Scaling BGP

Geoff Huston
APNIC

May 2016

Through the Routing Lens

Through the Routing Lens

There are very few ways to assemble a single view of the entire Internet

The lens of routing is one of the ways in which information relating to the entire reachable Internet is bought together

Even so, its not a perfect lens…

There is no Routing God!

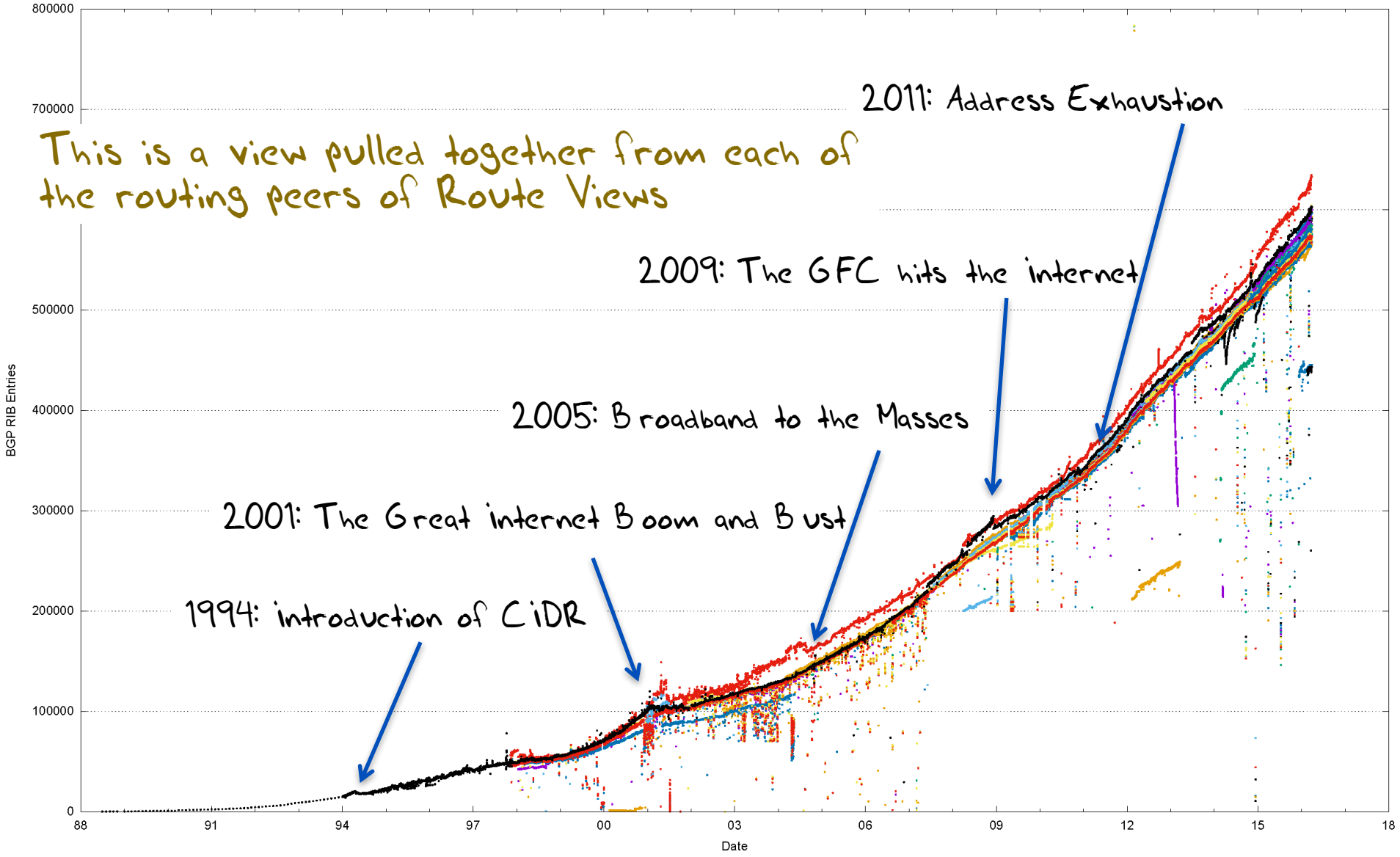
There is no single objective “out of the system” view of the Internet’s Routing environment.

