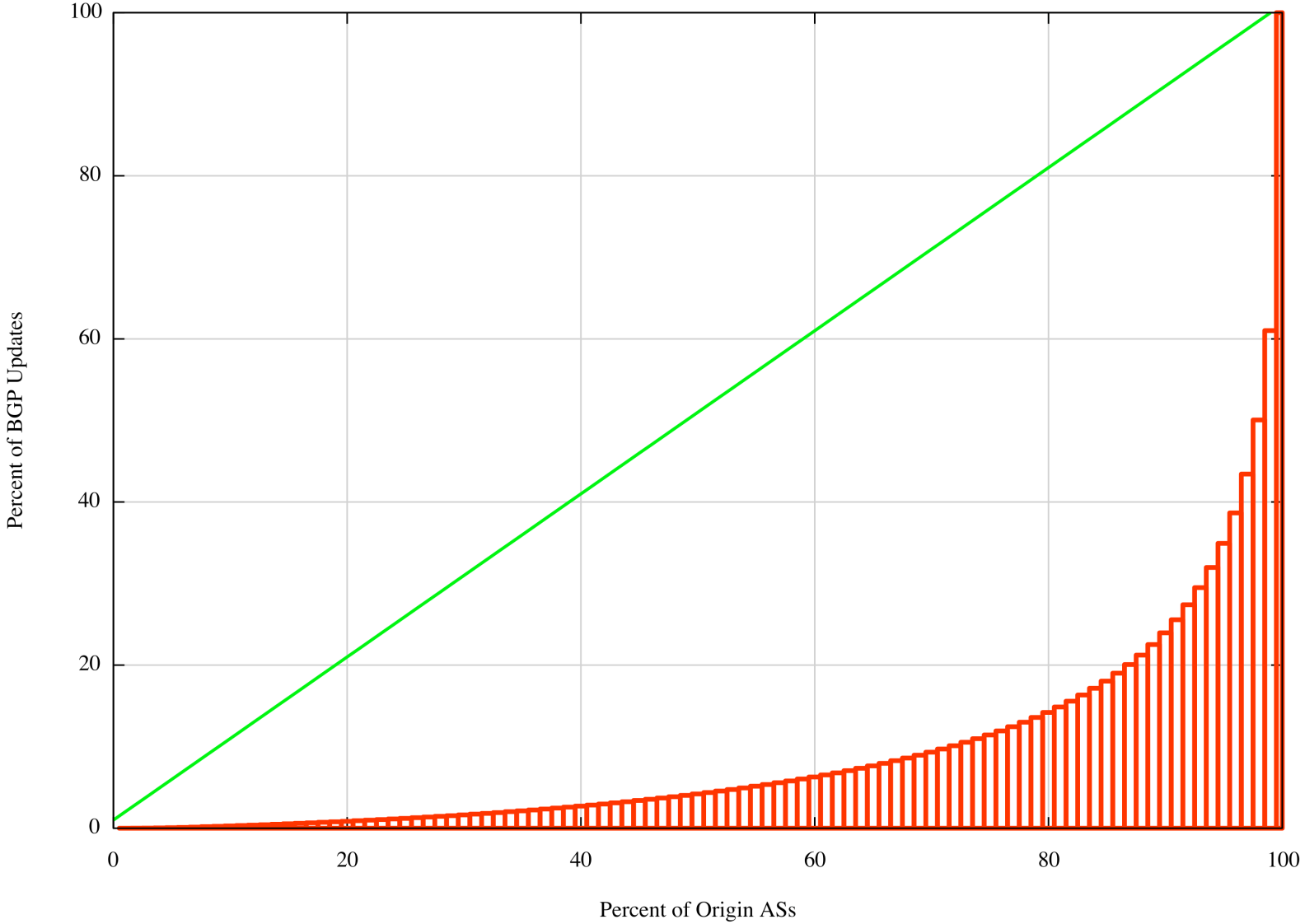
Average AS Path Length has remained constant



What is going on?

- The convergence instability factor for a distance vector protocol like BGP is related to the AS path length, and average AS Path length has remained steady in the Internet for some years
- Taking MRAI factors into account, the number of received Path Exploration Updates in advance of a withdrawal is related to the propagation time of the withdrawal message. This is approximately related to the average AS path length
- Today's Internet is more densely interconnected, but is not any more "stringier"
- This implies that the number of protocol path exploration transitions leading to a prefic withdrawal is relatively stable over time

The update distribution of BGP is heavily skewed



What is going on?

- The major component of dynamic BGP load is not an artifact of the larger routing space, but a case of relatively intense levels of BGP path manipulation at or close to origin for TE purposes from a very small subset of origin AS's
  - the dominant factor behind what is being measured in updates is not implicitly related to network component stability, but more likely to be related to path manipulation associated with TE

A prudent view of capacity planning for 4 - 5 years into BGP's future

- A FIB size of 750,000 entries of IPv4, and a total of 1M entries of IPv4 + IPv6 FIB space
- A BGP update load of 2x the current rates
- No major changes in the BGP use profile, no major changes in the distribution of BGP updates

Some Closing Opinions

- The BGP sky is **not** falling

  The 2008 BGP data appears to indicate that the prospects of the imminent death of BGP through routing table inflation appear to be vastly exaggerated

  - The inflation rate of the routing table remains well under Moore's law
  - The rate of increase of processed updates in the gather data appears to be minimal
  - The rate of increase of withdrawals is slightly larger, but is still slight

  – The network is, on the whole, very stable and BGP is not under immediate stress in terms of scaling pressures

Some Closing Opinions

- This does not model post-V4 address exhaustion scenarios
  - But there is no evidence to be alarmed about excessive fragmentation of the routing space at this stage

A Word of Caution

- This is a simple exercise in statistical curve fits, not a demand level simulation of the players routing environment.
- This exercise does not factor in any IPv4 address exhaustion considerations and scenarios around address movement that may alter the picture of fragmentation of the routing space.
- However the AS growth projections are a strong indicator of underlying industry dynamics in terms of discrete routing entities, and these projections show a modest growth component

  This means that while the projections are very weak in the period of 2011 and beyond, there are grounds to take a conservative view of BGP growth in this phase of the Internet's evolution

Thank You

Questions?

Additional Material

Top 10 AS Profile

- The 10 most active Origin ASs are associated with 3,069,860 prefix updates, or 5.53% of the total number of prefix updates for the year.
- Who are they and what are they doing?

