APNIC 34 CONFERENCE

PHNOM PENH
CAMBODIA

21 - 31 August 2012

A Progress Report on BGP

Geoff Huston

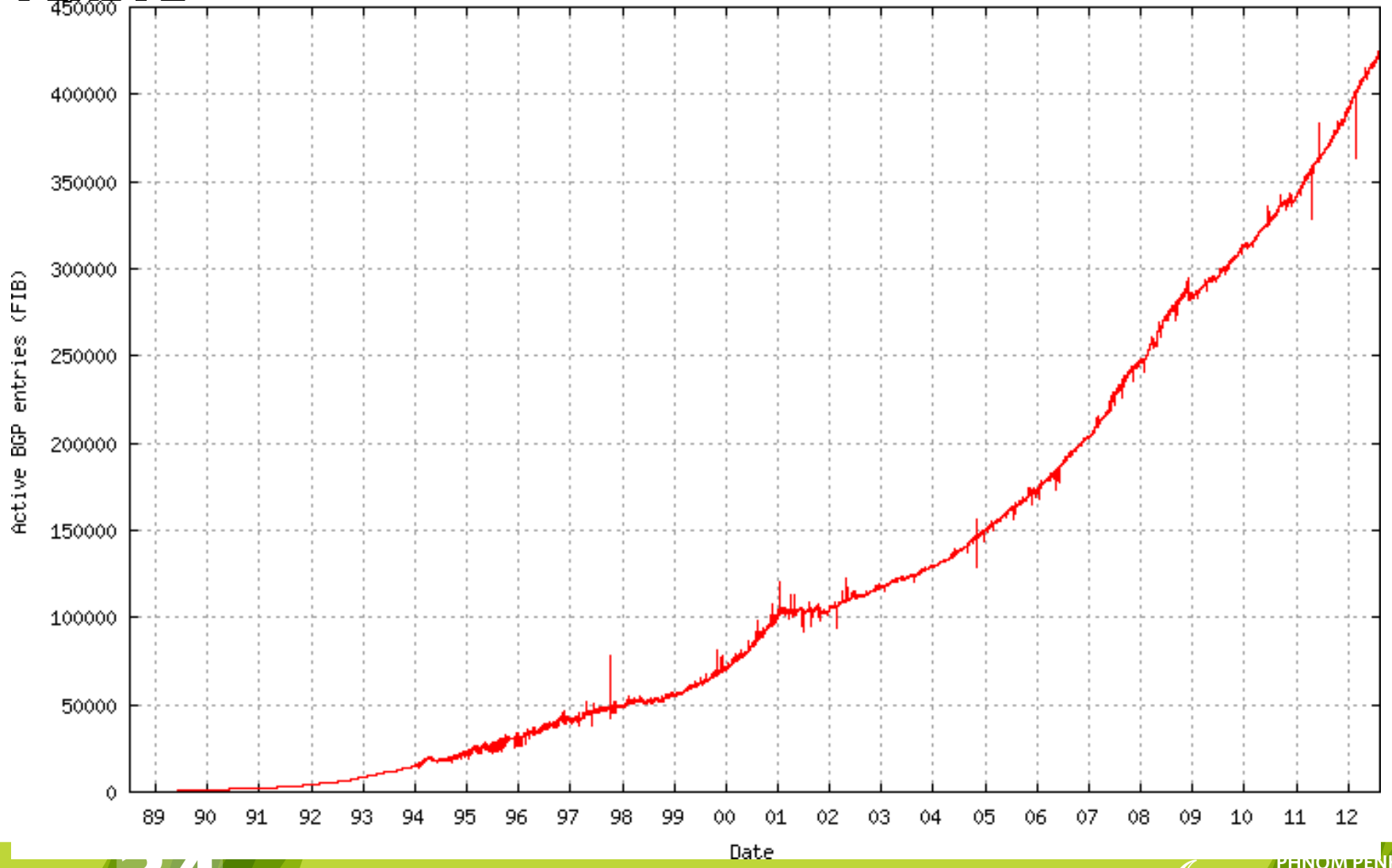
APNIC

Agenda

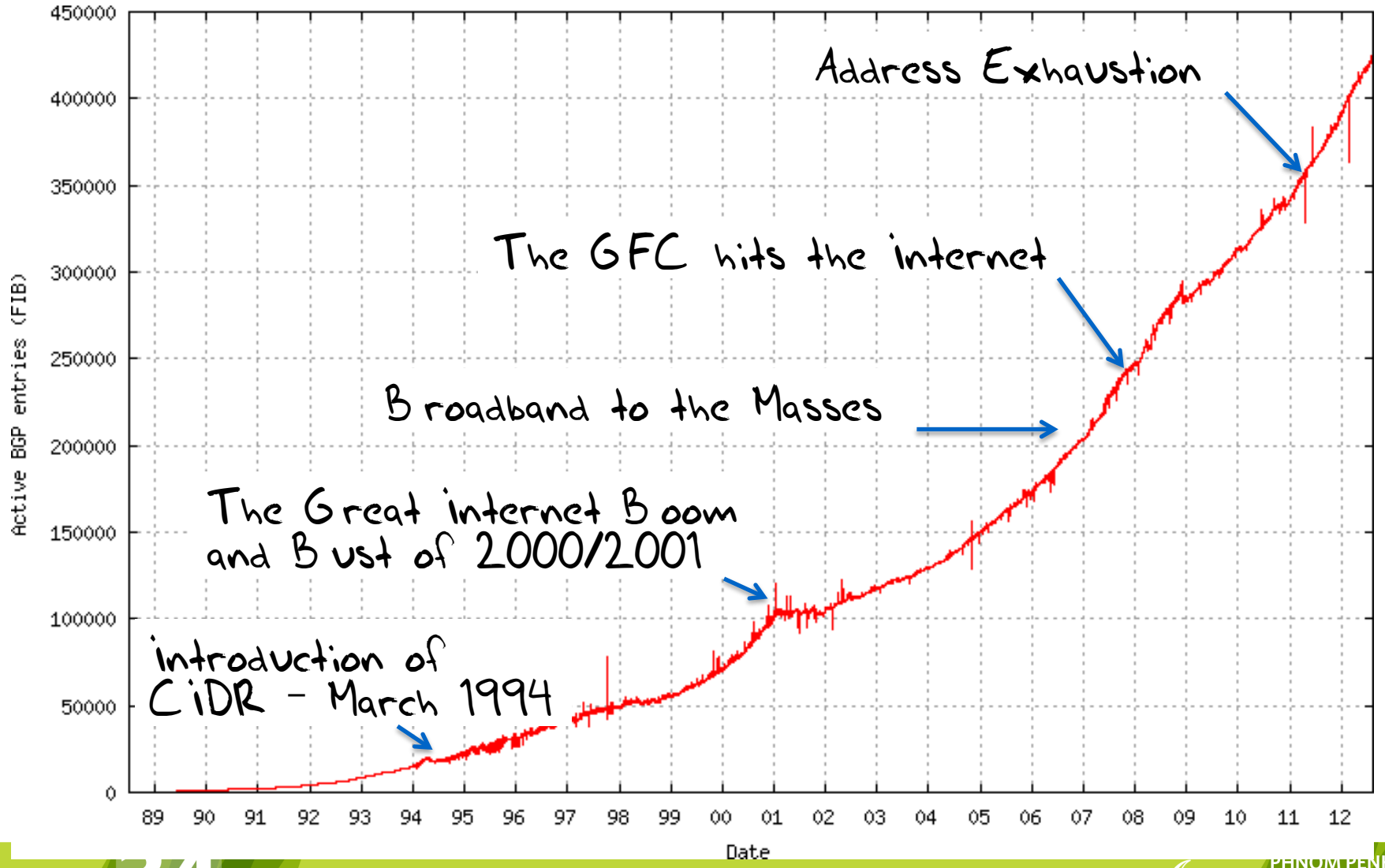
In this presentation we will explore the space of inter-domain routing (the Border Gateway Protocol – BGP

- We will look at the growth of the eBGP routing table over time and some projections for future growth
- Then we'll look at the dynamic behaviour of eBGP, and the extent to which more specifics are dominating routing table growth ... or not