BGP distributes a routing view that is modified as it is distributed, so every eBGP speaker will see a slightly different set of prefixes, and each view is relative to a given location

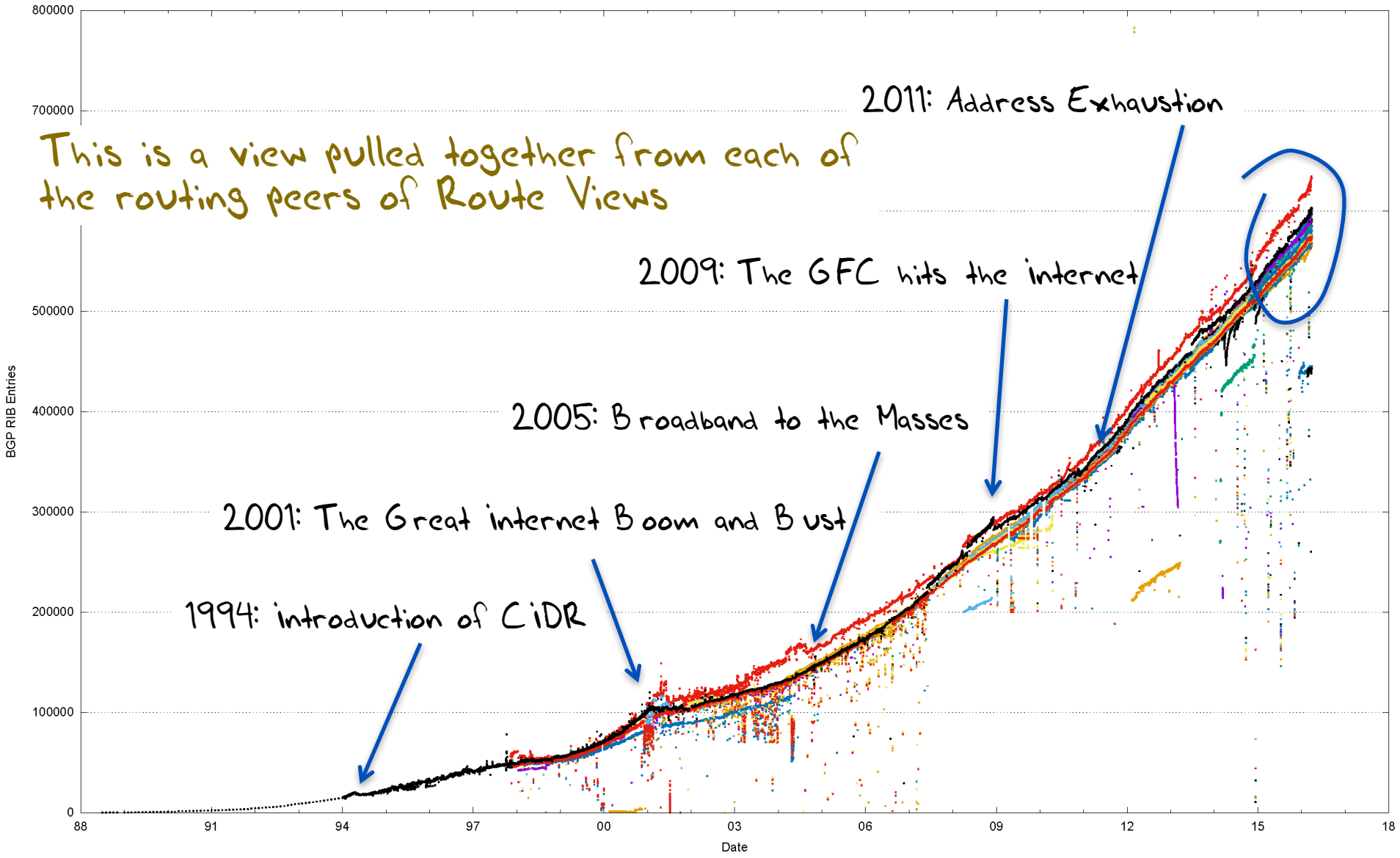
So the picture I will be painting here is one that is drawn from the perspective of AS131072. This is a stub AS at edge of the Internet, and this is an eBGP view.