Top 10 AS Profile

Rank	AS	Updates	Percent of Total
1	6389	418,945	0.75%
2	8159	373,164	0.67%
3	17974	348,835	0.63%
4	17488	338,762	0.61%
5	4621	325,462	0.59%
6	9835	297,464	0.54%
7	9498	266,930	0.48%
8	4323	248,126	0.45%
9	9829	226,666	0.41%
10	9583	225,506	0.41%
TOTAL		3,069,860	5.53%

#1 AS6389

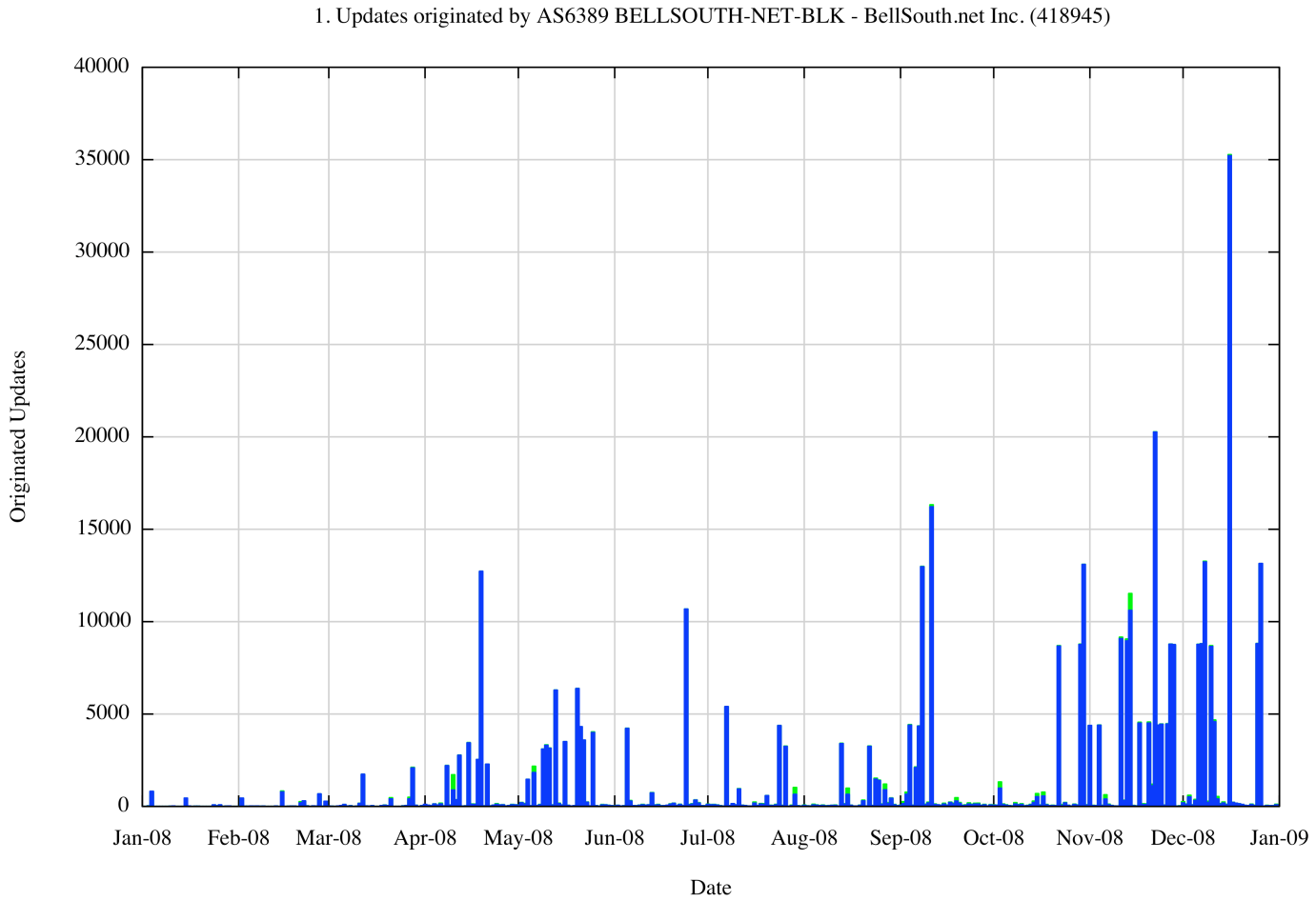
BELLSOUTH-NET-BLK - BellSouth.net Inc.AS (US)

AS Adjacency: Up: 1 Down: 44

Originated Prefixes: 4380

Aggregateable : 4096 (!)

#1 AS6389


1. Updates originated by AS6389 BELLSOUTH-NET-BLK - BellSouth.net Inc. (418945)
Originated Updates
40000
35000
30000
25000
20000
15000
10000
5000
0
Jan-08
Feb-08
Mar-08
Apr-08
May-08
Jun-08
Jul-08
Aug-08
Sep-08
Oct-08
Nov-08
Dec-08
Jan-09
Date

#2 AS8151

Uninet S.A. de C.V (MX)