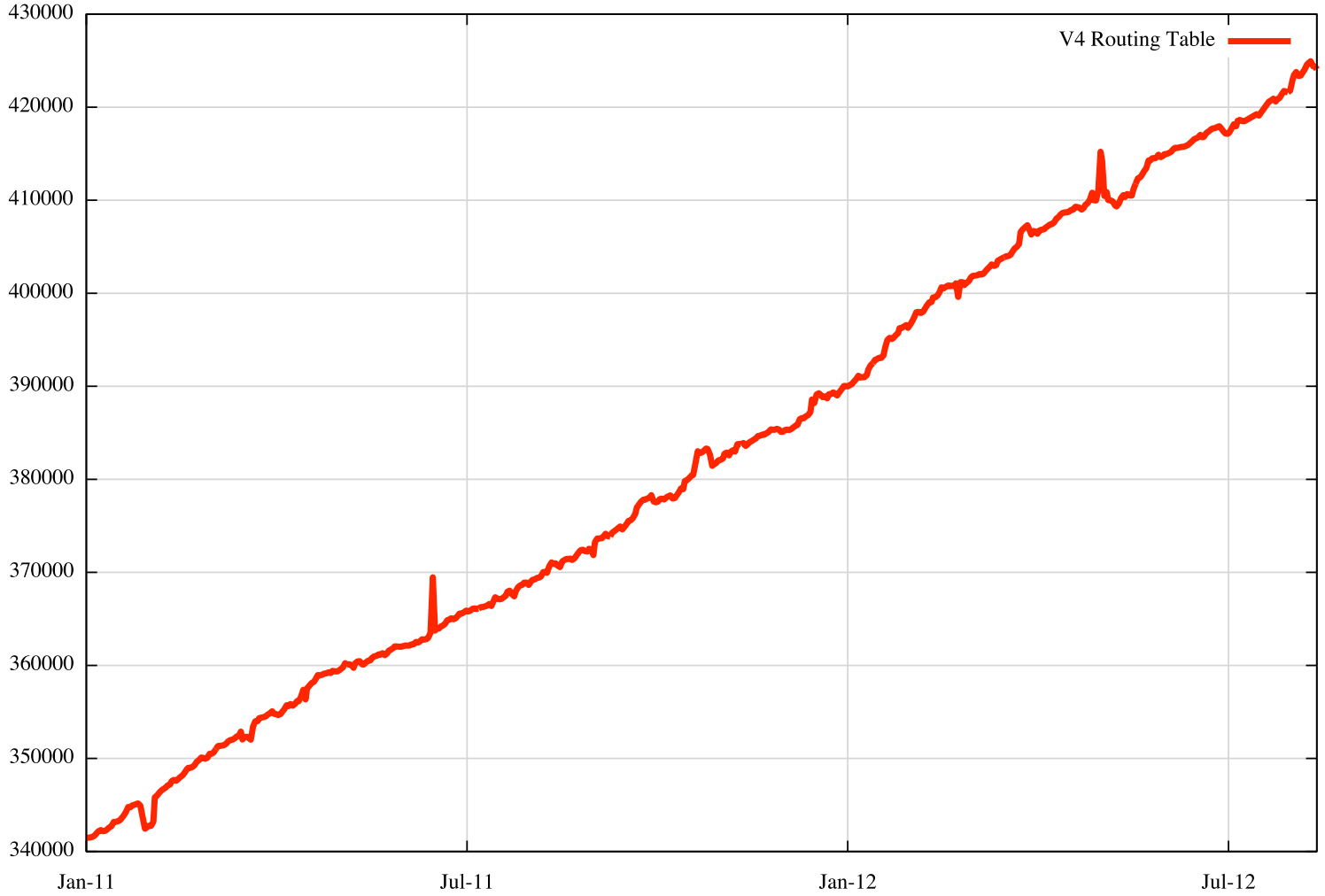
The Big Picture of the v4 Routing Table



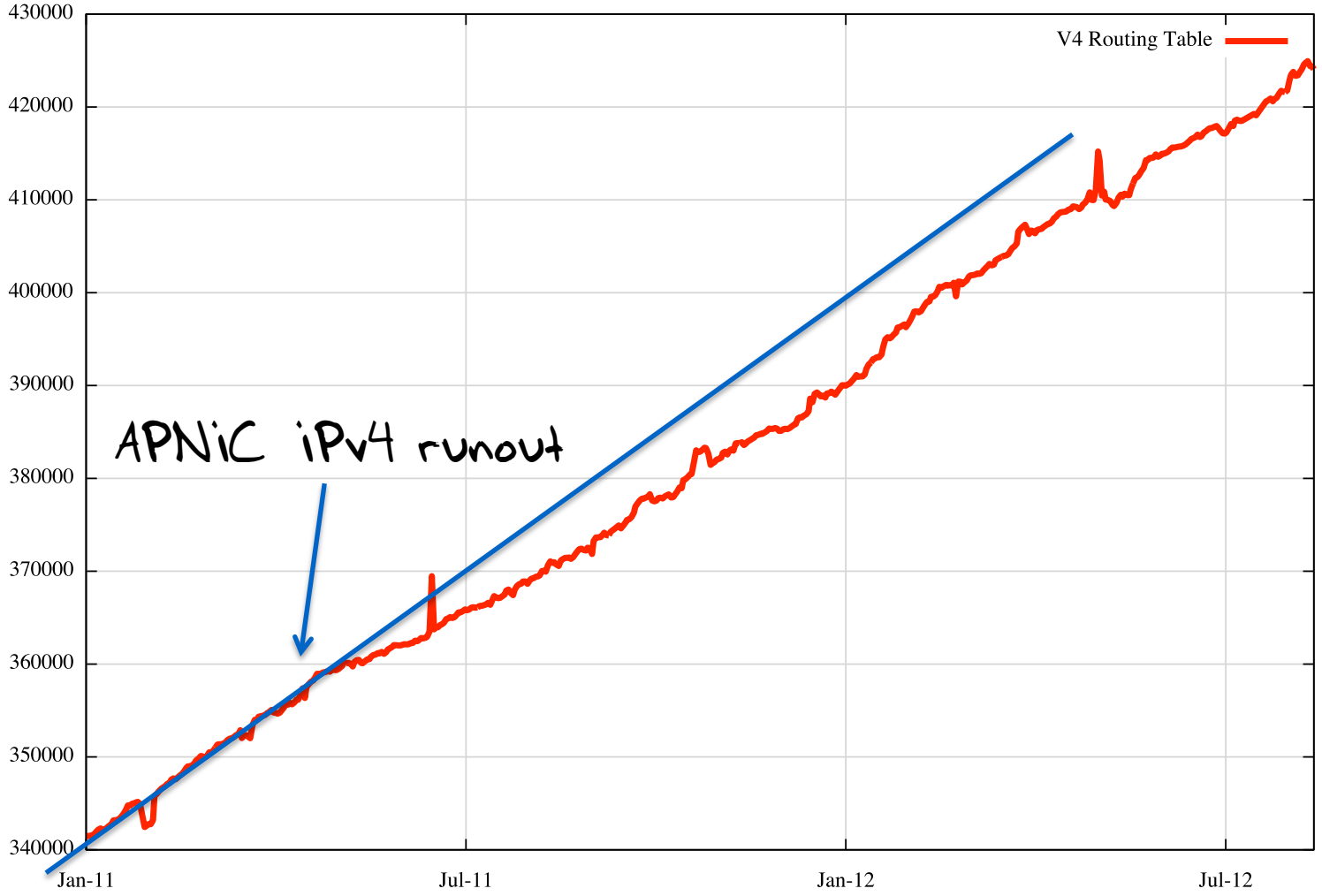
The Big Picture of the v4 Routing Table



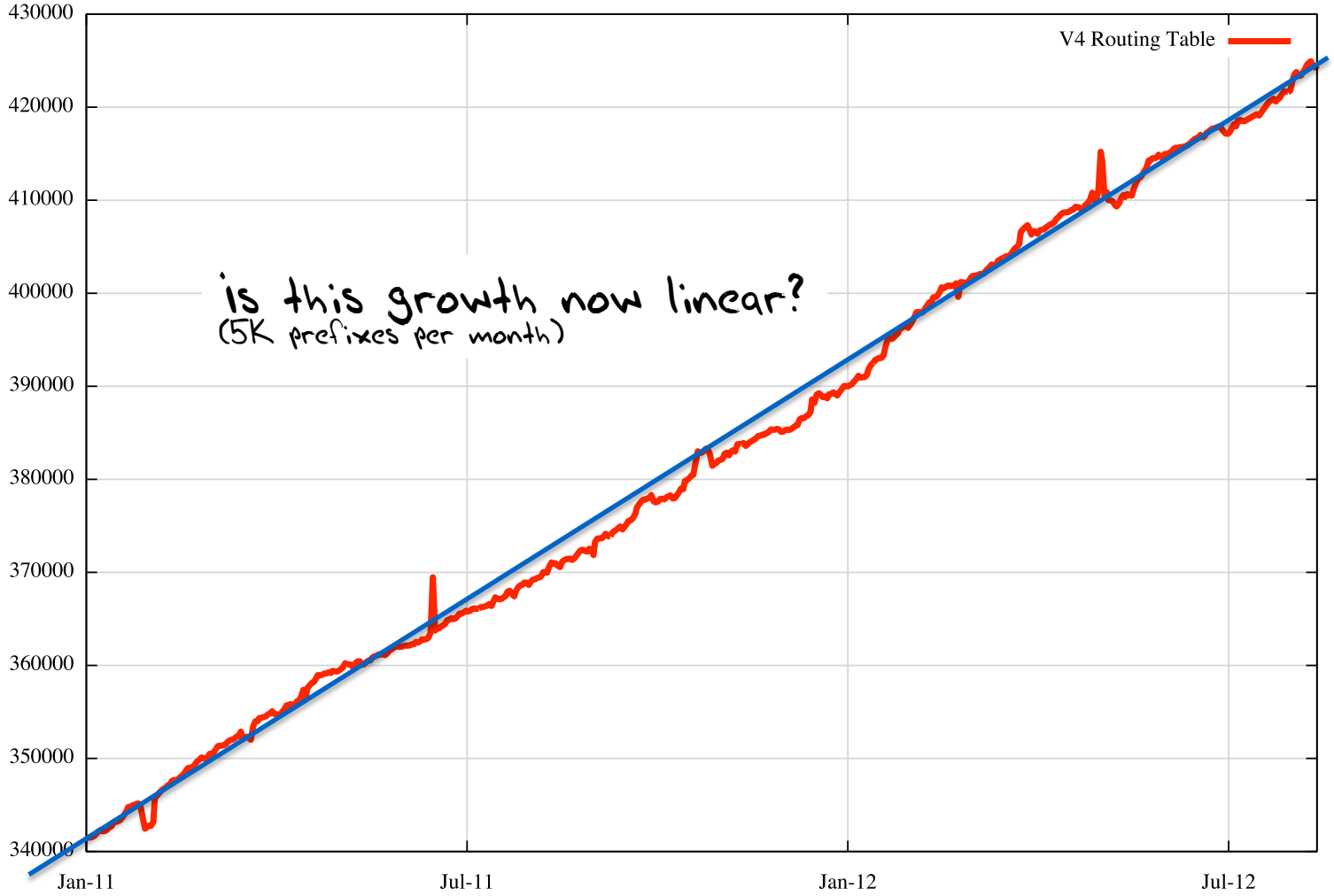
IPv4 BGP Prefix Count 2011 - 2012



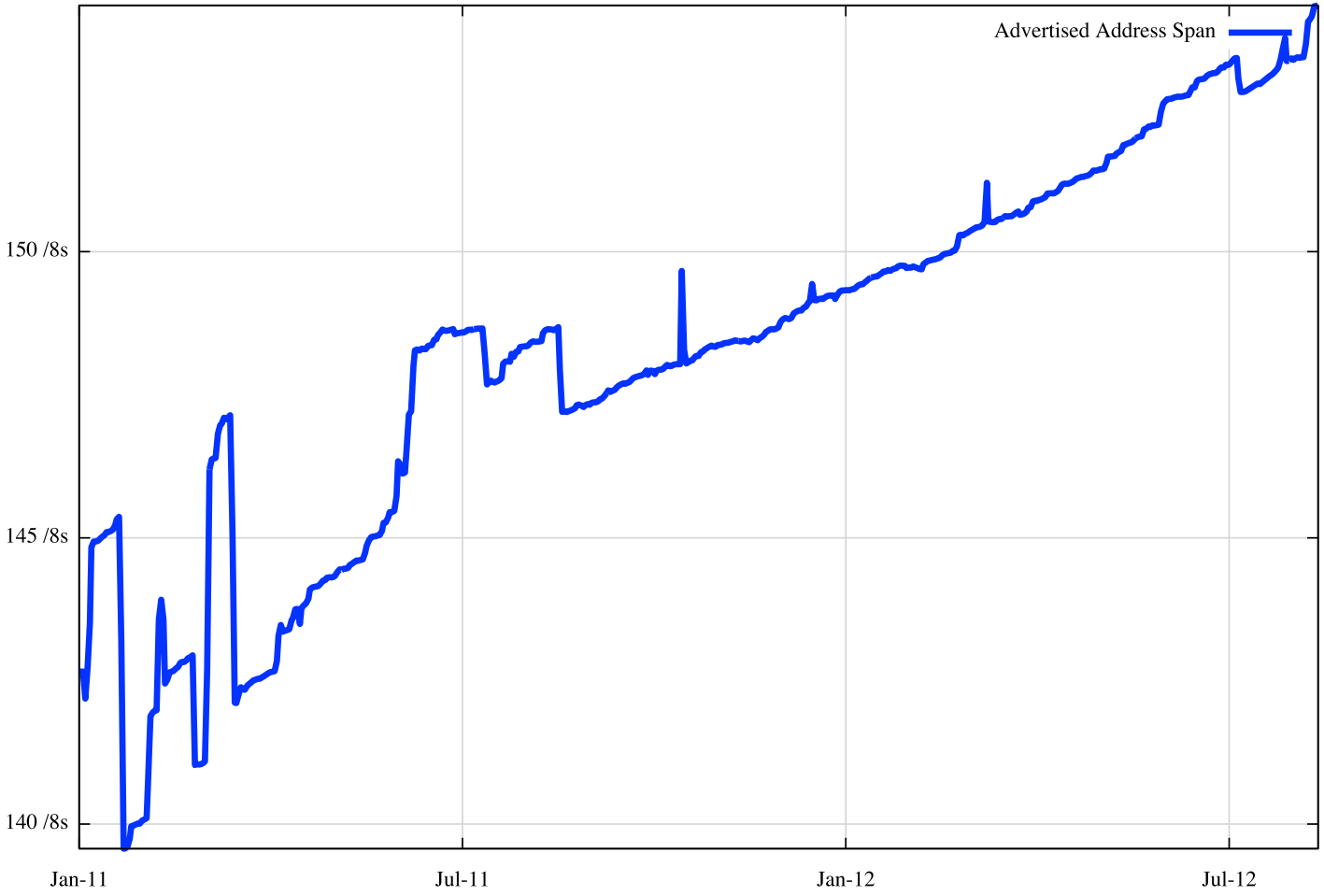
IPv4 BGP Prefix Count 2011 - 2012



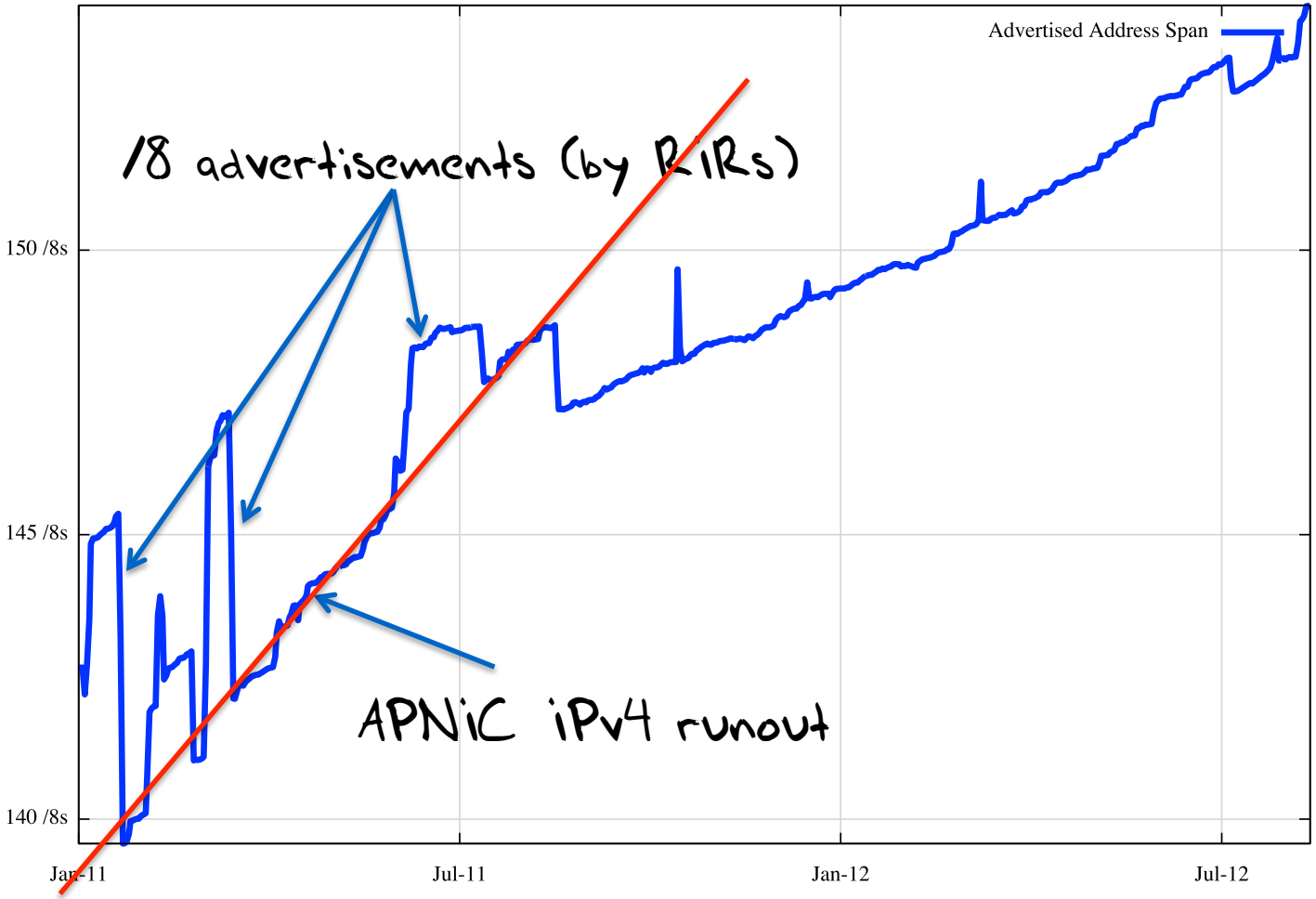
IPv4 BGP Prefix Count 2011 - 2012



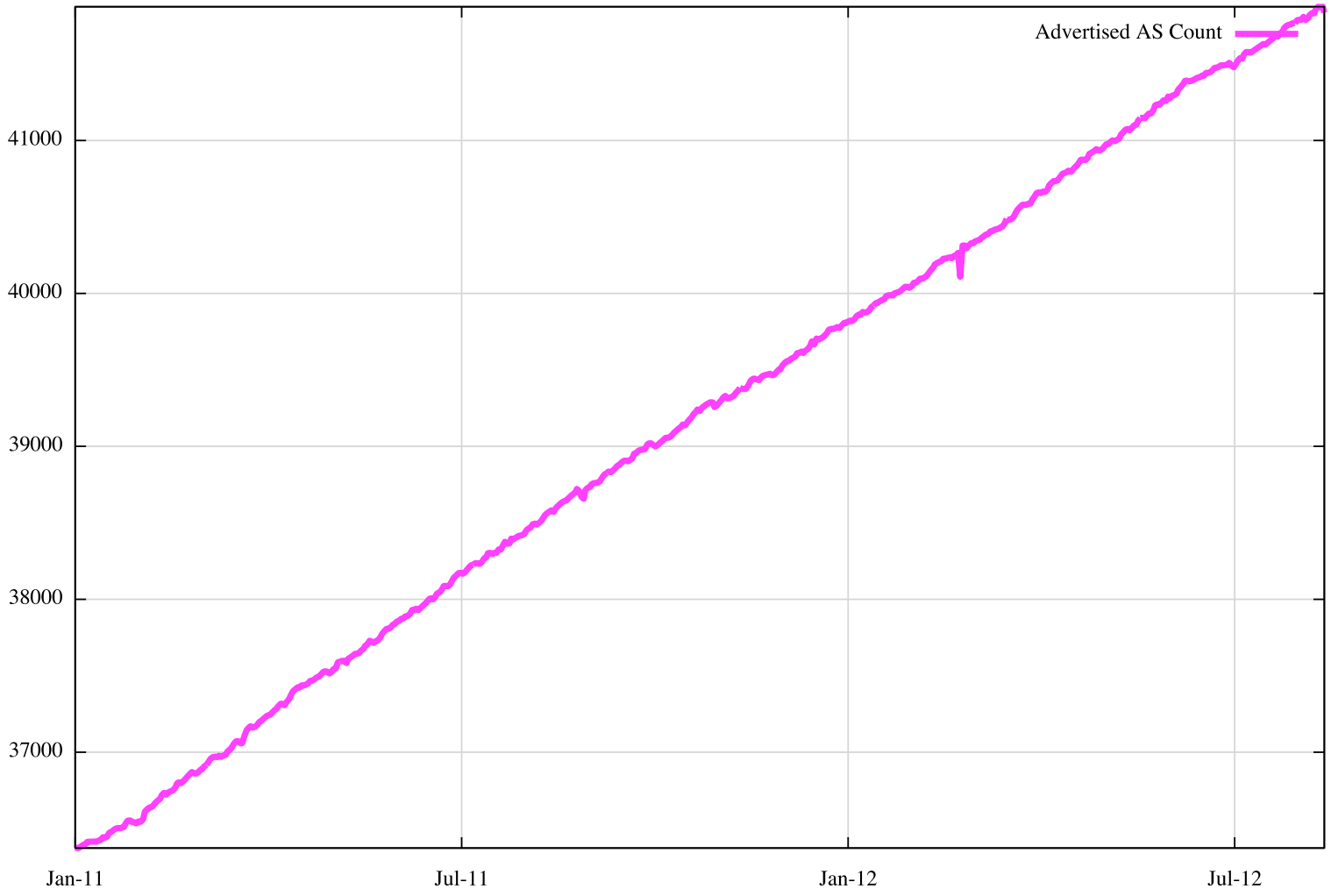
IPv4 Routed Address Span



IPv4 Routed Address Span



IPv4 Routed AS Count



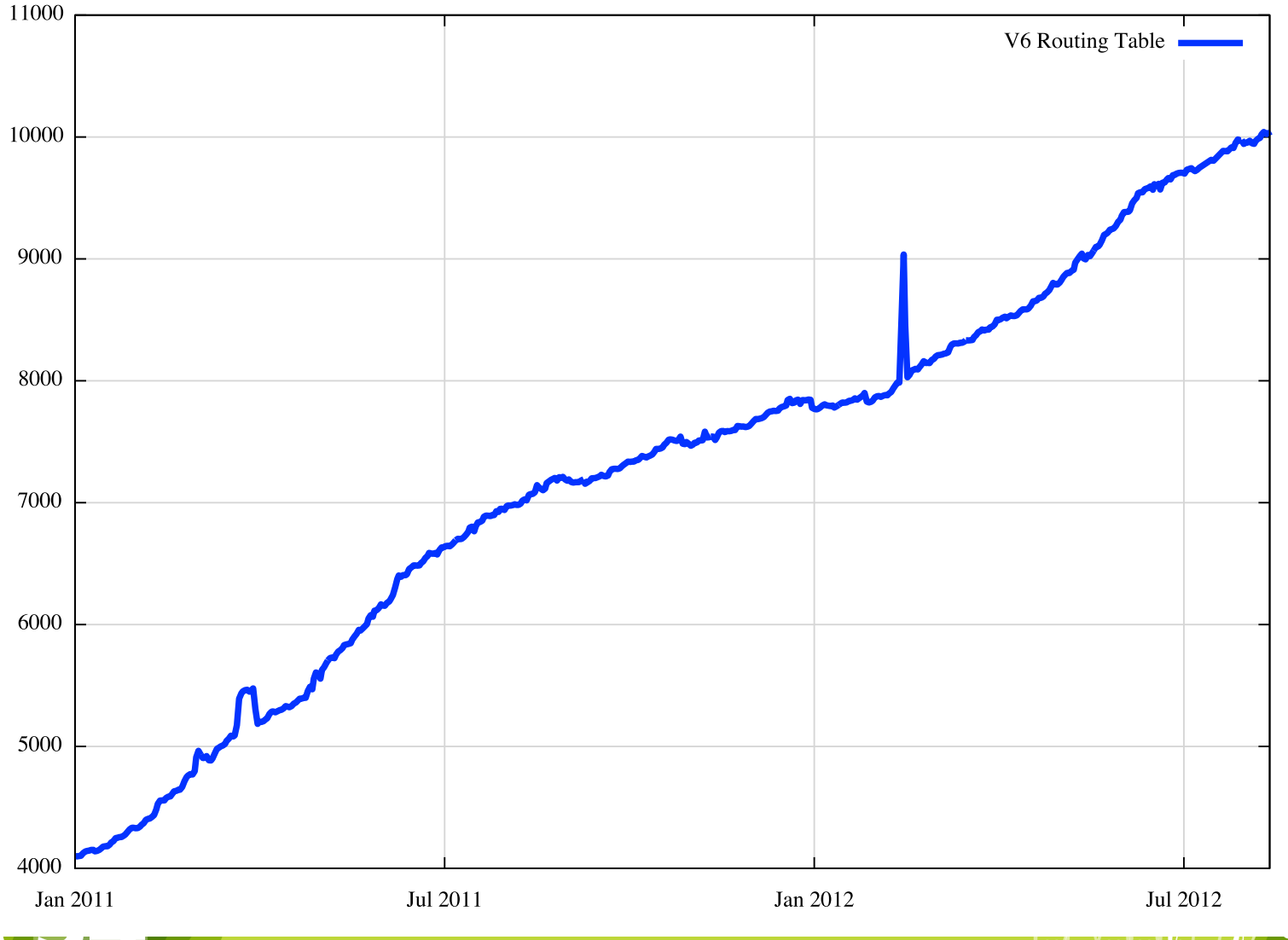
IPv4 2012 BGP Vital Statistics

	1-Jan-12	31-Jul-12	
Prefix Count	390,000	422,000	**+14% p.a.**
Roots	190,000	205,000	+13%
More Specifics	200,000	217,000	+15%
Address Span	149 /8s	153 /8s	**+ 5%**
AS Count	39,800	41,800	**+ 9%**
Transit	5,700	6,000	+ 9%
Stub	34,100	35,800	+ 9%