You may have a similar view from your network.

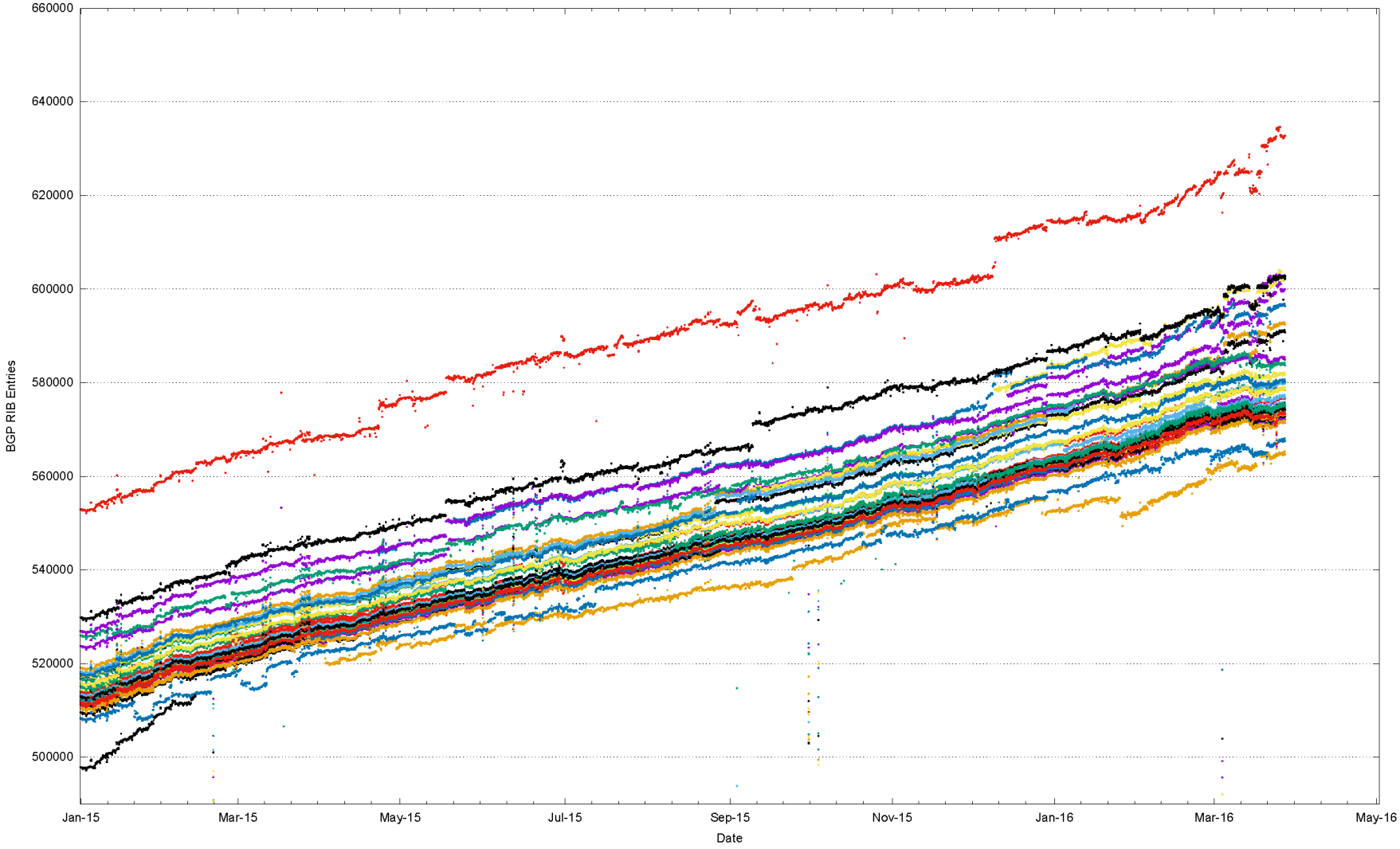