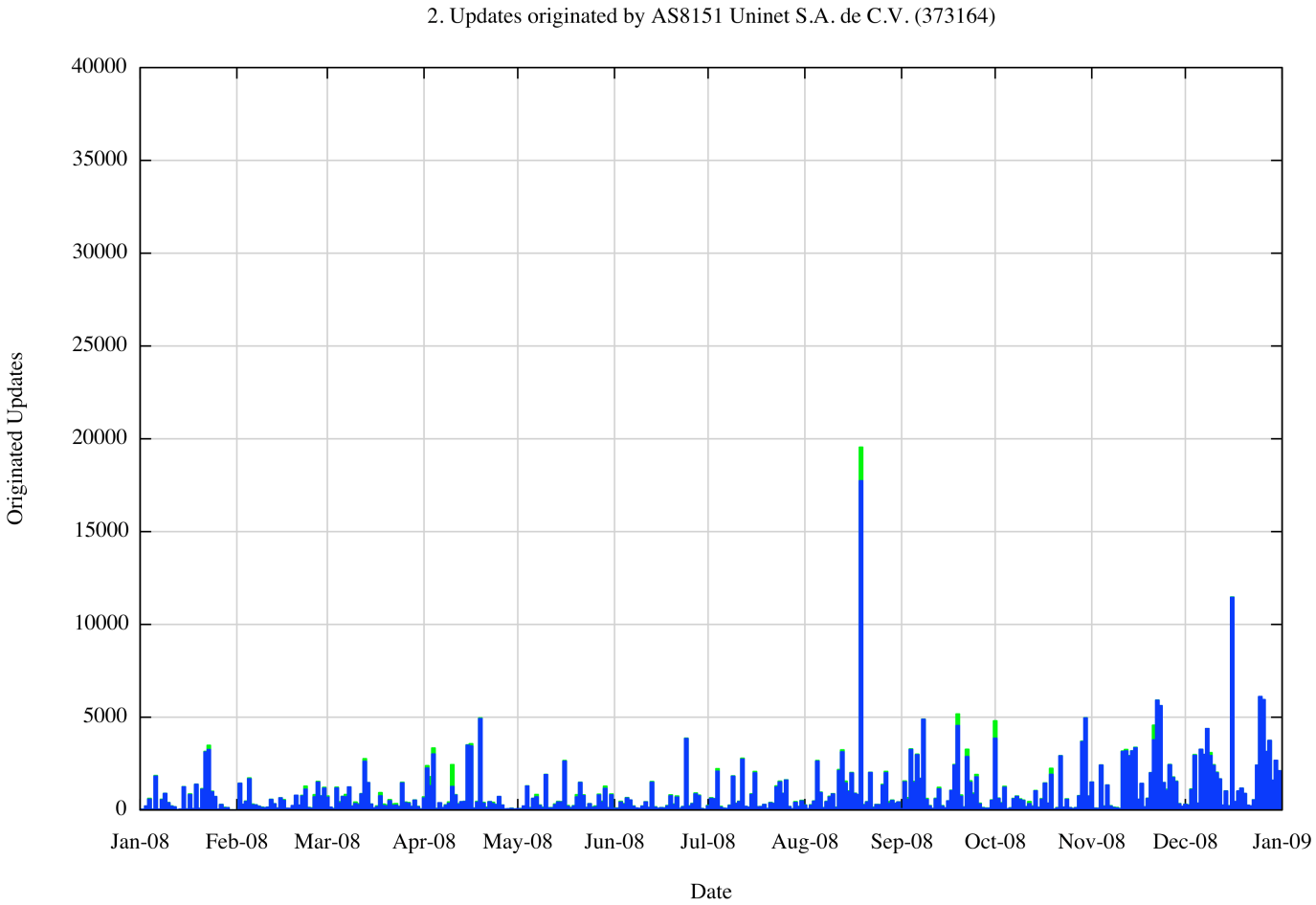
AS Adjacency: Up: 7 Down: 44

UP: 1239, 2914, 701, 7018, 18592, 3356, 3549

Originated Prefixes: 1480

Aggregateable : 1073

#2 AS8151


2. Updates originated by AS8151 Uninet S.A. de C.V. (373164)
Originated Updates
40000
35000
30000
25000
20000
15000
10000
5000
0
Jan-08
Feb-08
Mar-08
Apr-08
May-08
Jun-08
Jul-08
Aug-08
Sep-08
Oct-08
Nov-08
Dec-08
Jan-09
Date

#3 AS17974

TELKOMNET-AS2-AP PT Telekomunikasi Indonesia (ID)

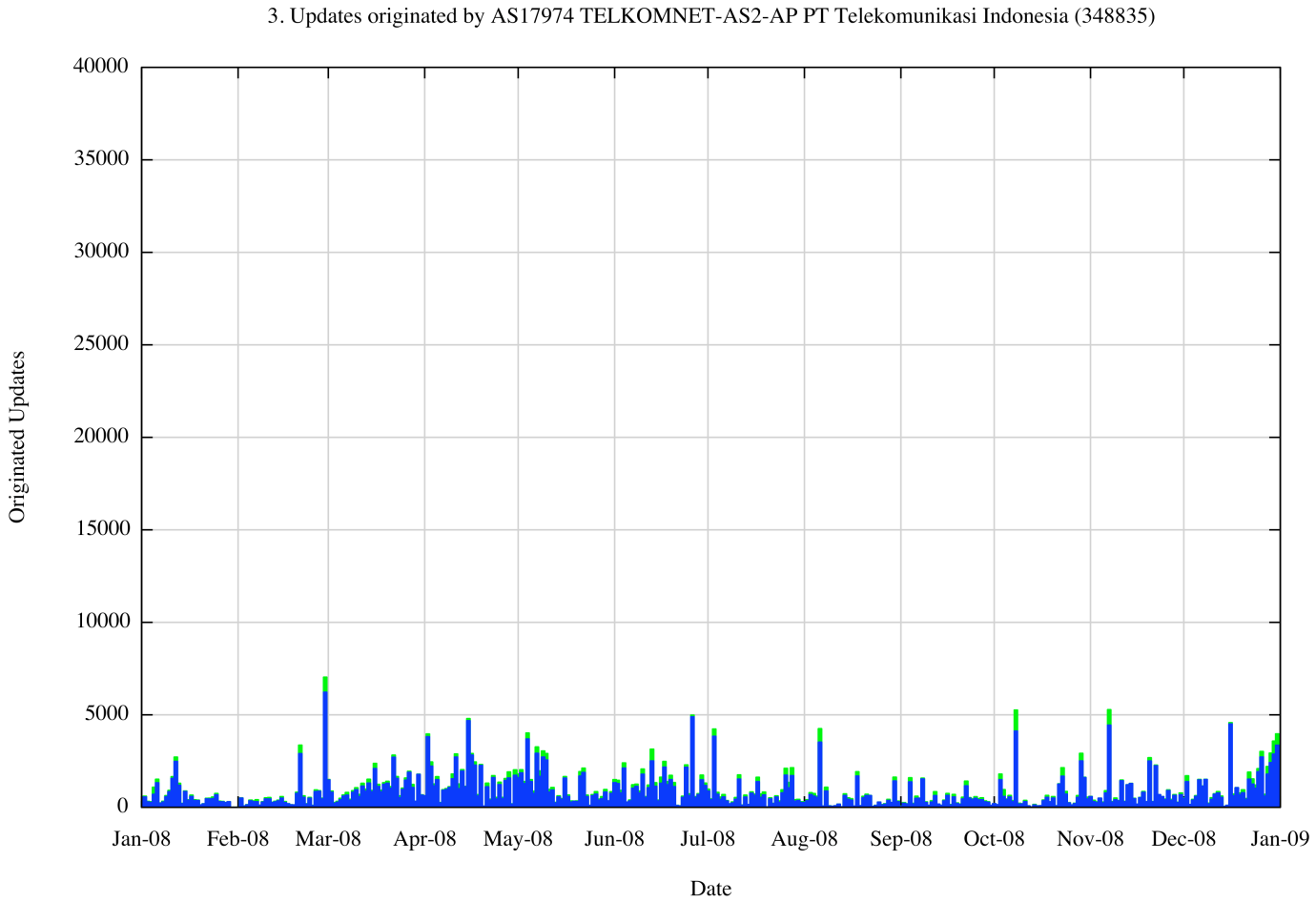
AS Adjacency: Up: 3, Down: 0

UP: 18051, **7713**, 38513

Originated Prefixes: 484

Aggregateable : 251

#3 AS17974


3. Updates originated by AS17974 TELKOMNET-AS2-AP PT Telekomunikasi Indonesia (348835)
Originated Updates
40000
35000
30000
25000
20000
15000
10000
5000
0
Jan-08
Feb-08
Mar-08
Apr-08
May-08
Jun-08
Jul-08
Aug-08
Sep-08
Oct-08
Nov-08
Dec-08
Jan-09
Date