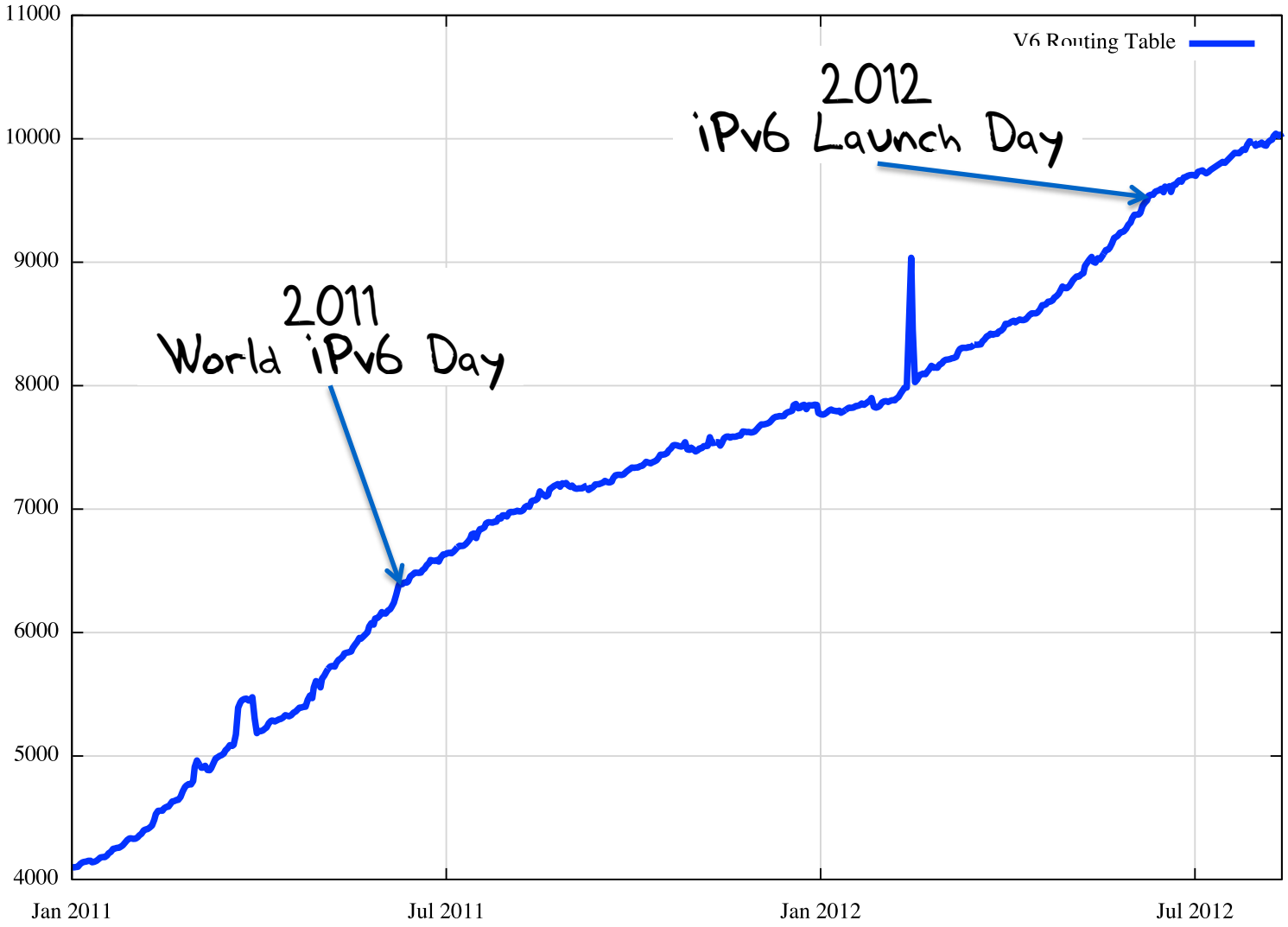
IPv4 in 2011

- Table growth remains surprisingly consistent
- Overall Internet growth in terms of BGP is at a rate of some ~14% p.a.
  - This is much the same as 2009 - 2011.
- Address span is now growing more slowly than the table size

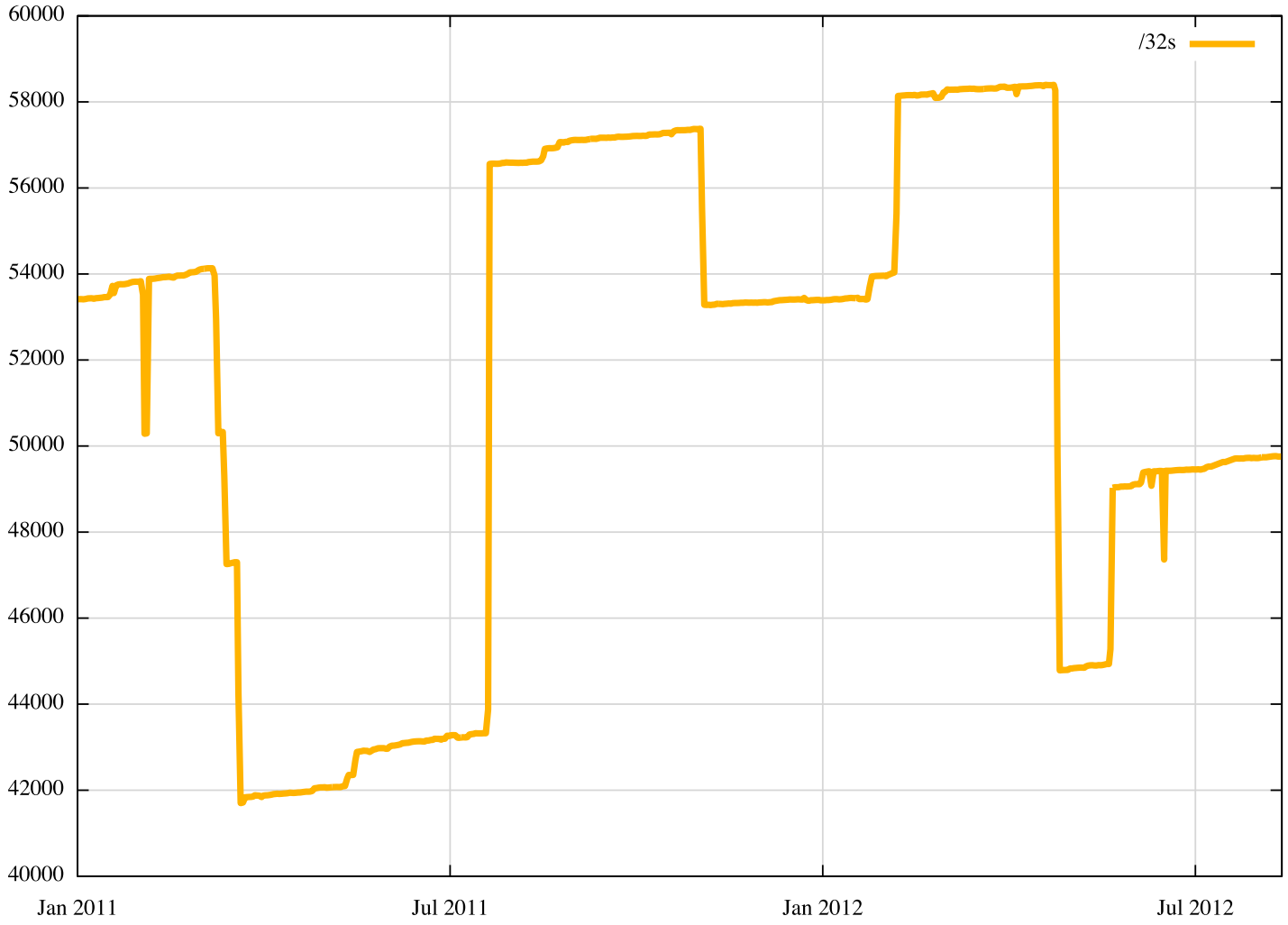
IPv6 BGP Prefix Count



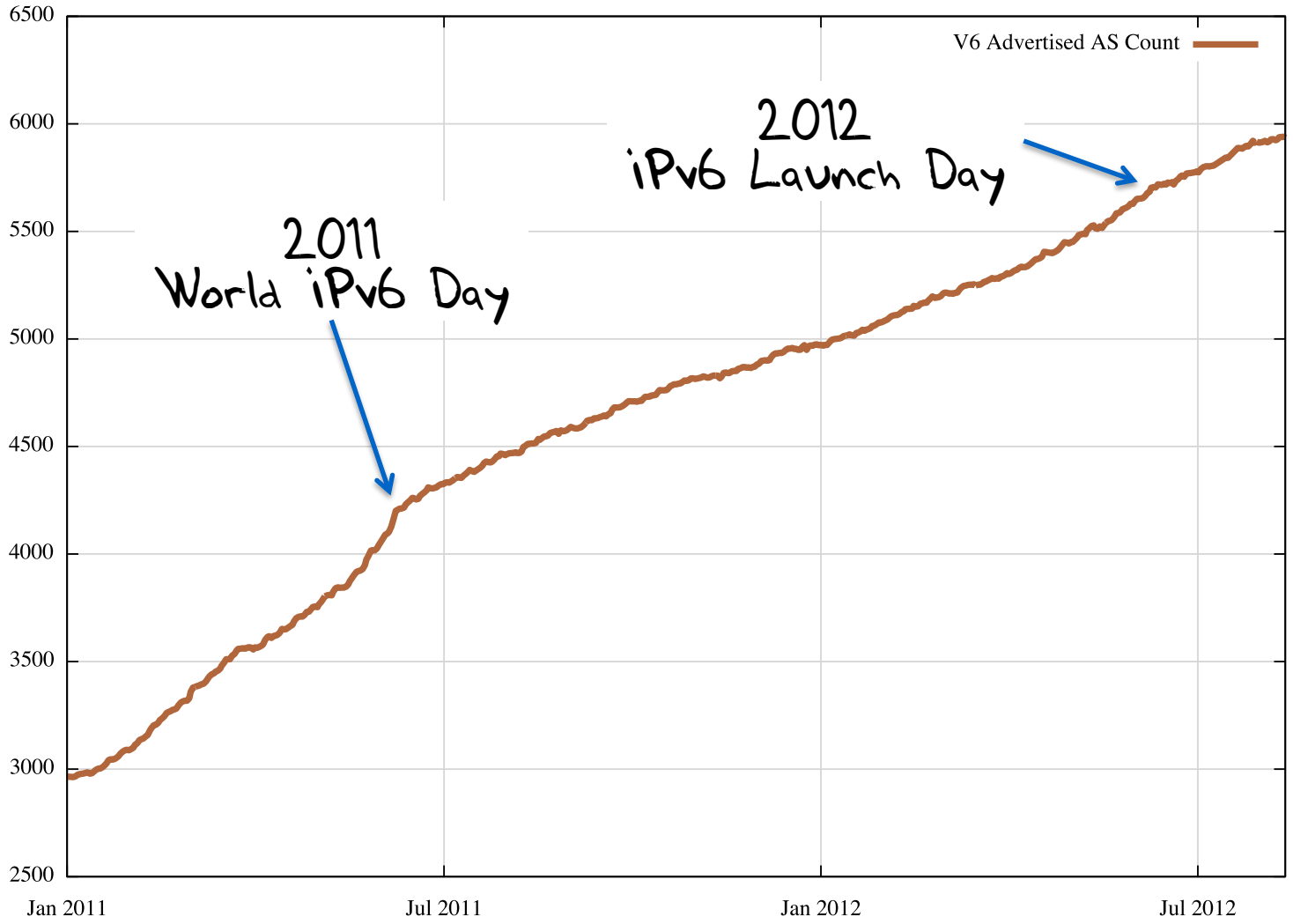
IPv6 BGP Prefix Count



IPv6 Routed Address Span



IPv6 Routed AS Count



IPv6 2011 BGP Vital Statistics

	Jan-12	Jul-12	p.a. rate
Prefix Count	7,759	9,950	+ 48%
Roots	5,751	6,420	+ 20%
More Specifics	2,008	3,530	+130%
Address Span (/32s)	53,387	49,803	- 12%
AS Count	4,968	5,915	+ 33%
Transit	985	985	+ 0%
Stub	3,983	4,730	+ 32%

IPv6 in 2010 - 2011

- Overall IPv6 Internet growth in terms of BGP is 50 % p.a.

(Looking at the AS count, if these relative growth rates persist then the IPv6 network would span the same network domain as IPv4 in 6 years time -- mid/late 2018)

Projections

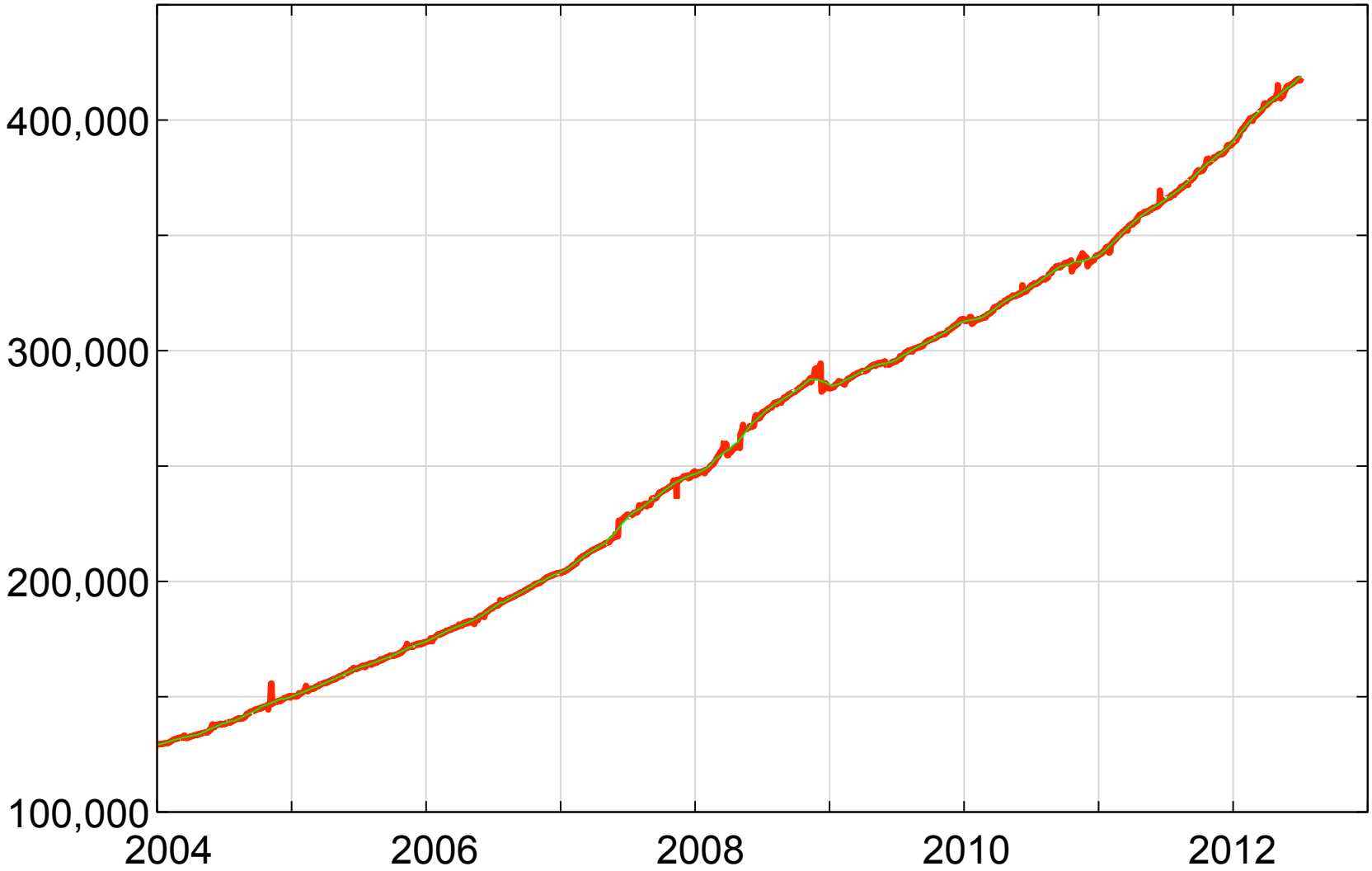
BGP Size Projections

- Generate a projection of the IPv4 routing table using a quadratic (O(2) polynomial) over the historic data
  - For IPv4 this is a time of **extreme uncertainty**
    - Registry IPv4 address run out
    - Uncertainty over the impacts of any after-market in IPv4 on the routing table

    which makes this projection even more

    speculative than normal!