28 Years of Routing the Internet



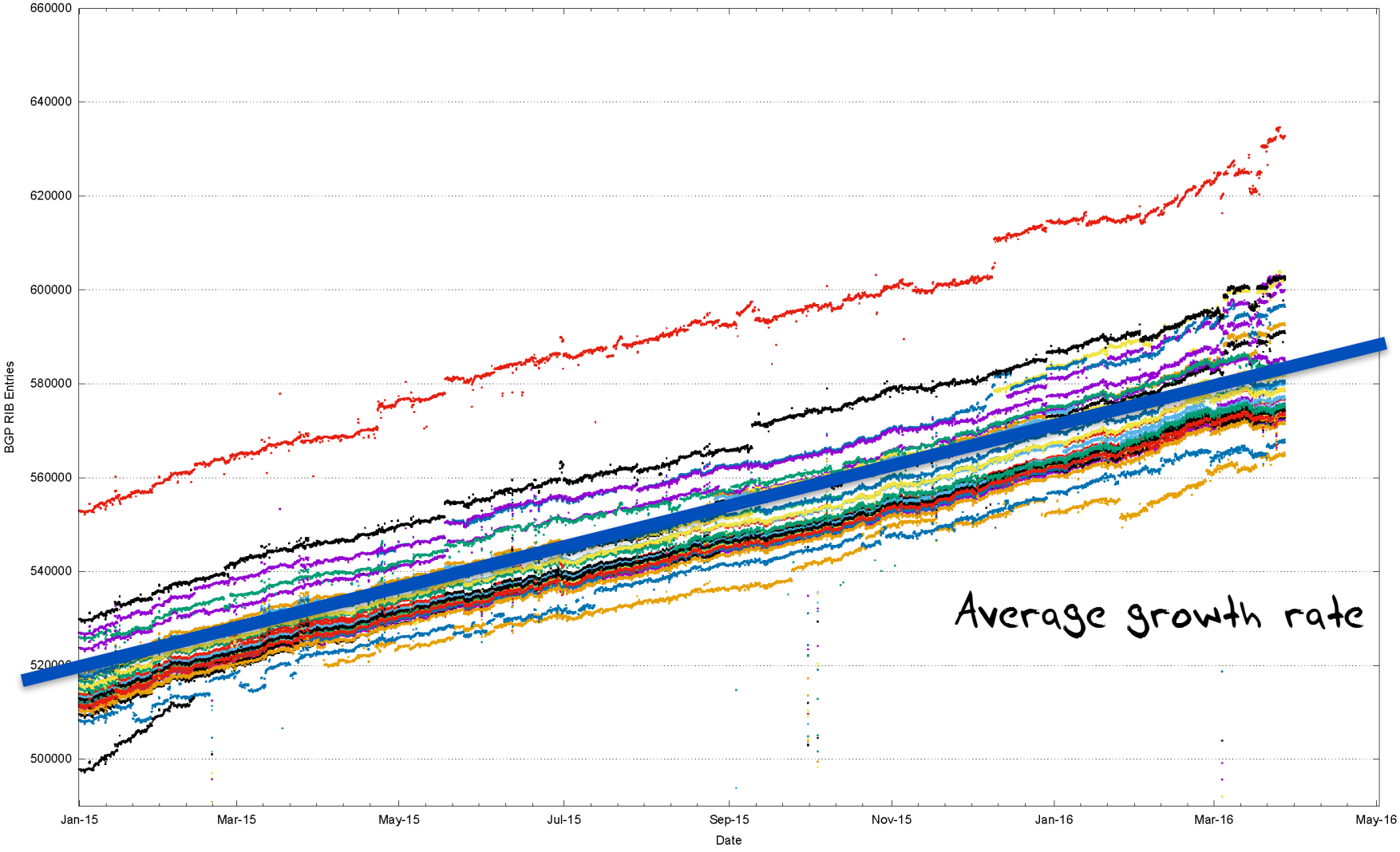