#4 AS17488

HATHWAY-NET-AP Hathway IP Over Cable Internet (IN)

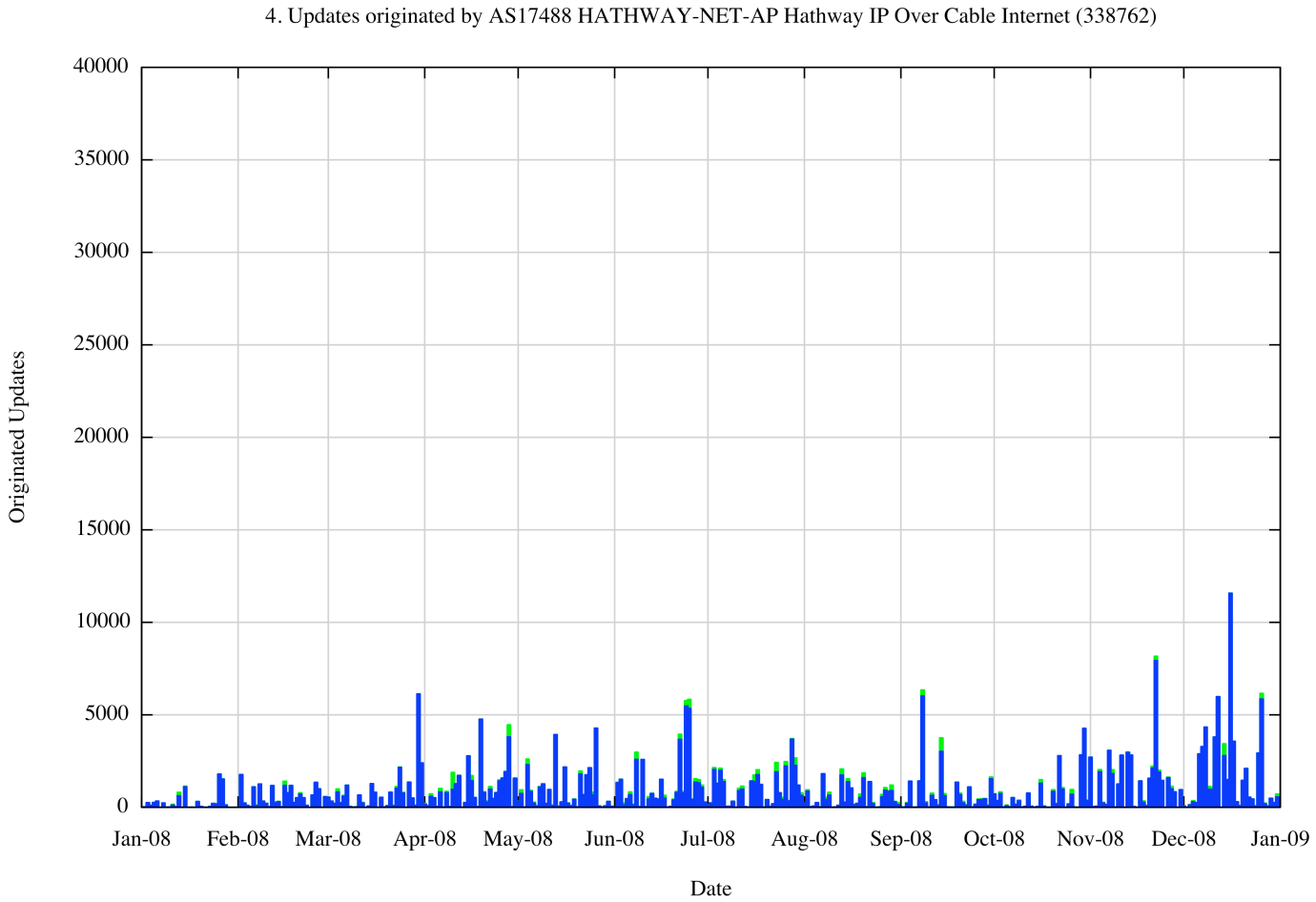
AS Adjacency: Up: 3 Down: 0

UP: 4755, 9730, 9498

Originated Prefixes: 1488

Aggregateable : 1321 (!)

#4 AS17488


4. Updates originated by AS17488 HATHWAY-NET-AP Hathway IP Over Cable Internet (338762)
Originated Updates
40000
35000
30000
25000
20000
15000
10000
5000
0
Jan-08
Feb-08
Mar-08
Apr-08
May-08
Jun-08
Jul-08
Aug-08
Sep-08
Oct-08
Nov-08
Dec-08
Jan-09
Date

#5 AS4621

UNINET-TH (TH)