IPv4 Table Size



Daily Growth Rates

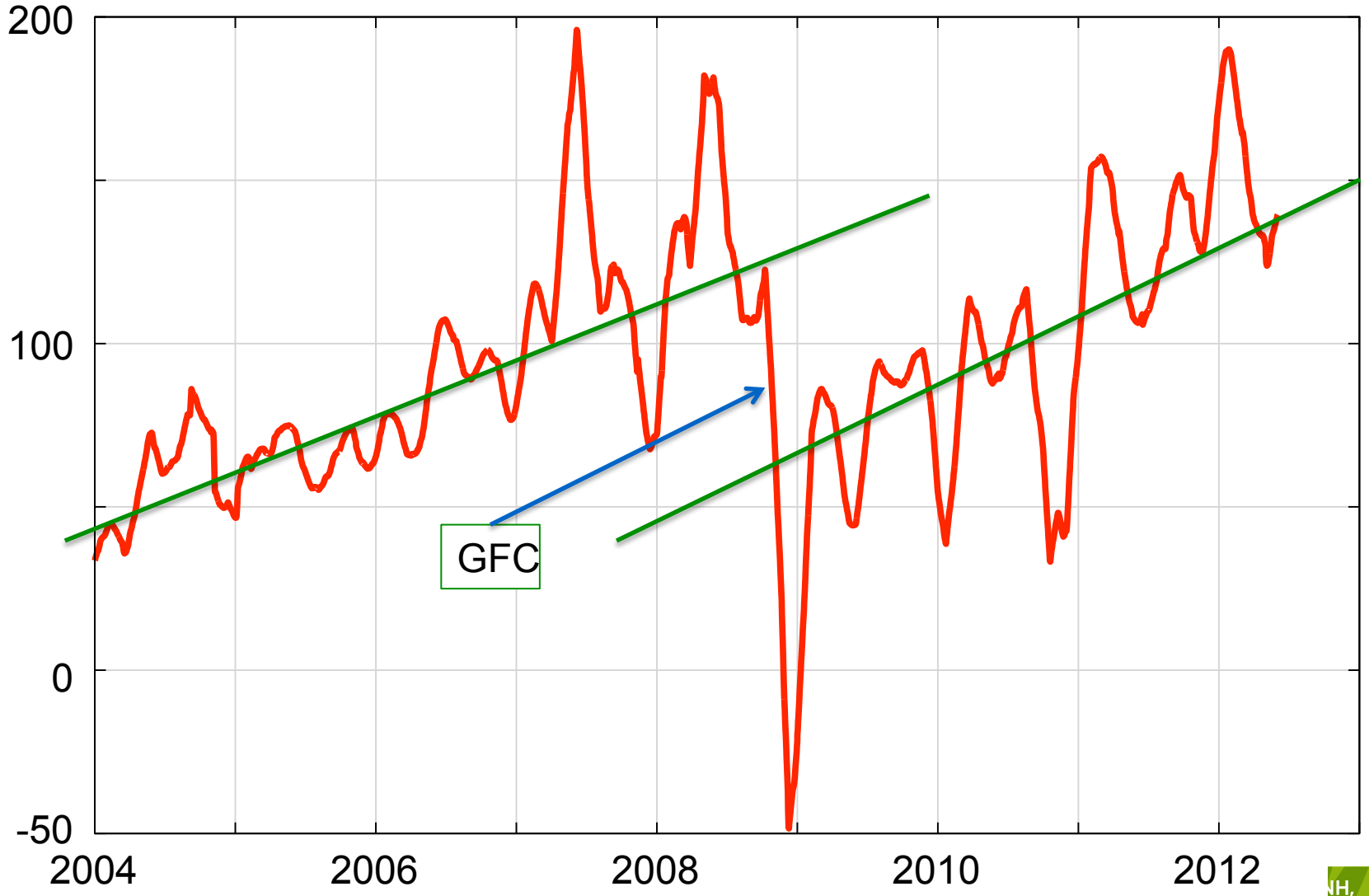
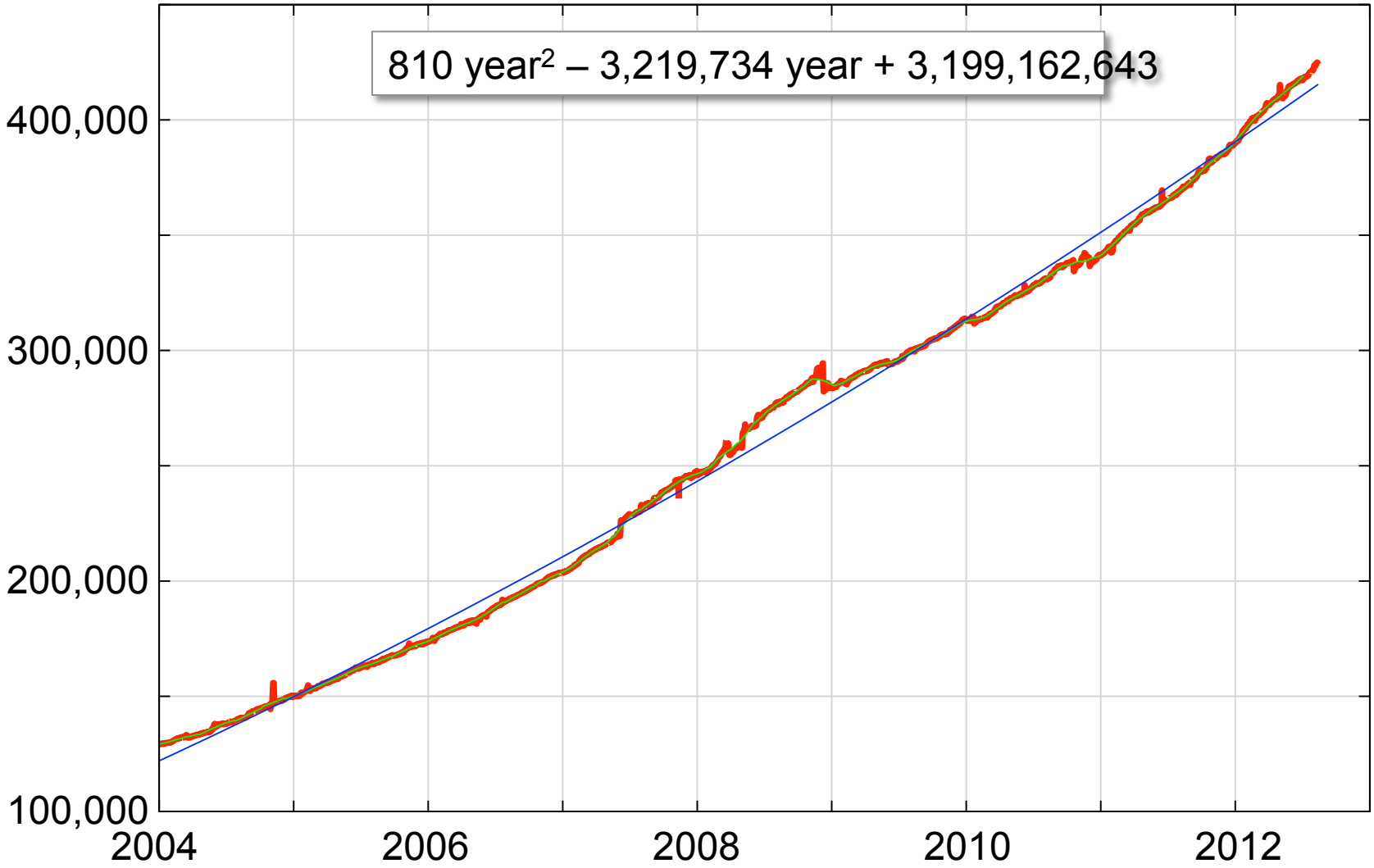
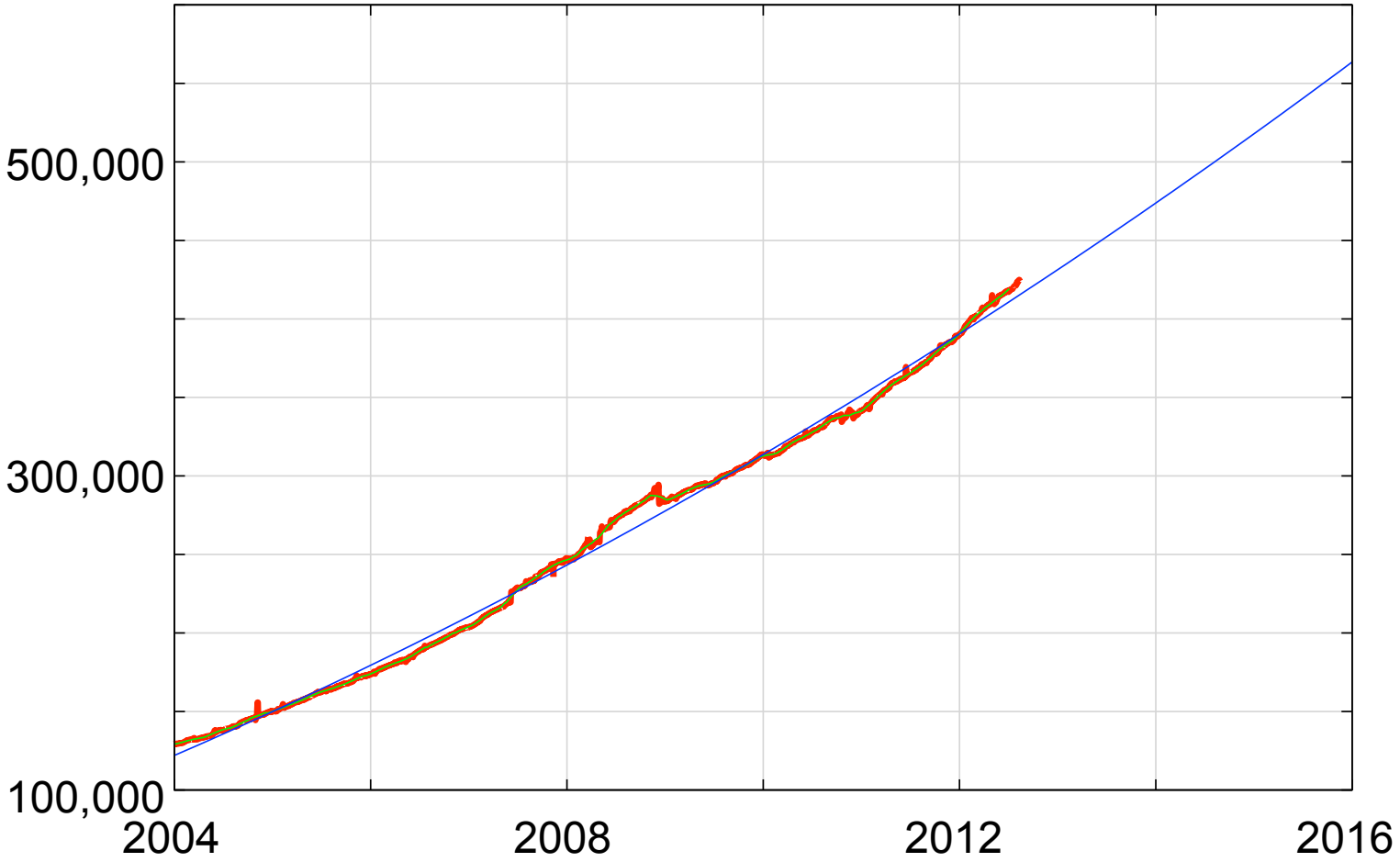