28 Years of Routing the Internet



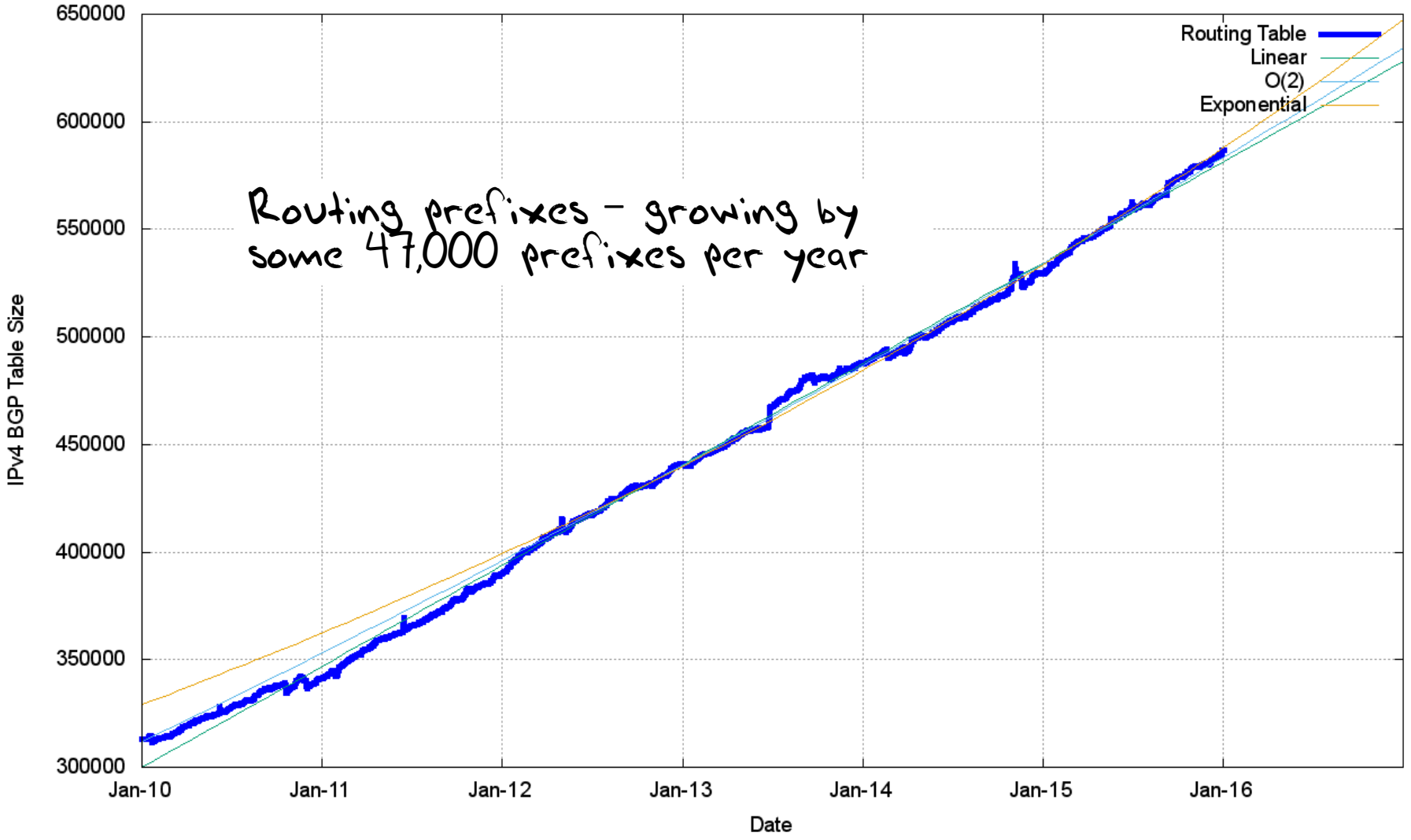