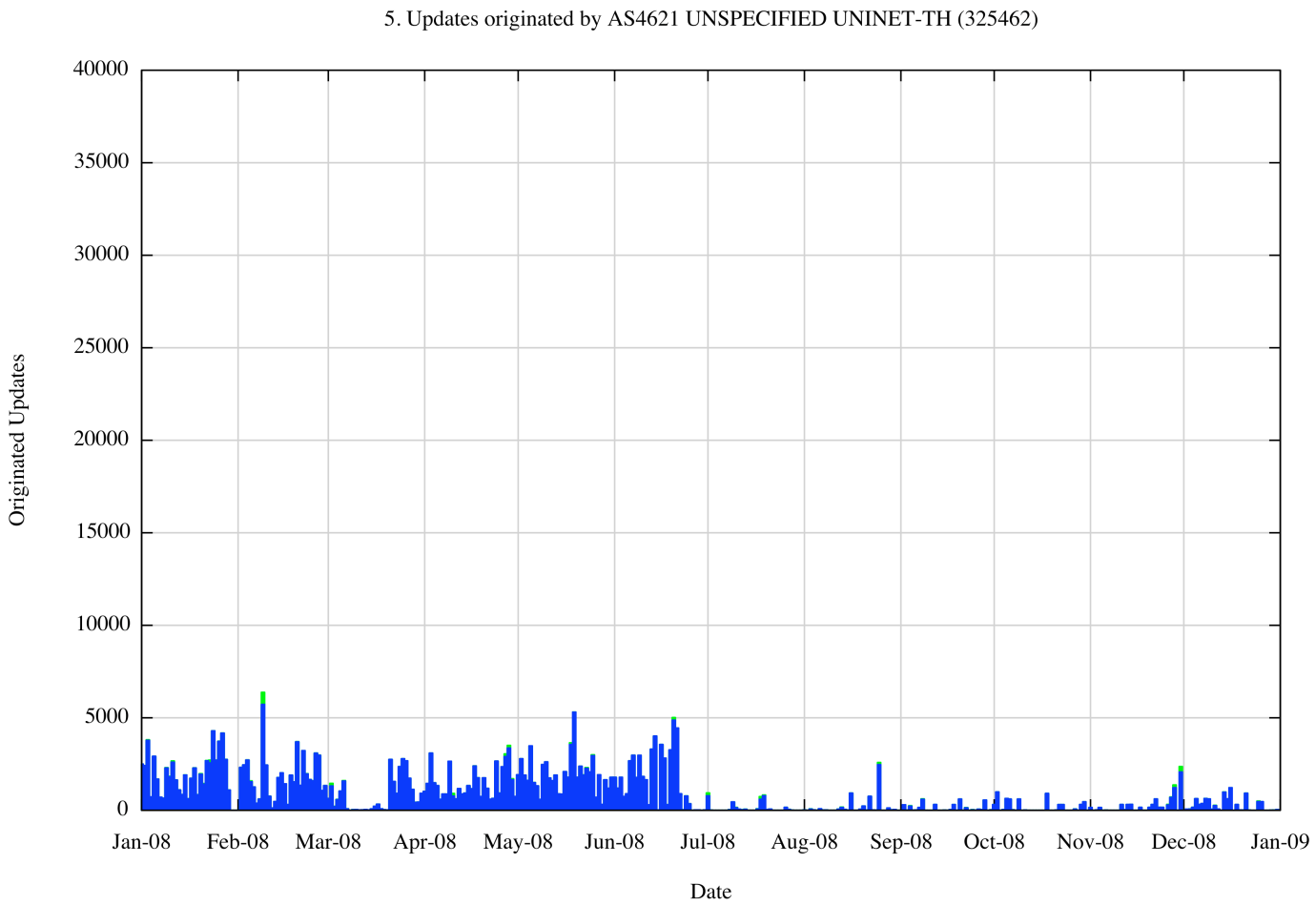
AS Adjacency: Up: 3, Down: 28

UP: 11537, 3491, 4561

Originated Prefixes: 184

Aggregateable : 95

#5 AS4621


5. Updates originated by AS4621 UNSPECIFIED UNINET-TH (325462)
Originated Updates
40000
35000
30000
25000
20000
15000
10000
5000
0
Jan-08
Feb-08
Mar-08
Apr-08
May-08
Jun-08
Jul-08
Aug-08
Sep-08
Oct-08
Nov-08
Dec-08
Jan-09
Date

#6 AS9385

GITS-TH-AS-AP Government Information Technology Services (TH)