Table Growth Model



IPv4 Table Projection

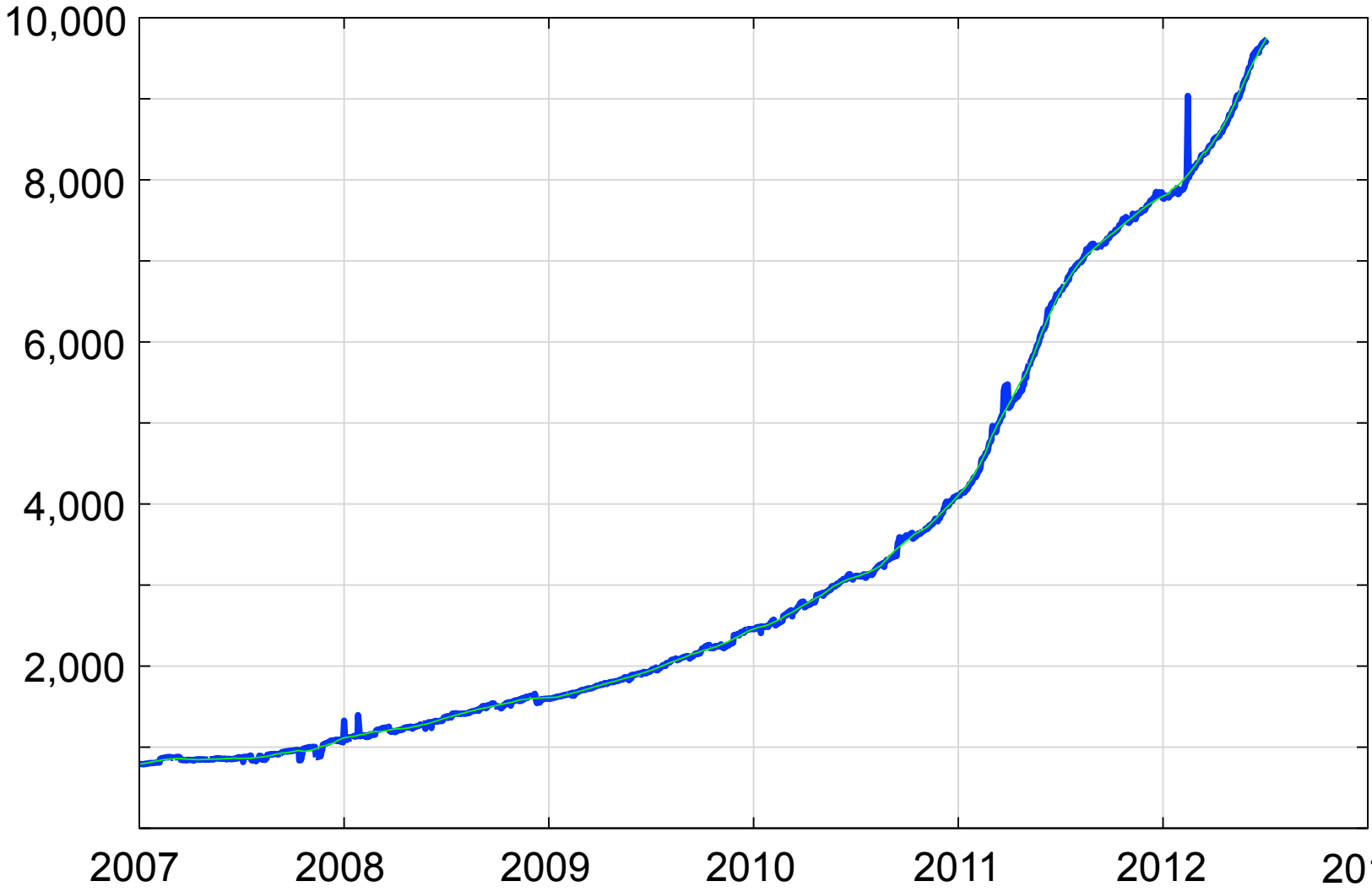


IPv4 BGP Table Size predictions

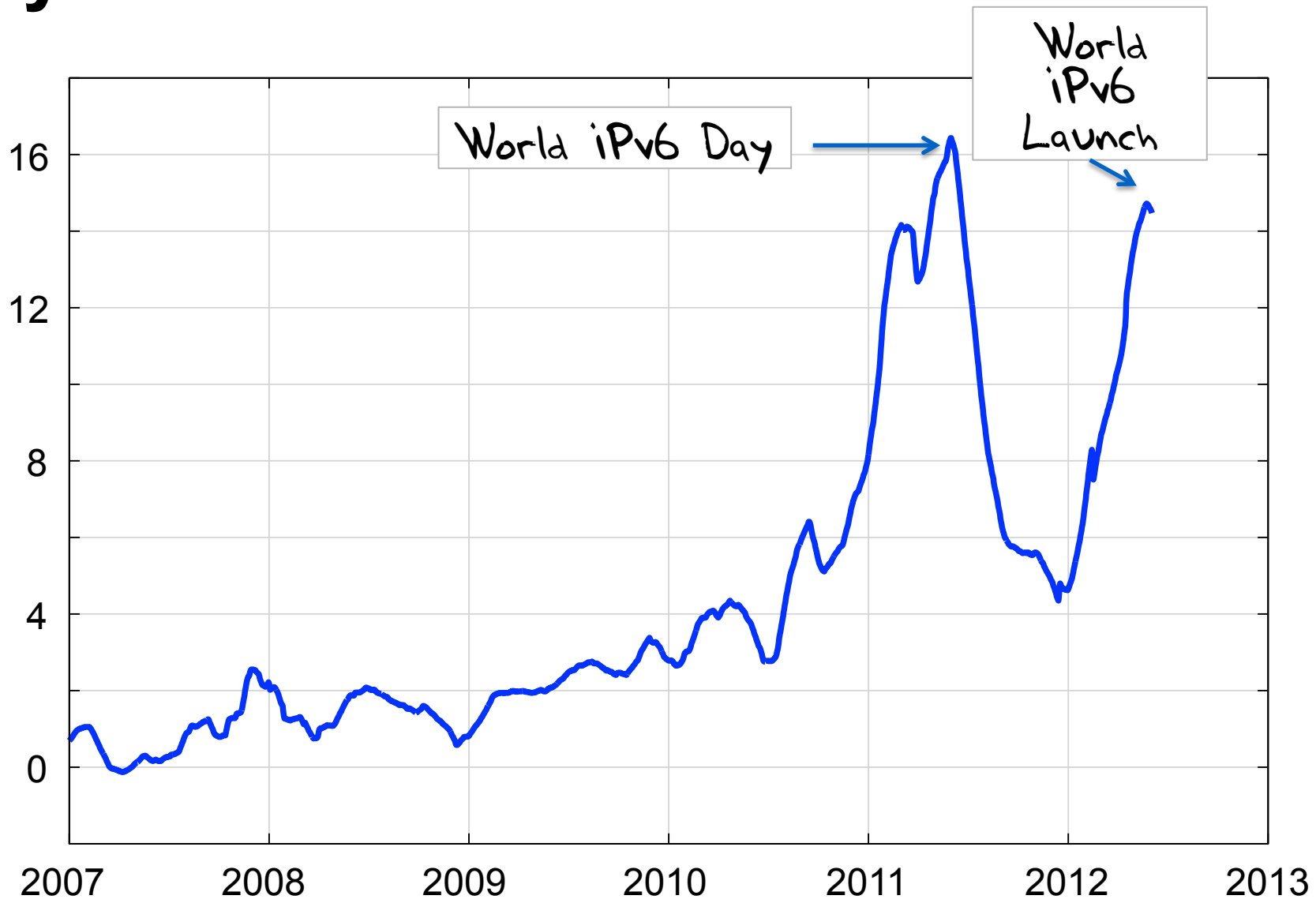
Jan 2011	347,000 entries
2012	390,000 entries
2013	431,000 entries
2014*	473,000 entries
2015*	517,000 entries
2016*	563,000 entries

* *These numbers are dubious due to uncertainties introduced by IPv4 address exhaustion pressures.*

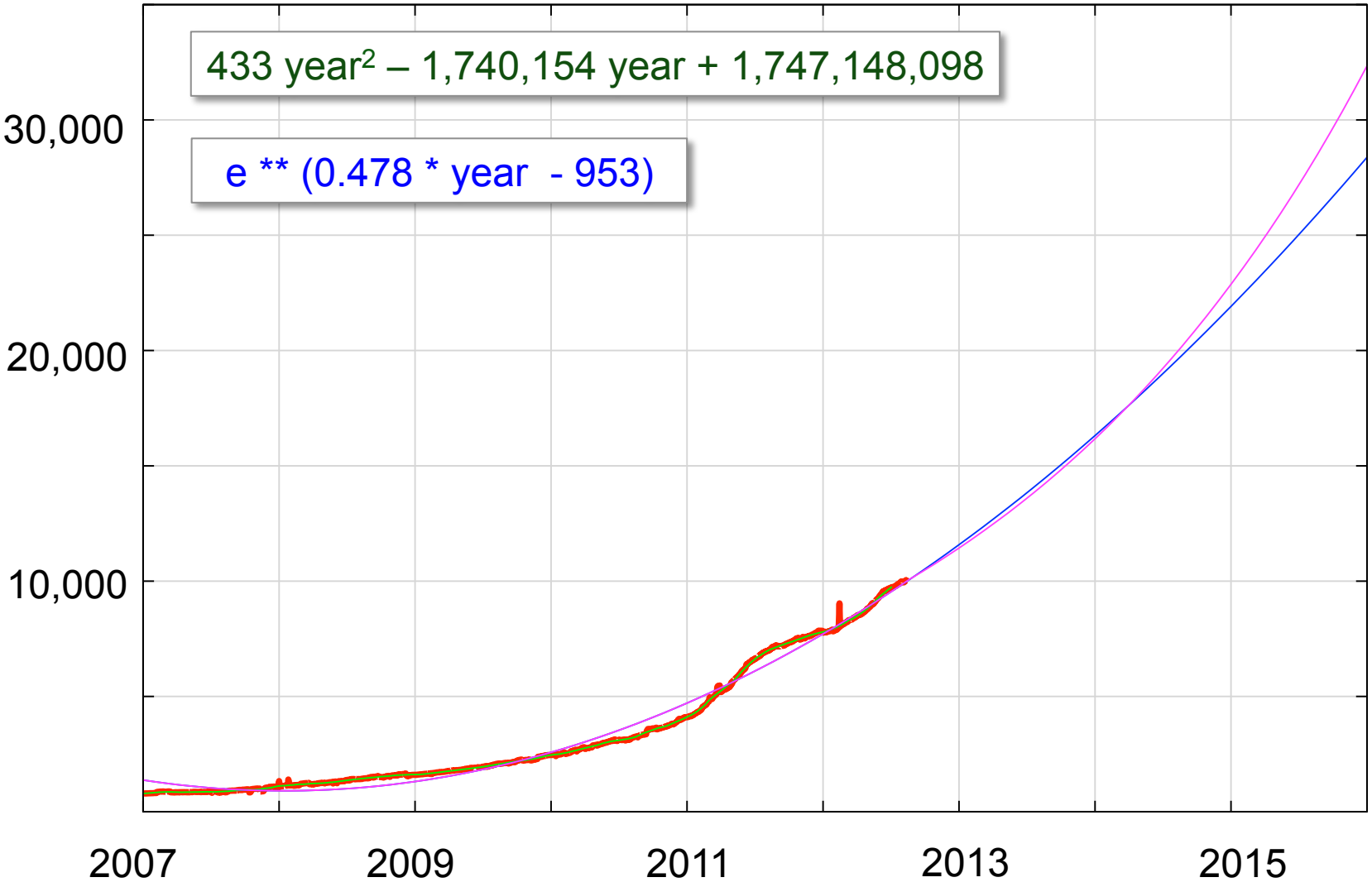
IPv6 Table Size



Daily Growth Rates



IPv6 Table Projection



IPv6 BGP Table Size predictions

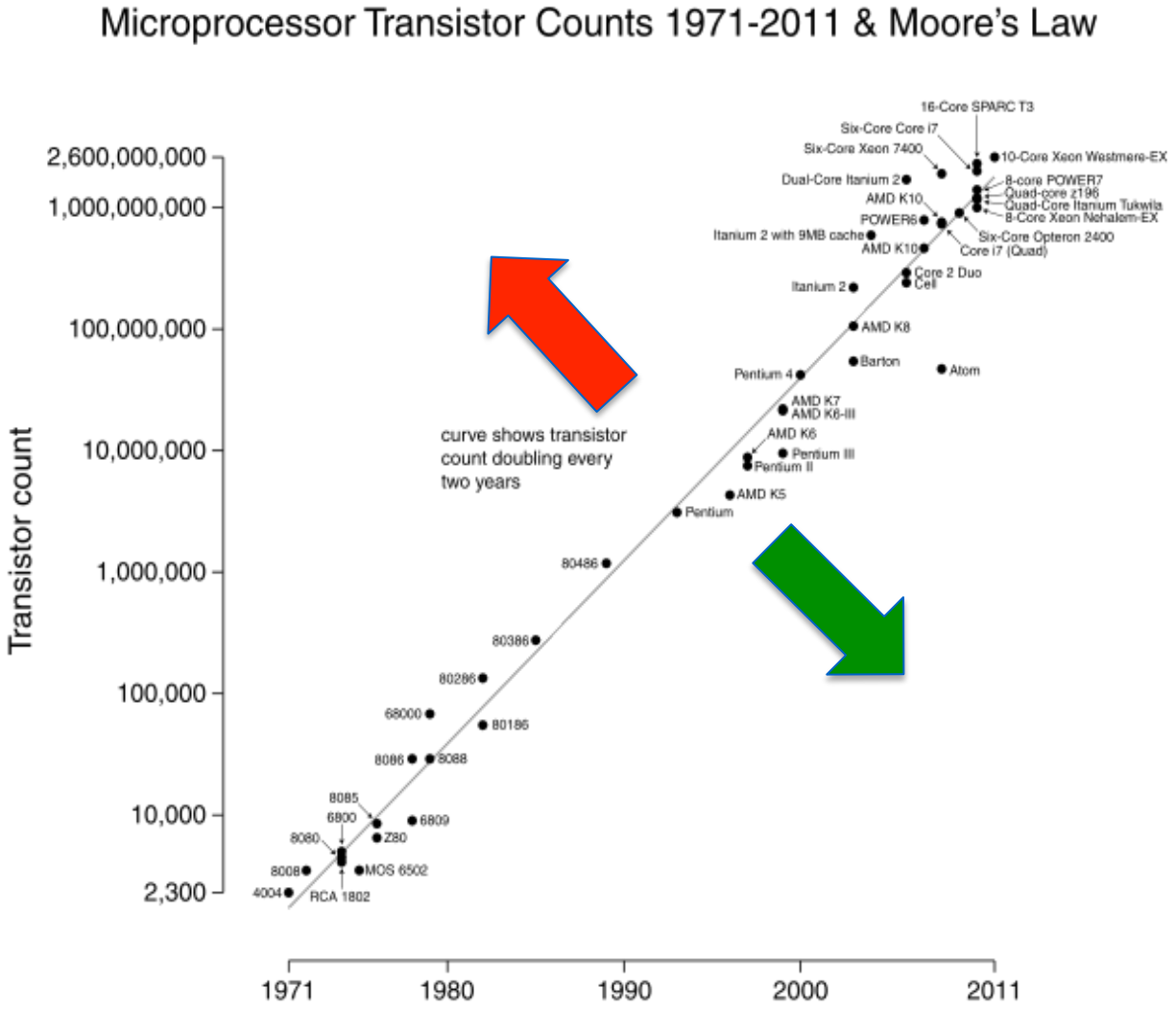
Jan 2011	4,000 entries
2012	8,000 entries
2013	11,600 entries
2014	16,300 entries
2015	21,900 entries
2016	28,300 entries

Up and to the Right

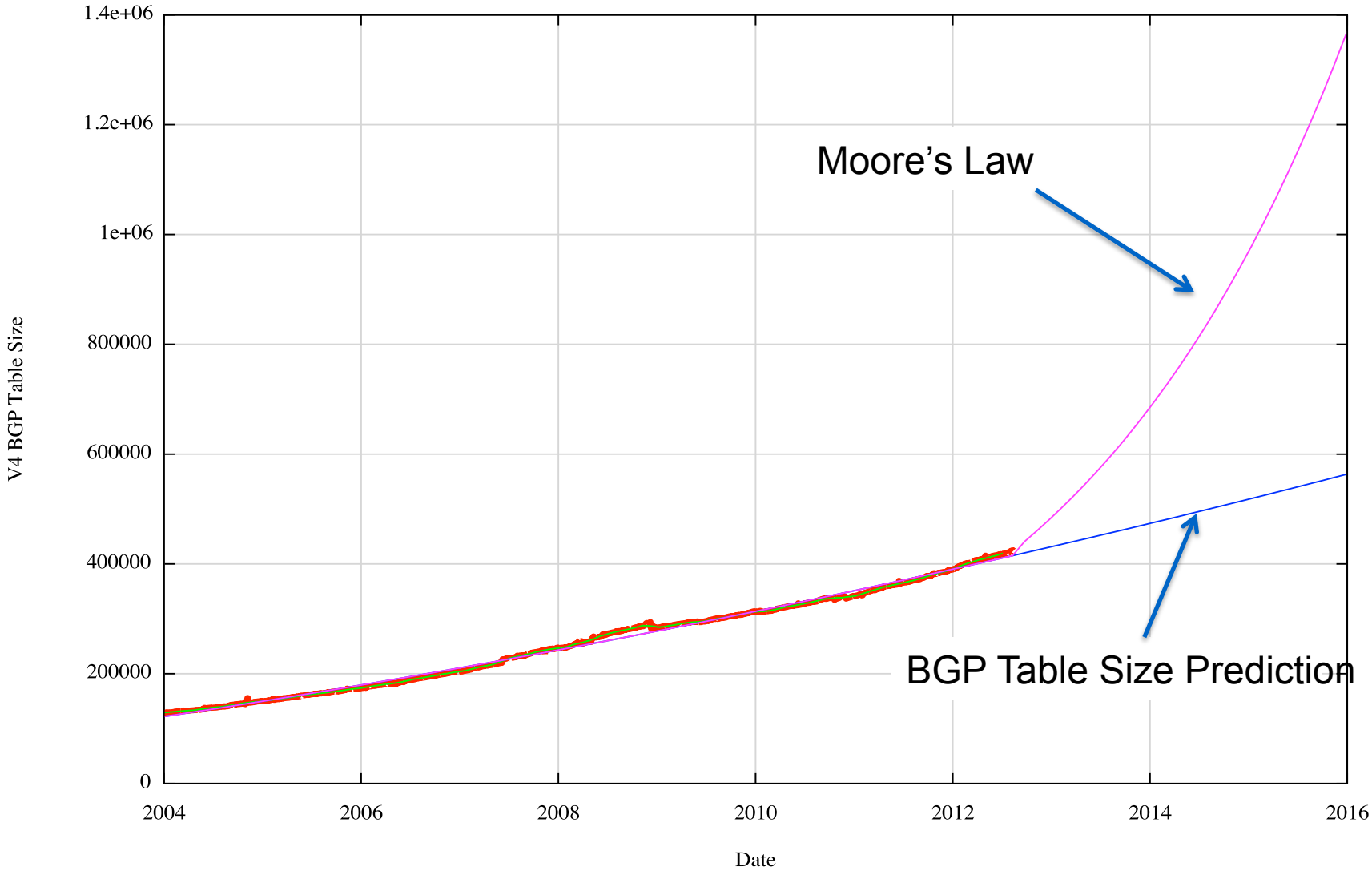
- Most Internet curves are “up and to the right”
- But what makes this curve painful?
  - The pain threshold is approximated by Moore’s Law

Moore's Law

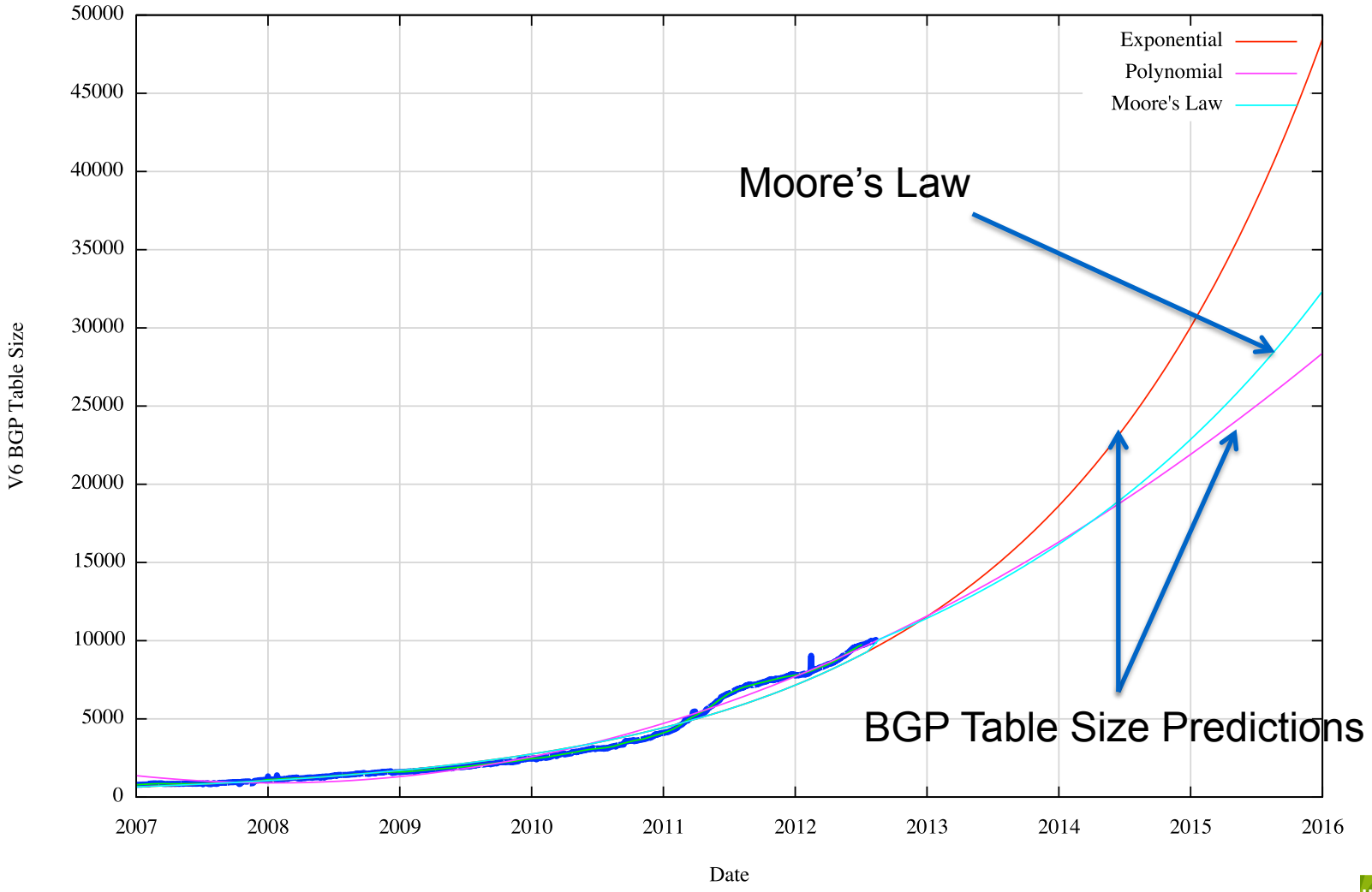
- As a rough rule of thumb, if the rate of growth of the table grows at a rate equal to, or less than Moore's Law, then the unit cost of storing the forwarding table should remain constant
  - Like all rough rules of thumb, there are many potential exceptions, and costs have many inputs as well as the raw cost of the the number of gates in a chip
  - Despite this, Moore's Law still a useful benchmark of a threshold of concern about routing growth