2015, as seen at Route Views



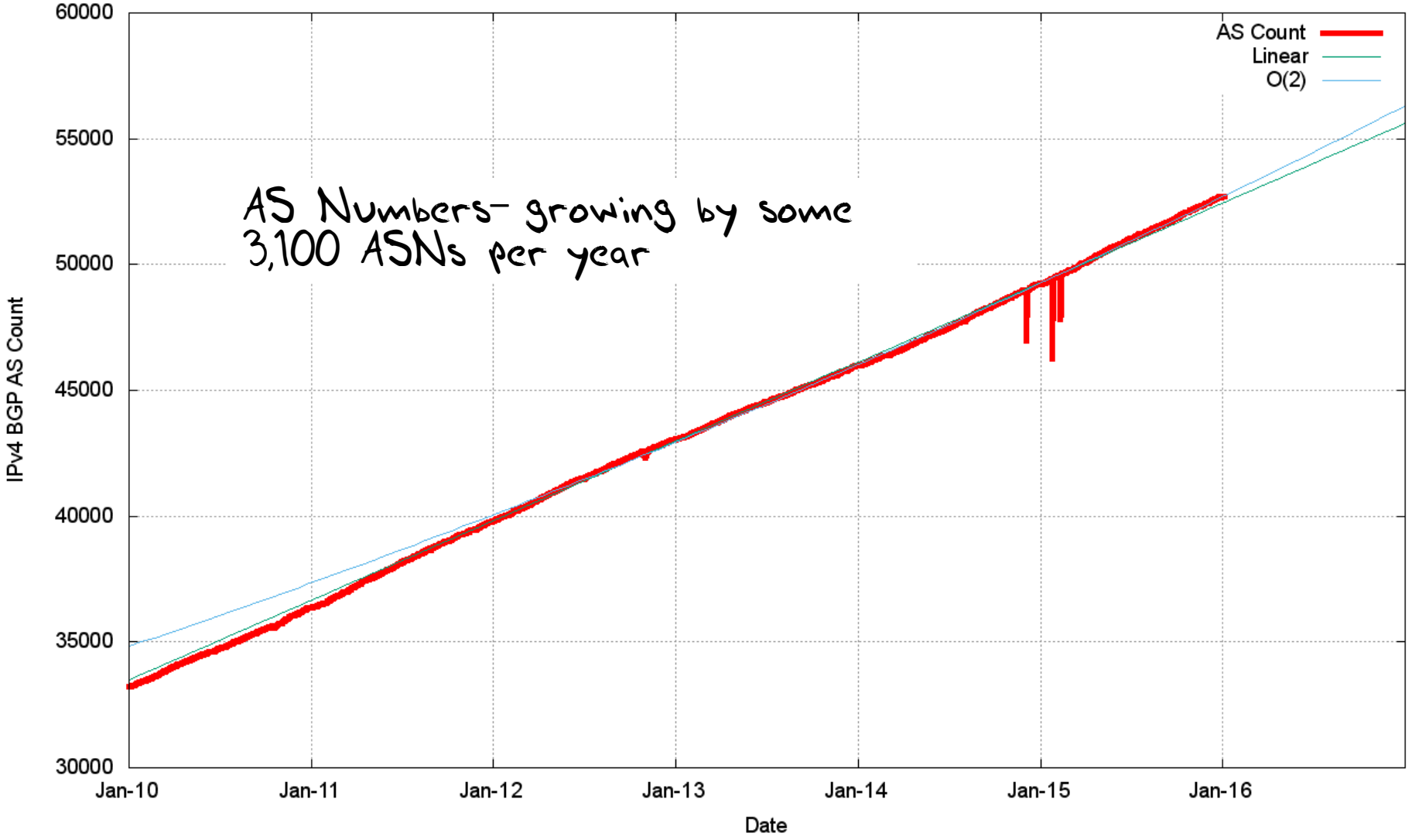