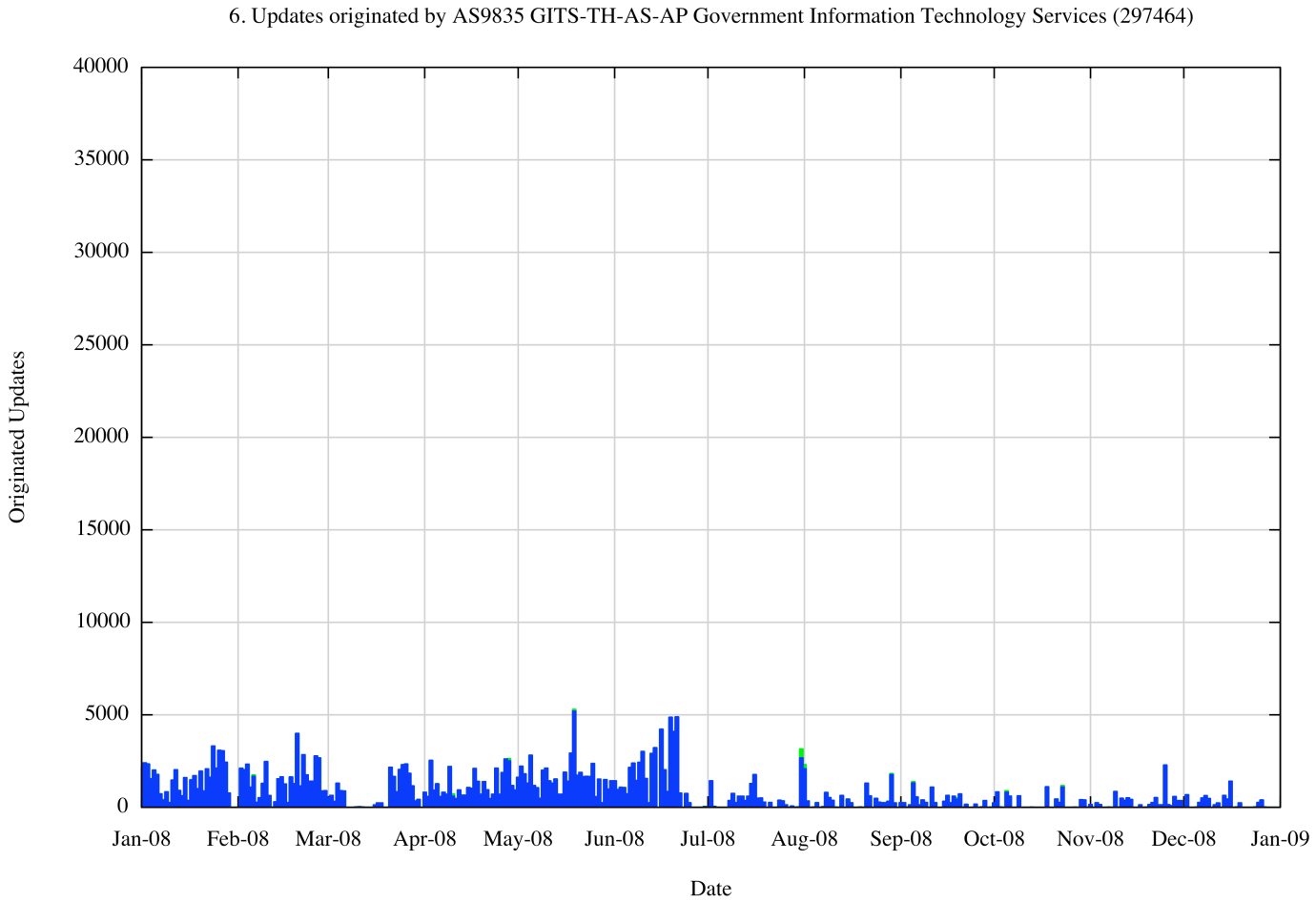
AS Adjacency: Up: 2 Down: 1

UP: 4750, 7470

Originated Prefixes: 129

Aggregateable: 78

#6 AS9835


6. Updates originated by AS9835 GITS-TH-AS-AP Government Information Technology Services (297464)
Originated Updates
40000
35000
30000
25000
20000
15000
10000
5000
0
Jan-08
Feb-08
Mar-08
Apr-08
May-08
Jun-08
Jul-08
Aug-08
Sep-08
Oct-08
Nov-08
Dec-08
Jan-09
Date

#7 AS9498

BBIL-AP BHARTI Airtel Ltd. (IN)

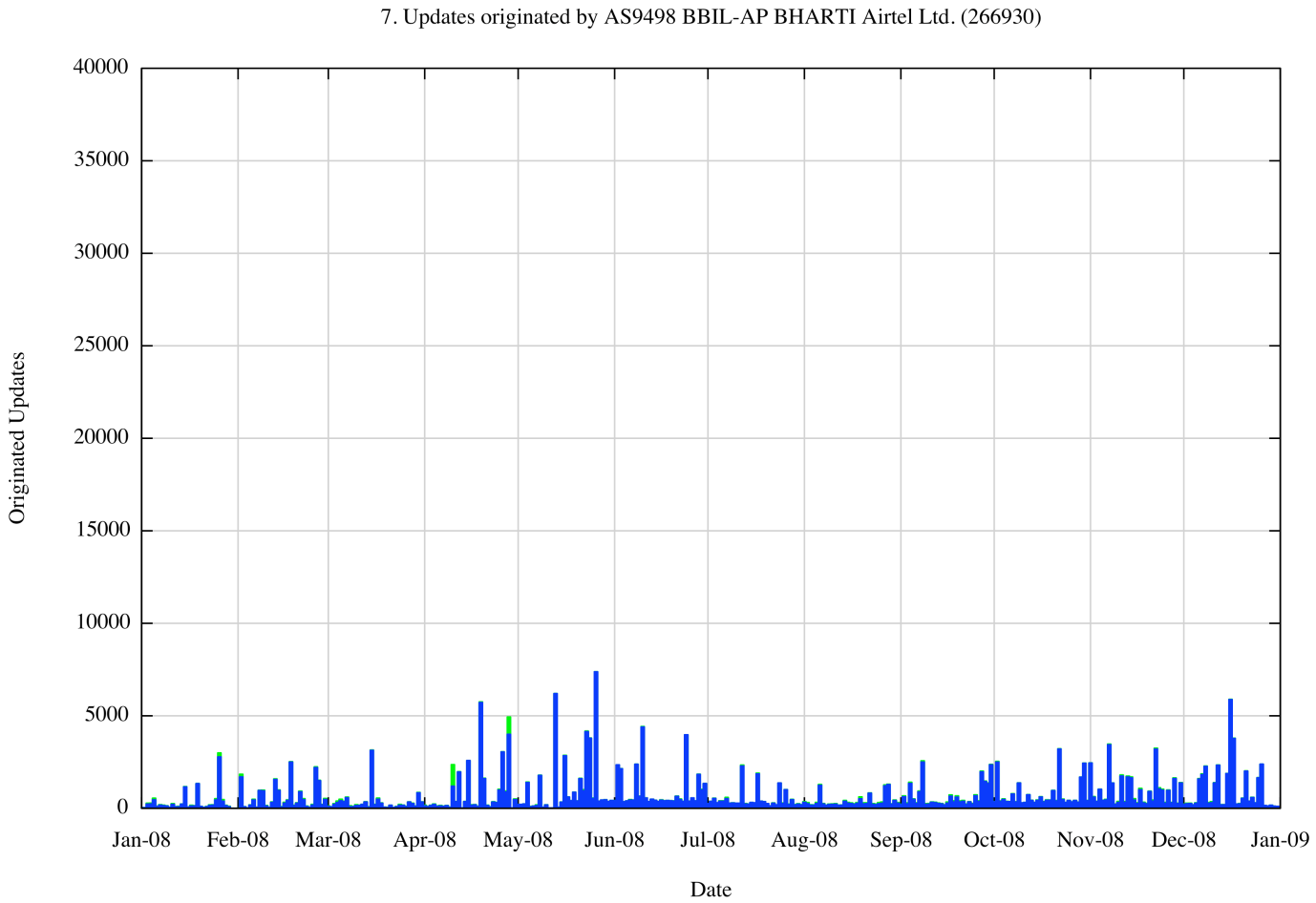
AS Adjacency: UP: 9 Down: 128

UP: 7473, 1239, 2914, 3356, 3491, 3320, 9730, 6453, 6762

Originated Prefixes: 678

Aggregateable : 407

#7 AS9498


7. Updates originated by AS9498 BBIL-AP BHARTI Airtel Ltd. (266930)
Originated Updates
40000
35000
30000
25000
20000
15000
10000
5000
0
Jan-08
Feb-08
Mar-08
Apr-08
May-08
Jun-08
Jul-08
Aug-08
Sep-08
Oct-08
Nov-08
Dec-08
Jan-09
Date