Microprocessor Transistor Counts 1971-2011 & Moore's Law
Transistor count
2,600,000,000
1,000,000,000
100,000,000
10,000,000
1,000,000
100,000
10,000
2,300
1971
1980
1990
2000
2011
Date of introduction
curve shows transistor count doubling every two years
16-Core SPARC T3
Six-Core Core i7
Six-Core Xeon 7400
10-Core Xeon Westmere-EX
Dual-Core Itanium 2
8-core POWER7
Quad-core z196
Quad-Core Itanium Tukwila
AMD K10
8-Core Xeon Nehalem-EX
POWER6
Six-Core Opteron 2400
Itanium 2 with 9MB cache
Core i7 (Quad)
AMD K10
Core 2 Duo
Cell
Itanium 2
AMD K8
Barton
Atom
Pentium 4
AMD K7
AMD K8-III
AMD K6
Pentium III
Pentium II
AMD K5
Pentium
80486
80386
80286
68000
80186
8086
8088
8085
6800
6809
Z80
8080
8008
MOS 6502
4004
RCA 1802

IPv4 BGP Table size and Moore's Law



IPv6 Projections and Moore's Law



eBGP Table Growth

- Nothing in these figures suggests that there is cause for urgent alarm -- at present
- The overall eBGP growth rates for IPv4 are holding at a modest level, and the IPv6 table, although it is growing rapidly, is still relatively small in size in absolute terms
- As long as we are prepared to live within the technical constraints of the current routing paradigm it will continue to be viable for some time yet

NATTing the Net

In the previous 12 months:

- The RIRs allocated 109M IPv4 addresses
- The routing table grew by 120M addresses
- The ISC host survey* indicates a growth of ~90M visible hosts
- BUT
  - Apple sold ~100M iphones and they have 24% market share
  - This implies that some ~400M mobile devices were deployed in the last 12 months
  - And that does not include Androids, Kindles, iPads, etc
- It appears that the NATTed Internet grew by ~600M devices in the last 12 months!

** http://www.isc.org/solutions/survey*

Aggregation

- To what extent do we still practice “conservative” routing and refrain from announcing more specifics into the routing table?
- Are we getting better or worse at aggregation in routing?
- What is the distribution of advertising more specifics? Are we seeing a significant increase in the number of more specific /24s in the routing table?

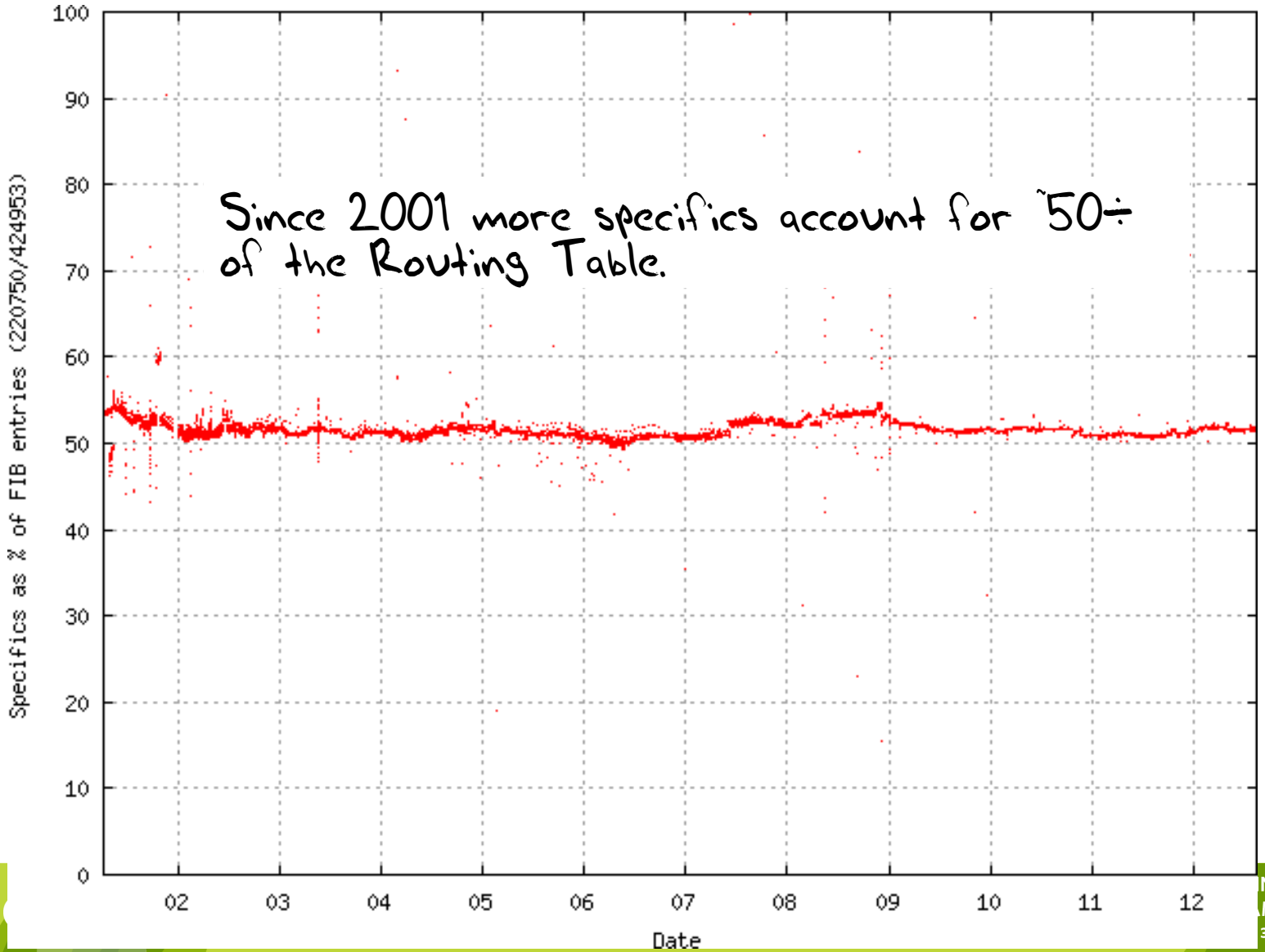
Who is doing this the most?

www.cidr-report.org

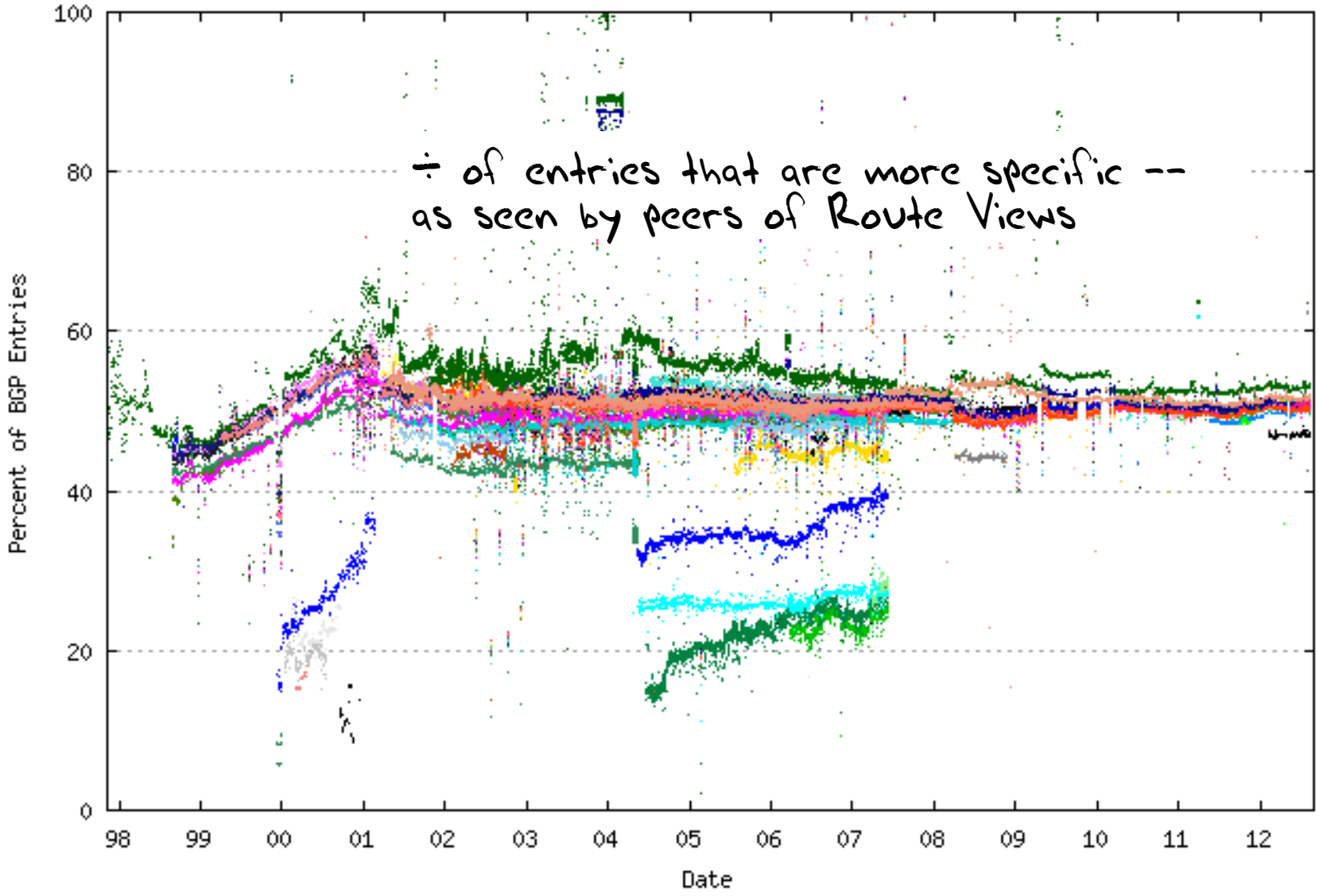
--- 14Aug12 ---

ASnum	NetsNow	NetsAggr	NetGain	% Gain	Description
Table	424954	244536	180418	42.5%	All ASes
AS6389	3374	195	3179	94.2%	BELLSOUTH-NET-BLK - BellSouth.net Inc.
AS28573	2042	63	1979	96.9%	NET Servicos de Comunicao S.A.
AS17974	2342	477	1865	79.6%	TELKOMNET-AS2-AP PT Telekomunikasi Indonesia
AS7029	3441	1770	1671	48.6%	WINDSTREAM - Windstream Communications Inc
AS18566	2088	417	1671	80.0%	COVAD - Covad Communications Co.
AS22773	1723	160	1563	90.7%	ASN-CXA-ALL-CCI-22773-RDC - Cox Communications Inc.
AS4766	2766	1294	1472	53.2%	KIXS-AS-KR Korea Telecom
AS10620	2077	761	1316	63.4%	Telmex Colombia S.A.
AS4323	1576	389	1187	75.3%	TWTC - tw telecom holdings, inc.
AS1785	1939	819	1120	57.8%	AS-PAETEC-NET - PaeTec Communications, Inc.
AS7303	1561	449	1112	71.2%	Telecom Argentina S.A.
AS4755	1605	543	1062	66.2%	TATACOMM-AS TATA Communications formerly VSNL is Leading ISP
AS7552	1110	149	961	86.6%	VIETEL-AS-AP Vietel Corporation
AS6458	882	42	840	95.2%	Telgua
AS8151	1470	667	803	54.6%	Uninet S.A. de C.V.
AS18101	941	159	782	83.1%	RELIANCE-COMMUNICATIONS-IN Reliance Communications Ltd.DAKC MUMBAI
AS4808	1123	354	769	68.5%	CHINA169-BJ CNCGROUP IP network China169 Beijing Province Network
AS9394	896	174	722	80.6%	CRNET CHINA RAILWAY Internet(CRNET)
AS13977	840	124	716	85.2%	CTELCO - FAIRPOINT COMMUNICATIONS, INC.
AS2118	685	14	671	98.0%	RELCOM-AS OOO "NPO Relcom"
AS15557	1215	552	663	54.6%	LDCOMNET Societe Francaise du Radiotelephone S.A
AS855	684	52	632	92.4%	CANET-ASN-4 - Bell Aliant Regional Communications, Inc.
AS3356	1102	472	630	57.2%	LEVEL3 Level 3 Communications
AS17676	703	79	624	88.8%	GIGAINFRA Softbank BB Corp.
AS22561	1030	431	599	58.2%	DIGITAL-TELEPORT - Digital Teleport Inc.
AS19262	1001	403	598	59.7%	VZGNI-TRANSIT - Verizon Online LLC
AS30036	1448	856	592	40.9%	MEDIACOM-ENTERPRISE-BUSINESS - Mediacom Communications Corp
AS24560	1039	448	591	56.9%	AIRTELBROADBAND-AS-AP Bharti Airtel Ltd., Telemedia Services
AS3549	1002	441	561	56.0%	GBLX Global Crossing Ltd.
AS4780	811	258	553	68.2%	SEEDNET Digital United Inc.
Total	44516	13012	31504	70.8%	Top 30 total

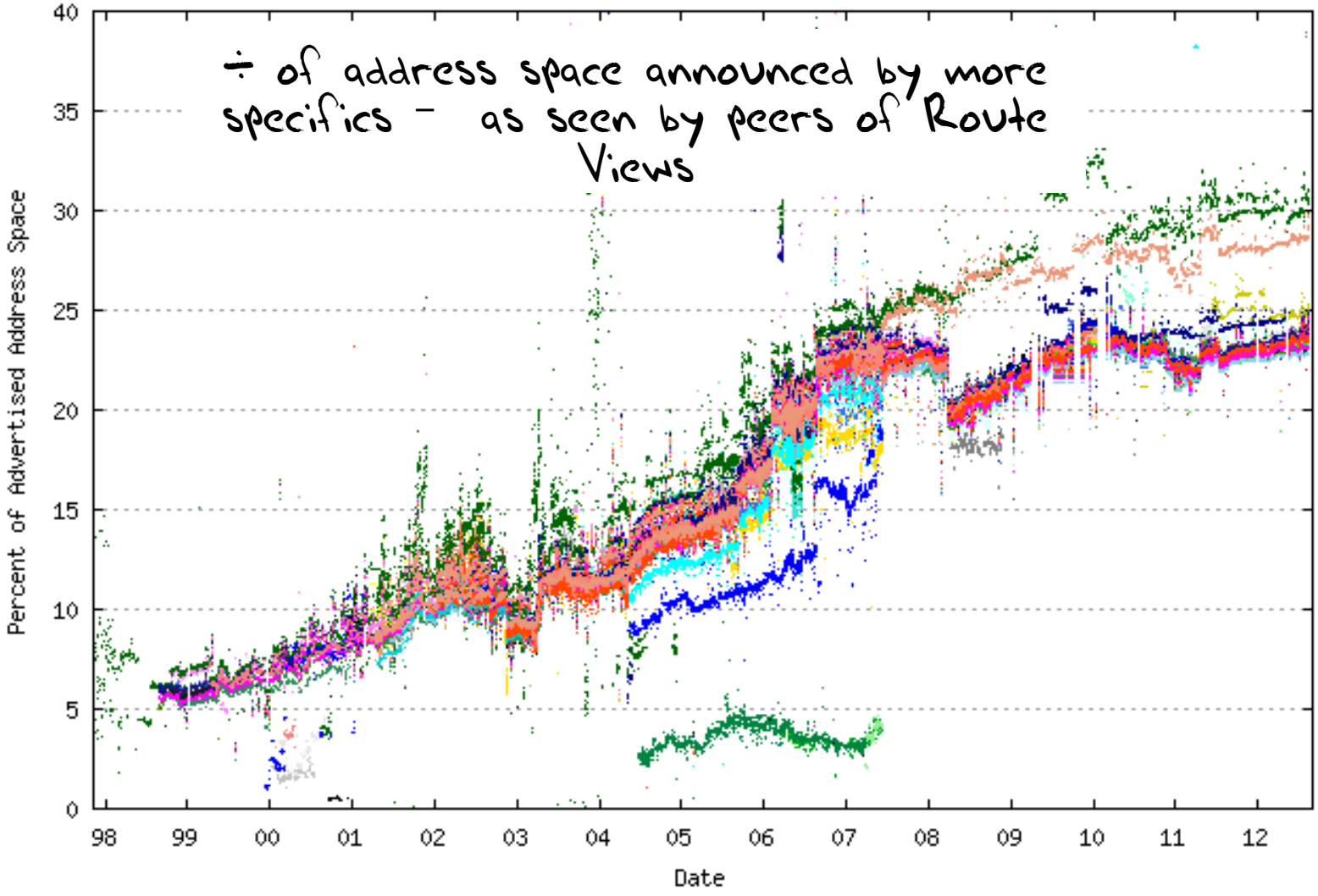
More specifics in the Routing Table



Does everyone see this?