2015, as seen at Route Views



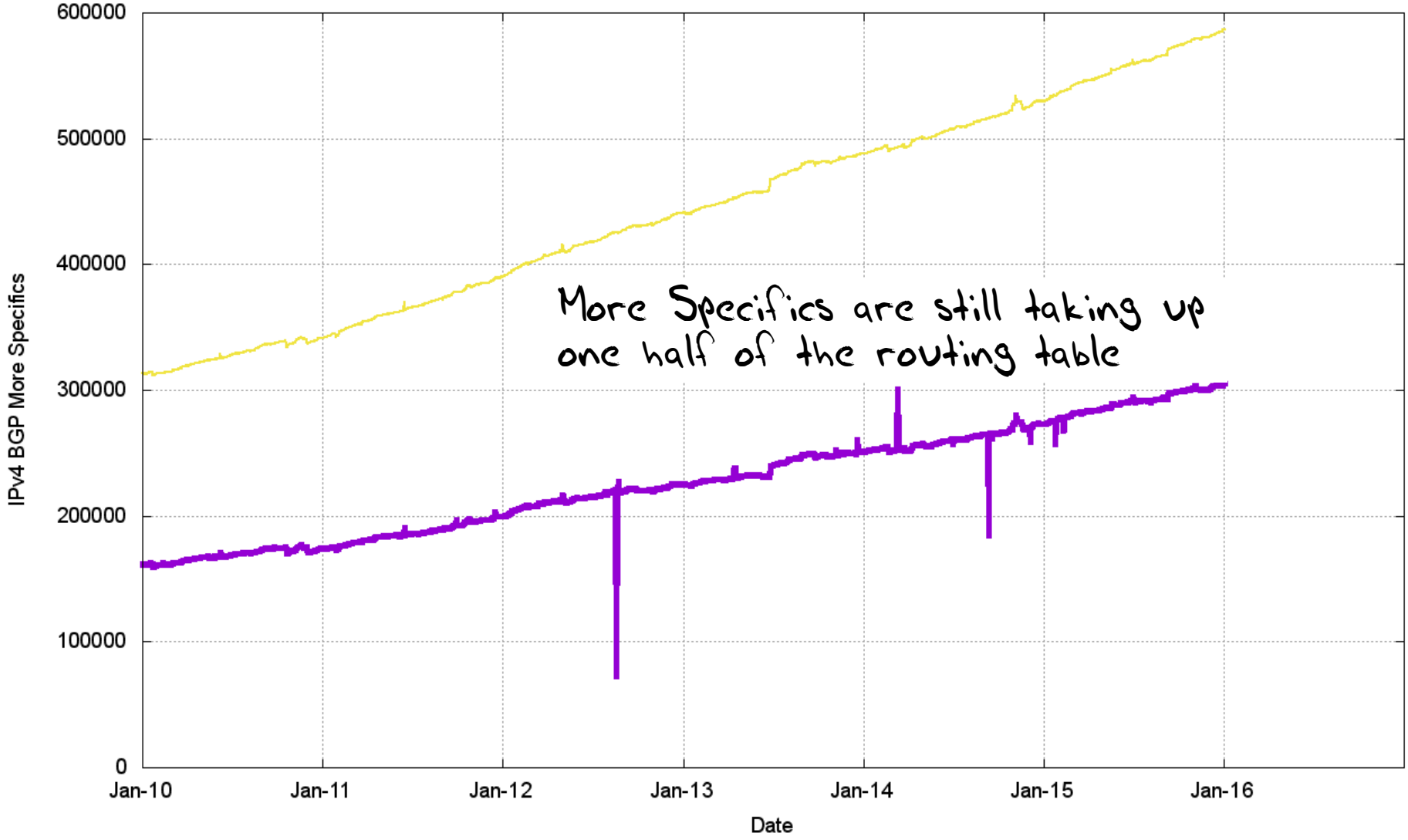
Routing Indicators for IPv4