#8 AS4323

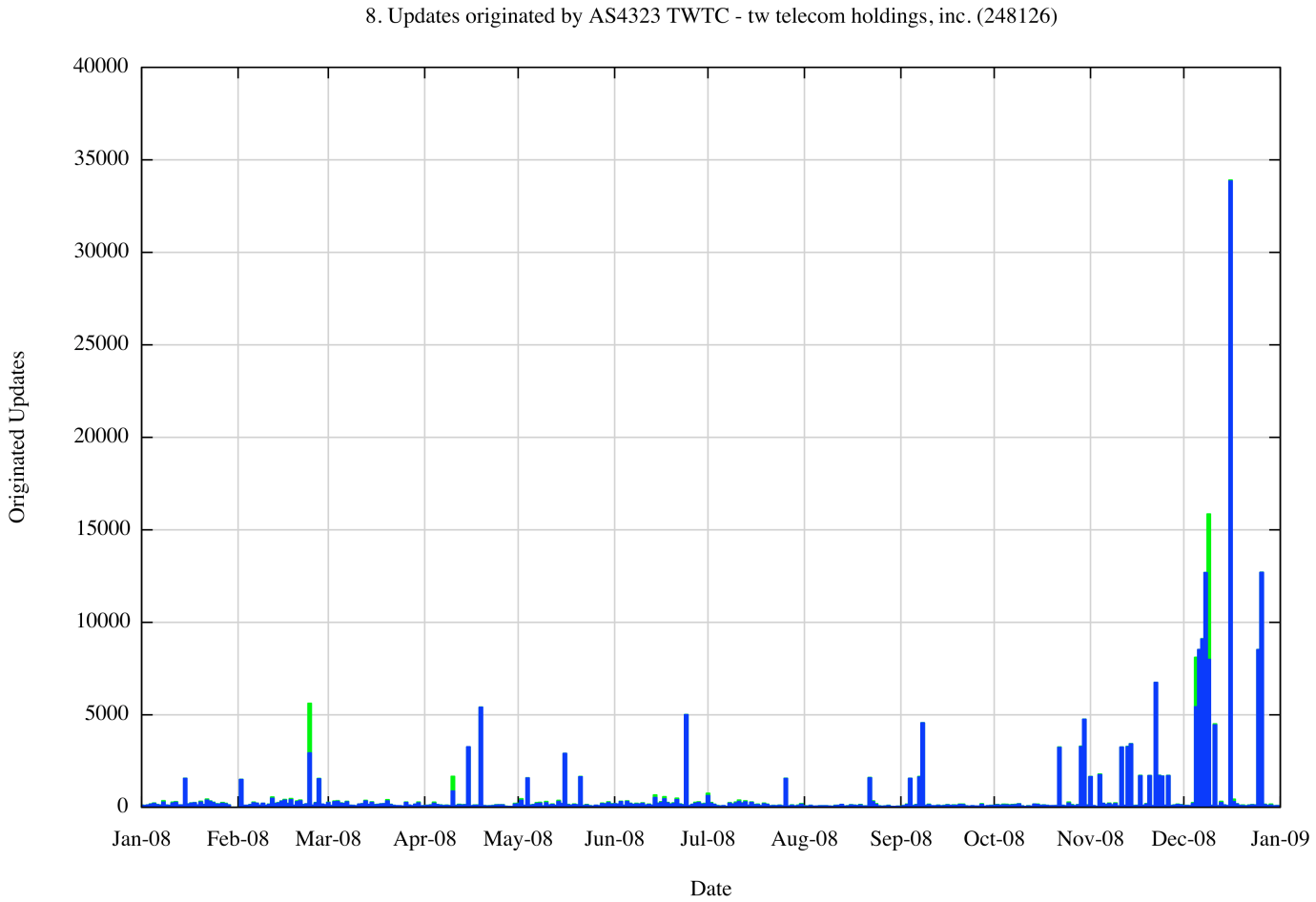
TWTC – Telecom Peering tw telecom holdings, inc. (US)

AS Adjacency: Up: 5 Down: 976

Originated Prefixes: 4203

Aggregatable: 3060 (!)

#8 AS4323


8. Updates originated by AS4323 TWTC - tw telecom holdings, inc. (248126)
Originated Updates
40000
35000
30000
25000
20000
15000
10000
5000
0
Jan-08
Feb-08
Mar-08
Apr-08
May-08
Jun-08
Jul-08
Aug-08
Sep-08
Oct-08
Nov-08
Dec-08
Jan-09
Date

#9 AS9829

BSNL-NIB National Internet Backbone (IN)