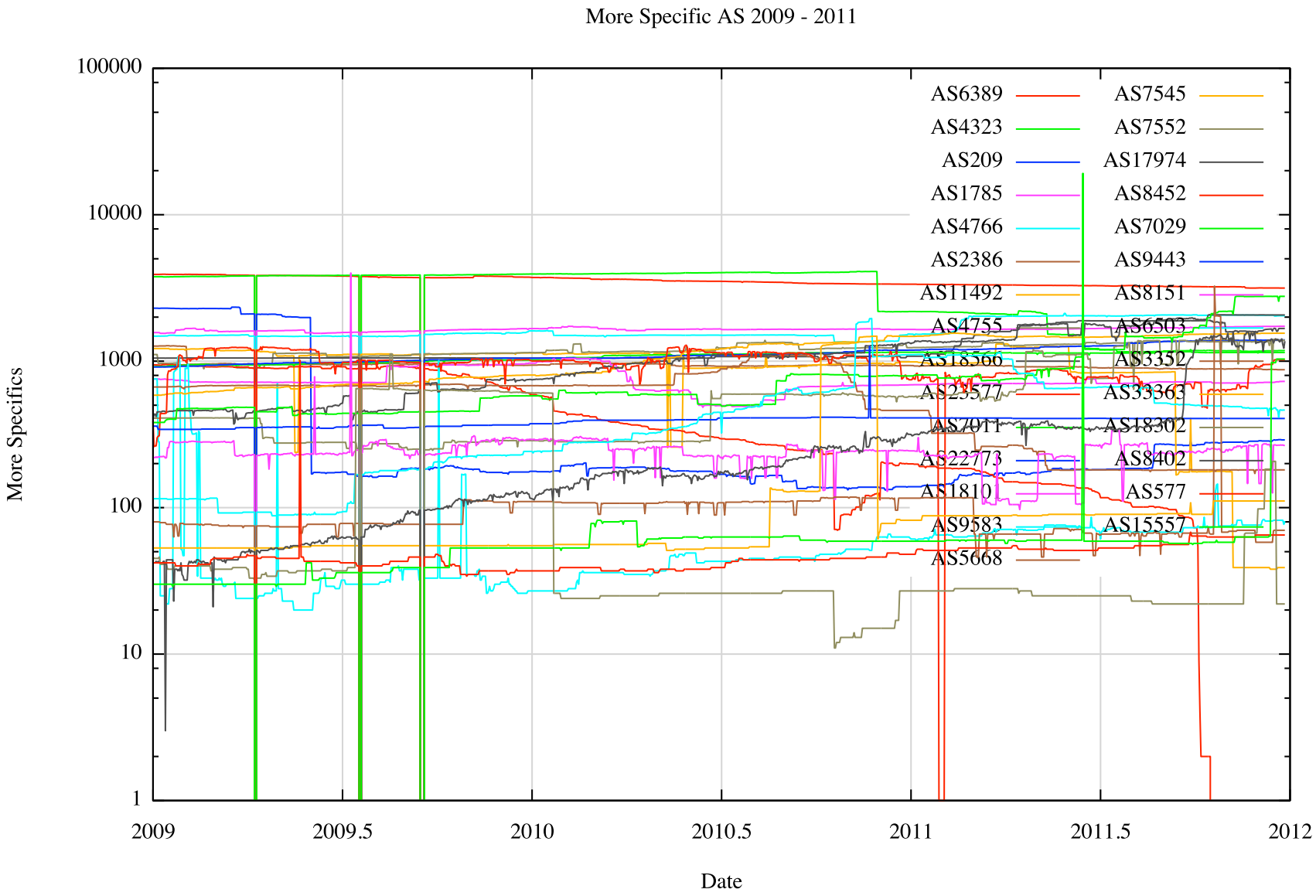
How much address space is announced by more specifics?



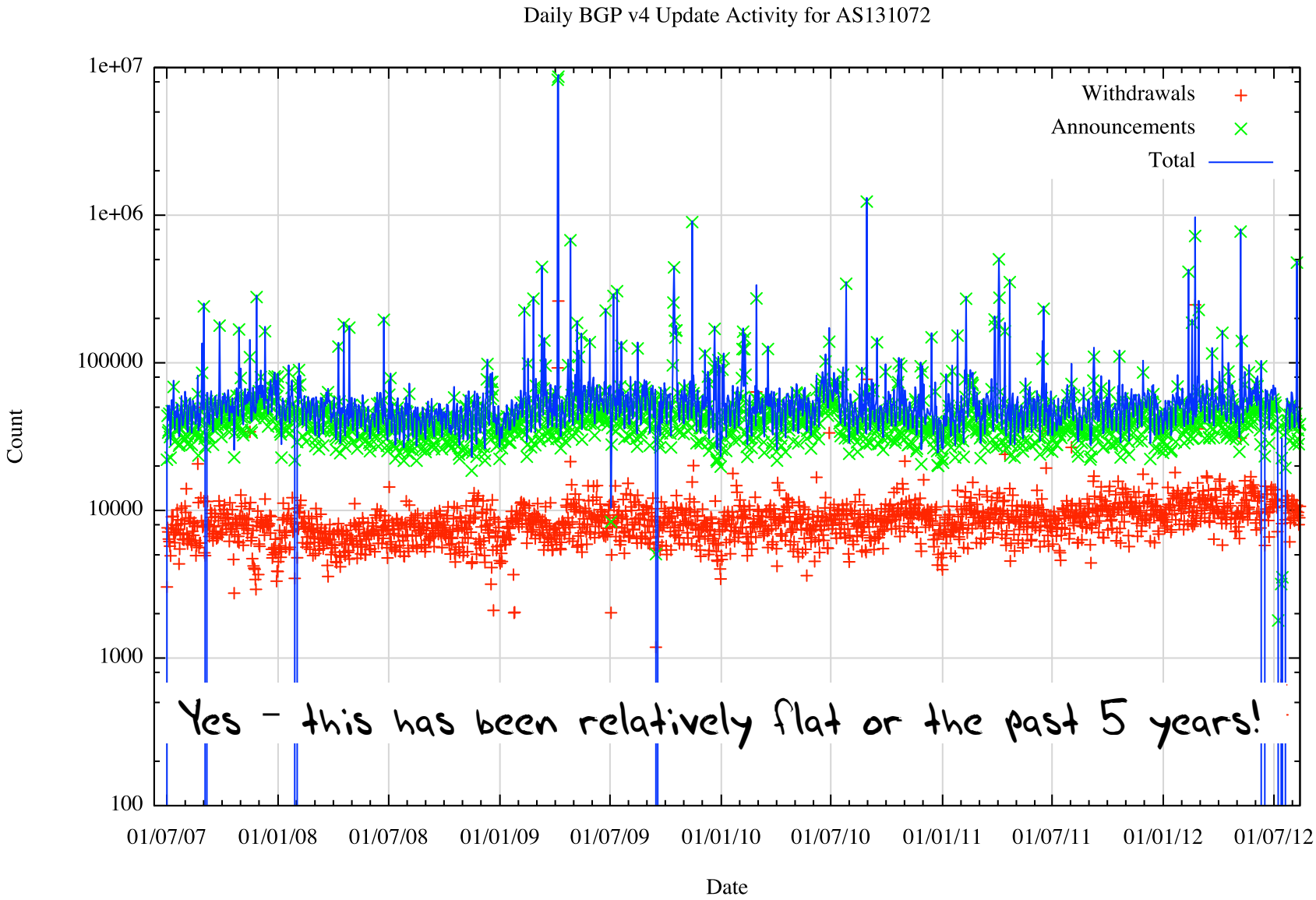
Are We Getting Any Better?

- Take the daily top 10 ASes over the past 3 years and track the number of more specifics advertised by these ASes over the entire period

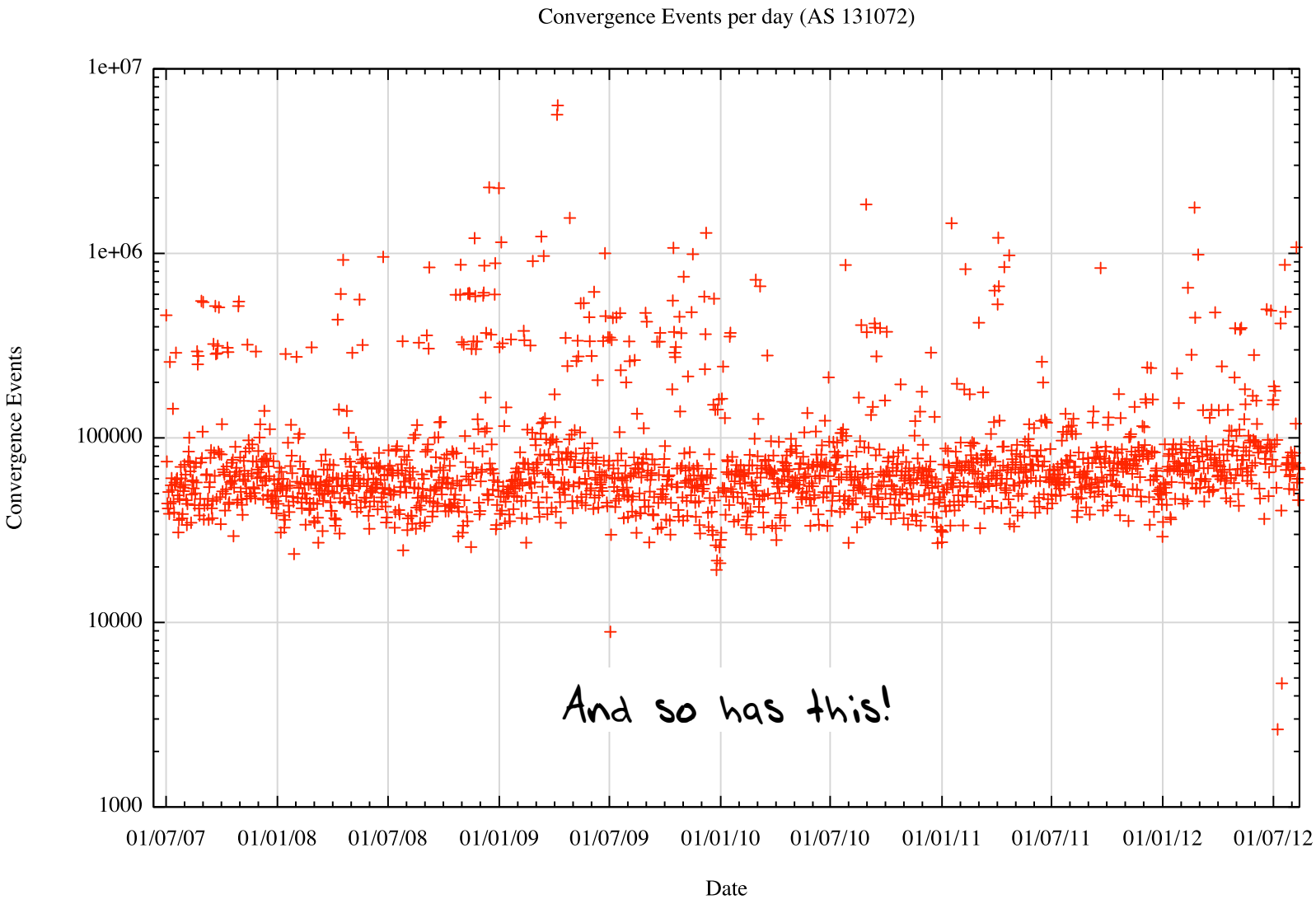
Yes ... and No


More Specific AS 2009 - 2011
More Specifics
Date
100000
10000
1000
100
10
1
2009
2009.5
2010
2010.5
2011
2011.5
2012
AS6389
AS4323
AS209
AS1785
AS4766
AS2386
AS11492
AS4755
AS18566
AS23577
AS7911
AS22773
AS1810
AS9583
AS5668
AS7545
AS7552
AS17974
AS8452
AS7029
AS9443
AS8151
AS6503
AS3352
AS33363
AS18302
AS8402
AS577
AS15557

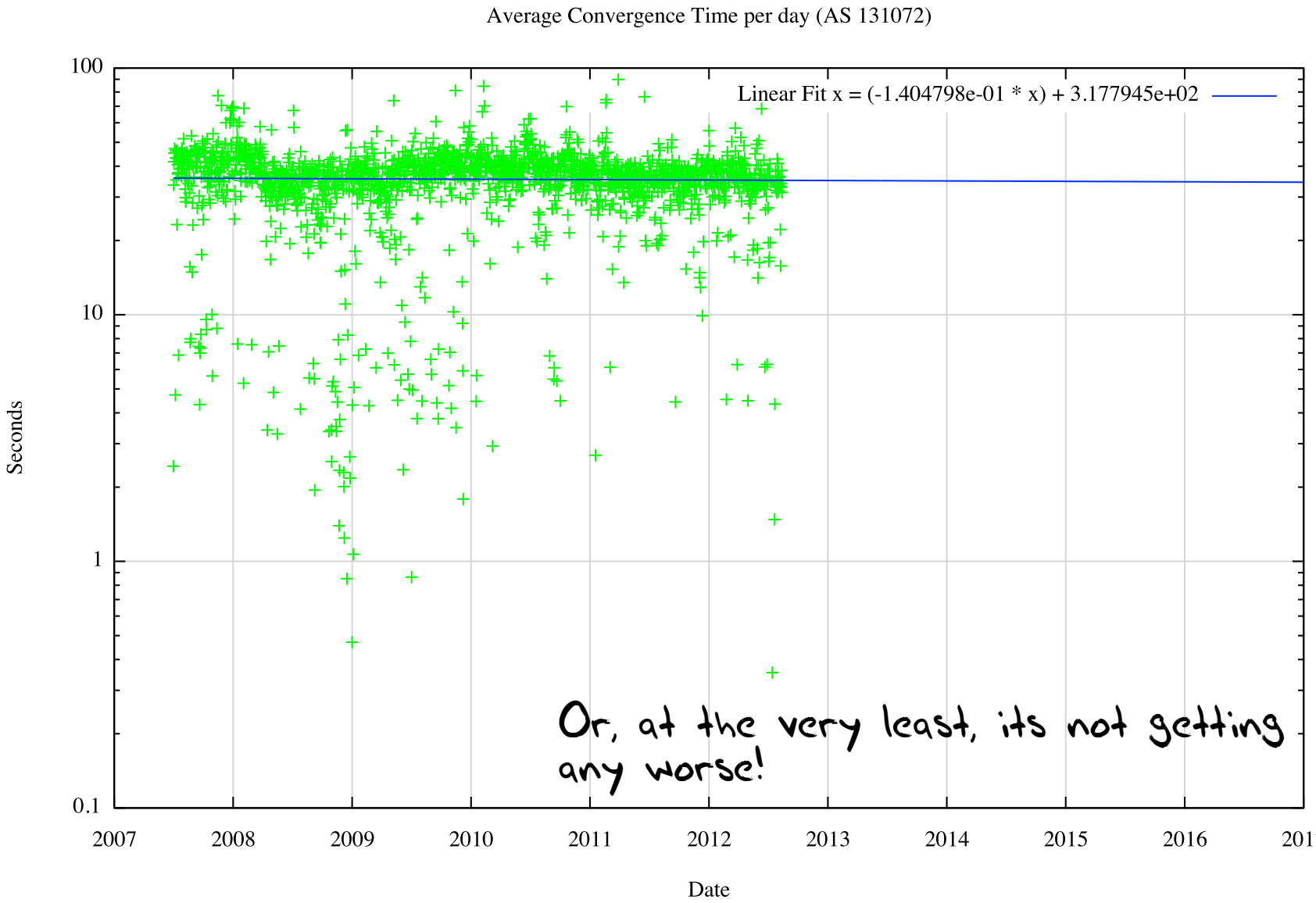
Is Routing still Scaling?