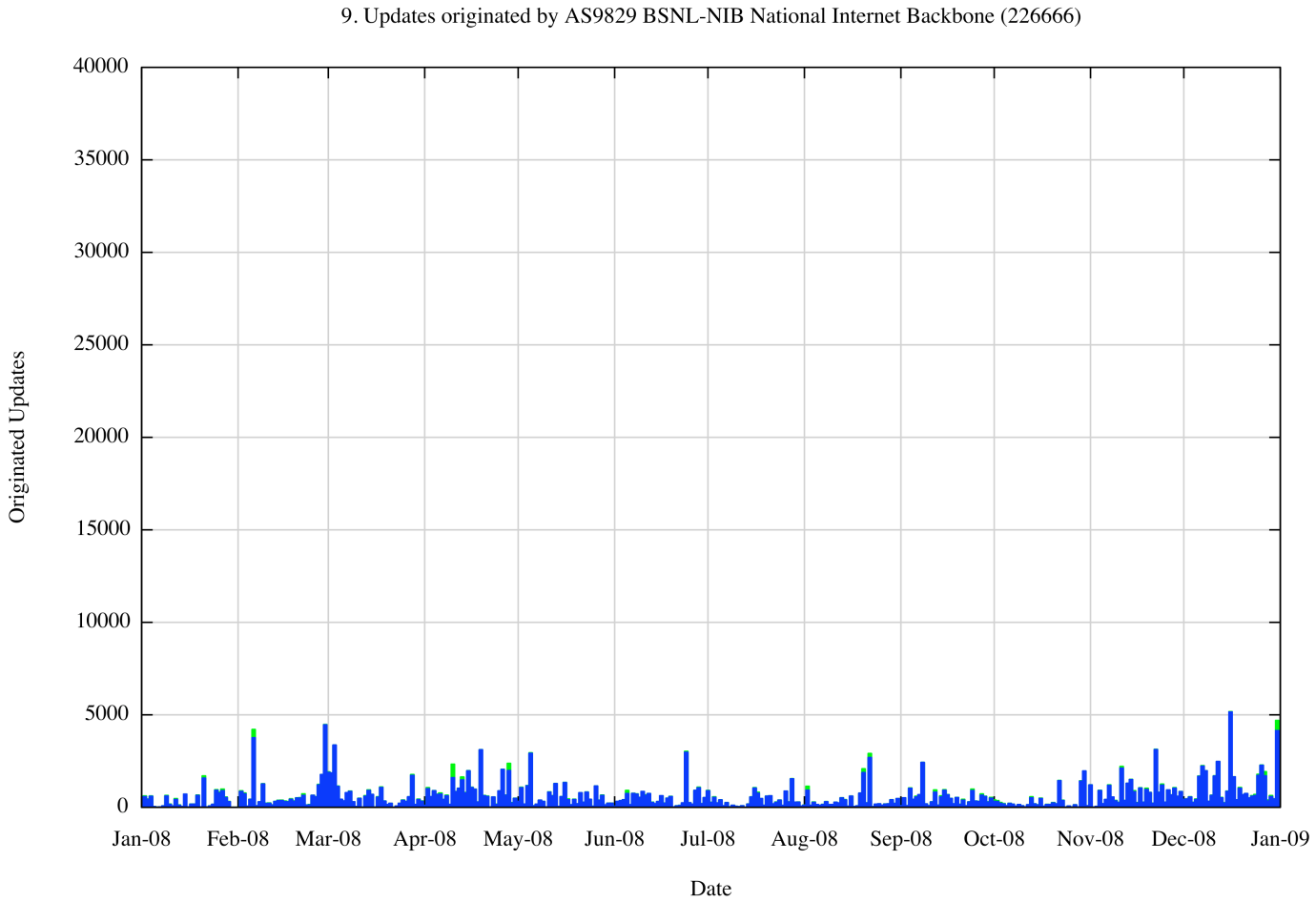
AS Adjacency: Up: 13 Down: 1

UP: 1239, 1273, 6453, 7018, 18101, 15412, 3561, 3356, 9910, 5400, 4755, 9498

Originated Prefixes: 624

Aggregateable: 273

#9 AS9829


9. Updates originated by AS9829 BSNL-NIB National Internet Backbone (226666)
Originated Updates
40000
35000
30000
25000
20000
15000
10000
5000
0
Jan-08
Feb-08
Mar-08
Apr-08
May-08
Jun-08
Jul-08
Aug-08
Sep-08
Oct-08
Nov-08
Dec-08
Jan-09
Date

#10 AS9583

SIFY-AS-IN Sify Limited (IN)

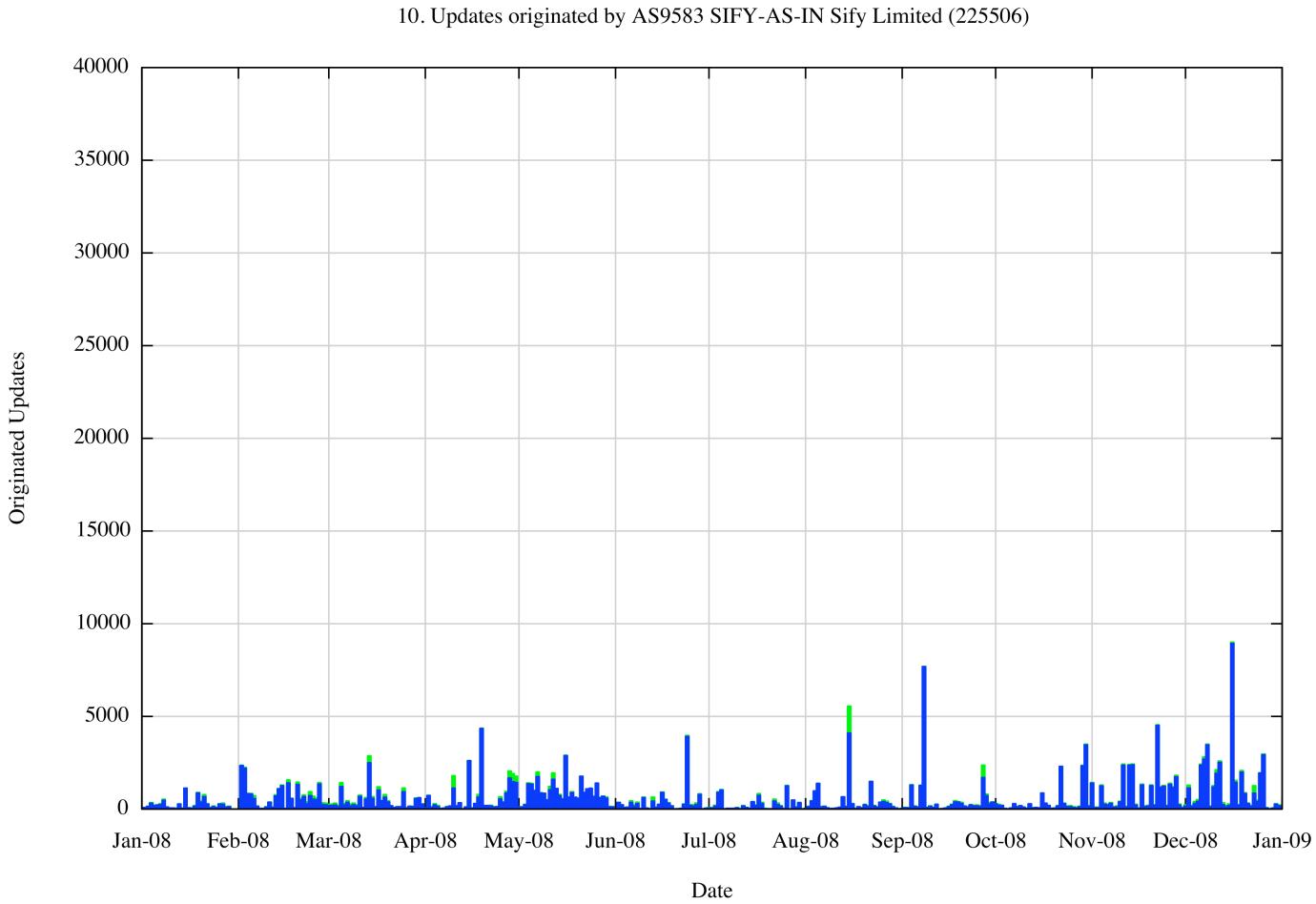
AS Adjacency: Up: 14 Down: 18

UP: 1239, 1299, 701, 7018, 2697, 10026, 3320, 3356, 3491, 5400, 4755, 9498, 6453, 7473

Originated Prefixes: 1111

Aggregateable: 275

#10 AS9583


10. Updates originated by AS9583 SIFY-AS-IN Sify Limited (225506)
Originated Updates
40000
35000
30000
25000
20000
15000
10000
5000
0
Jan-08
Feb-08
Mar-08
Apr-08
May-08
Jun-08
Jul-08
Aug-08
Sep-08
Oct-08
Nov-08
Dec-08
Jan-09
Date