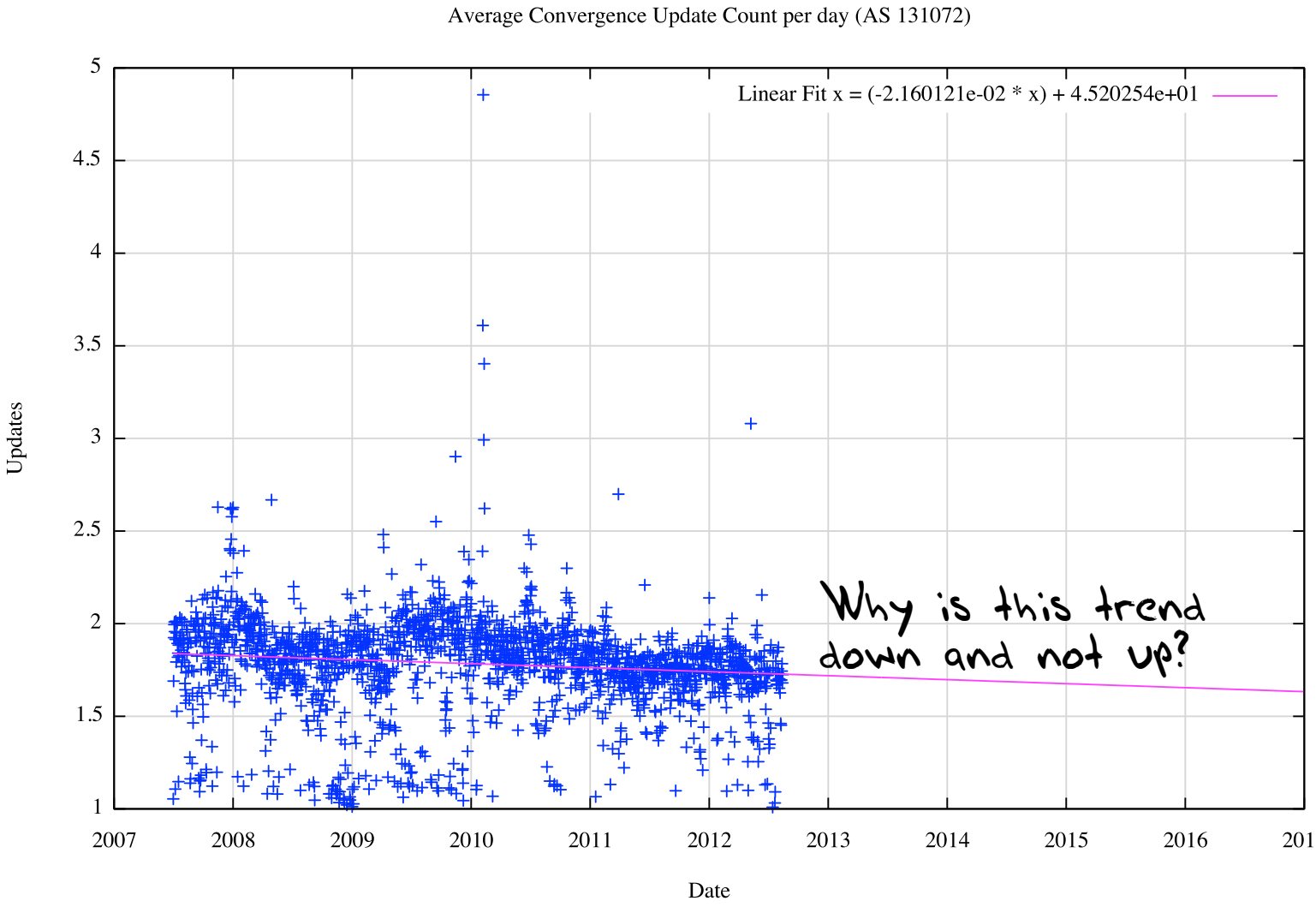
Instability Events per day



Convergence is improving!



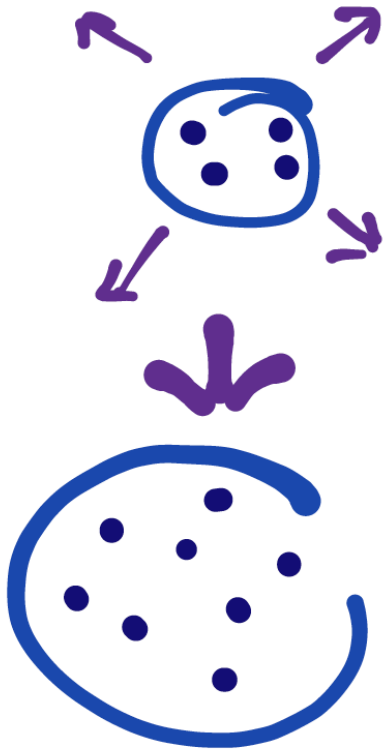
Convergence is improving!



What is going on?

- The “shape” of the Internet determines the scalability of the Internet’s routing system

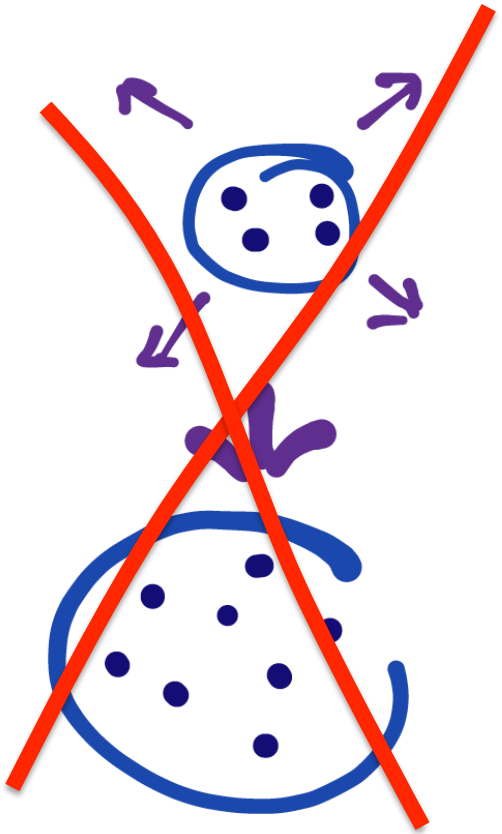
if the internet “expands” then the AS paths lengthen. This means that a distance vector algorithm will take longer to converge, and send more updates in the process



What is going on?

- The “shape” of the Internet determines the scalability of the Internet’s routing system

if the internet “expands” then the AS paths lengthen. This means that a distance vector algorithm will take longer to converge, and send more updates in the process – this is NOT the case for the internet



What is going on?

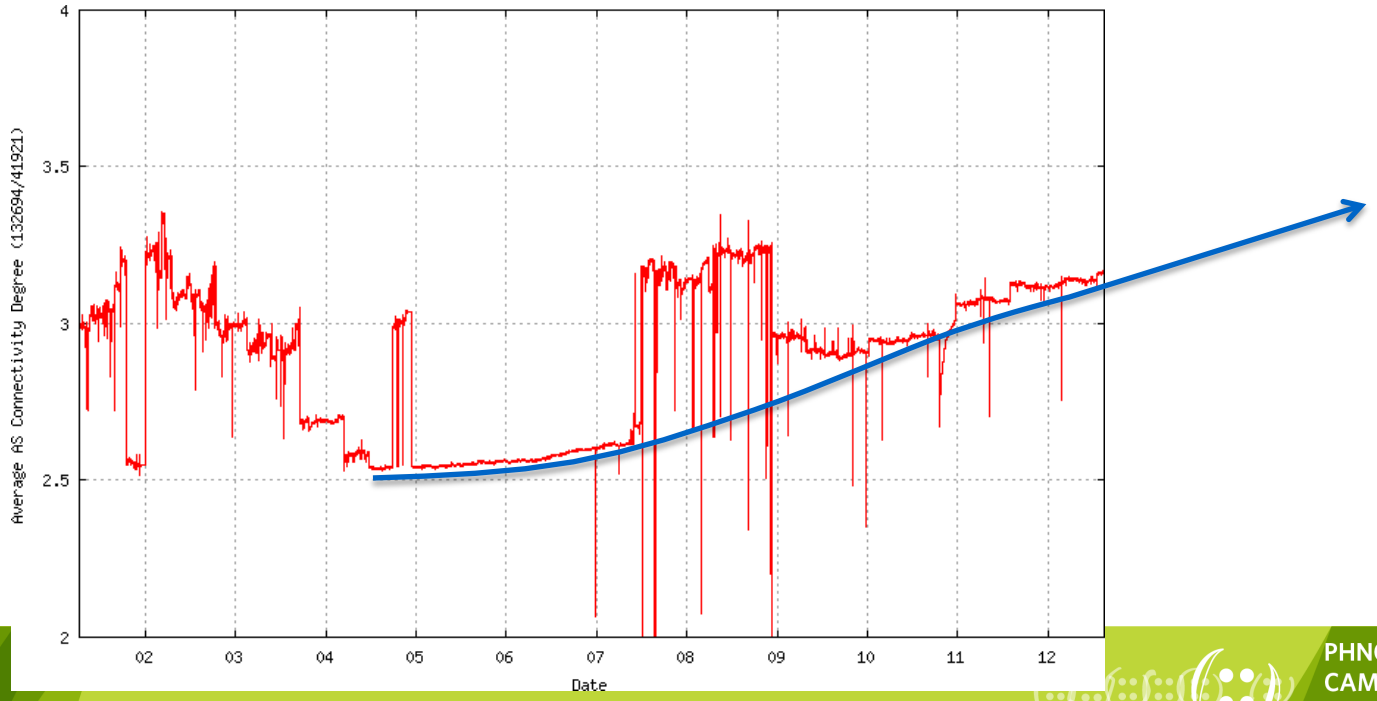
- The “shape” of the Internet determines the scalability of the Internet’s routing system

if the internet “compresses” then the AS paths length remain steady. This means that a distance vector algorithm will behave consistently over time. This compression is achieved by “clustering” of new entrants around existing AS’s, increasing the connectivity density of the larger AS’s

The Shape of the Internet

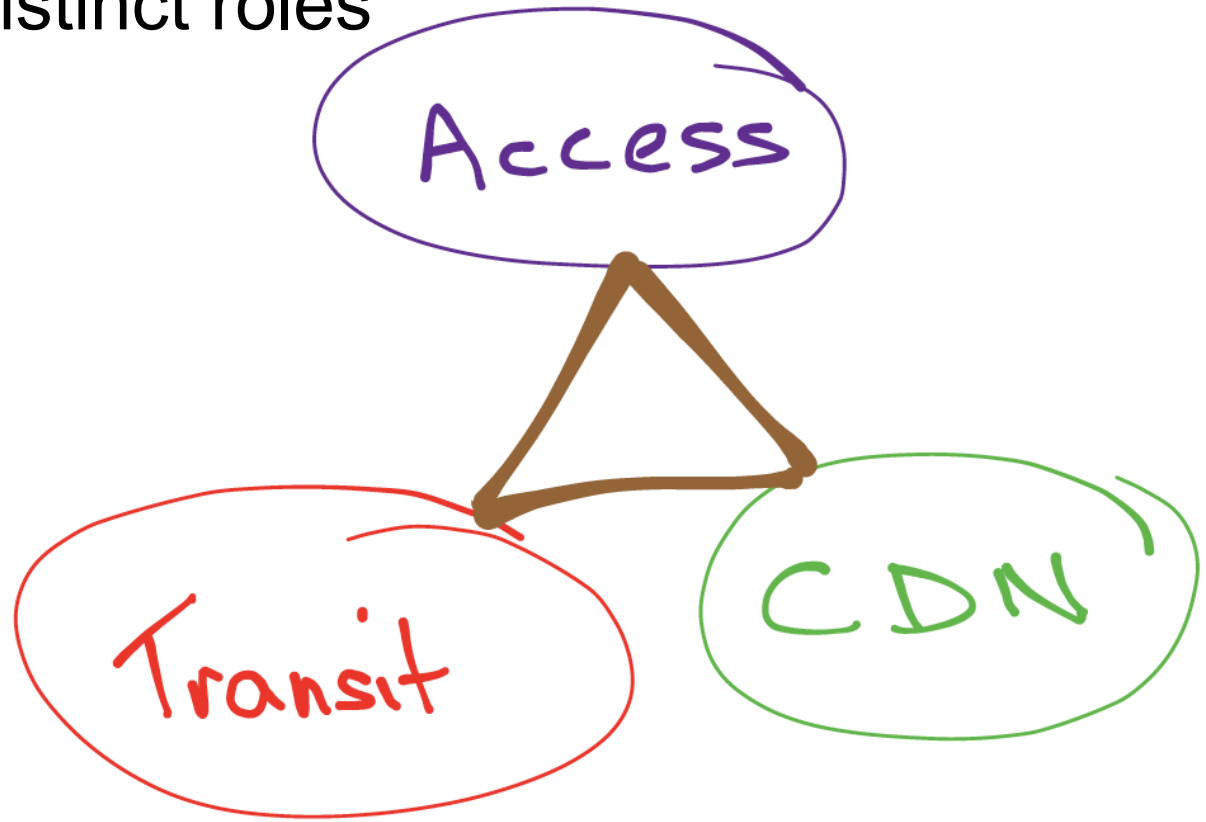
The Internet's carriage transit environment is becoming less diverse as the transit tiers are being compressed – AS connectivity is becoming denser

Average AS Connectivity



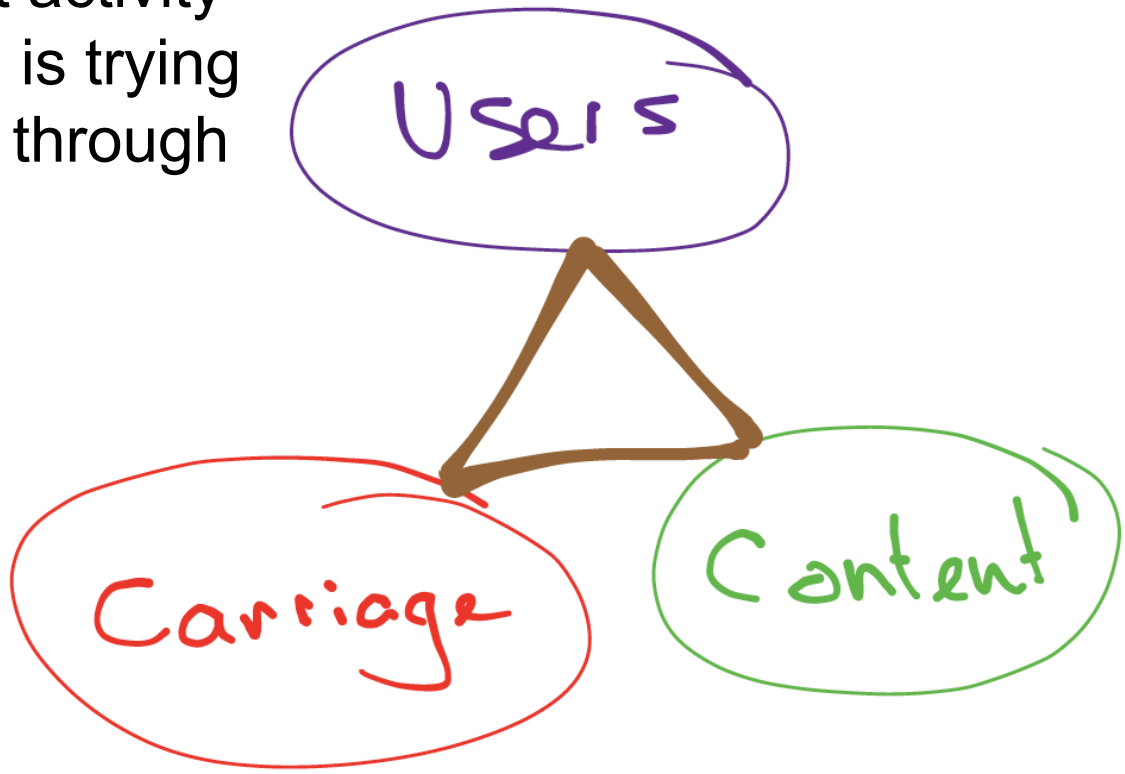
The Shape of the Internet

AS's are no longer all the same - they are now specialized into distinct roles



The Shape of the Internet

This AS specialization mirrors the underlying segmentation of the market into distinct activity sectors, each of which is trying to maximize efficiency through economies of scale



Thank You

Questions?