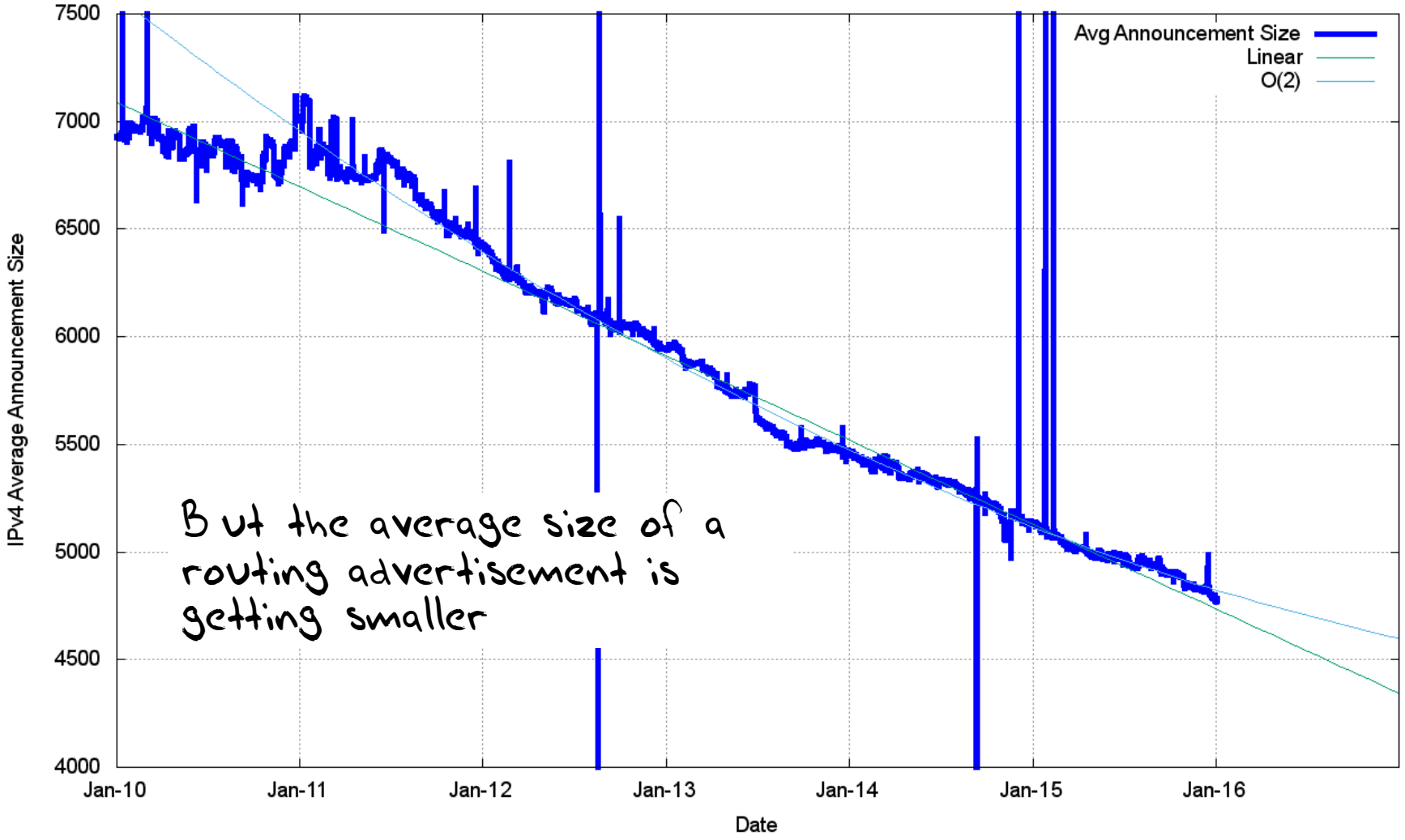
Routing Indicators for IPv4



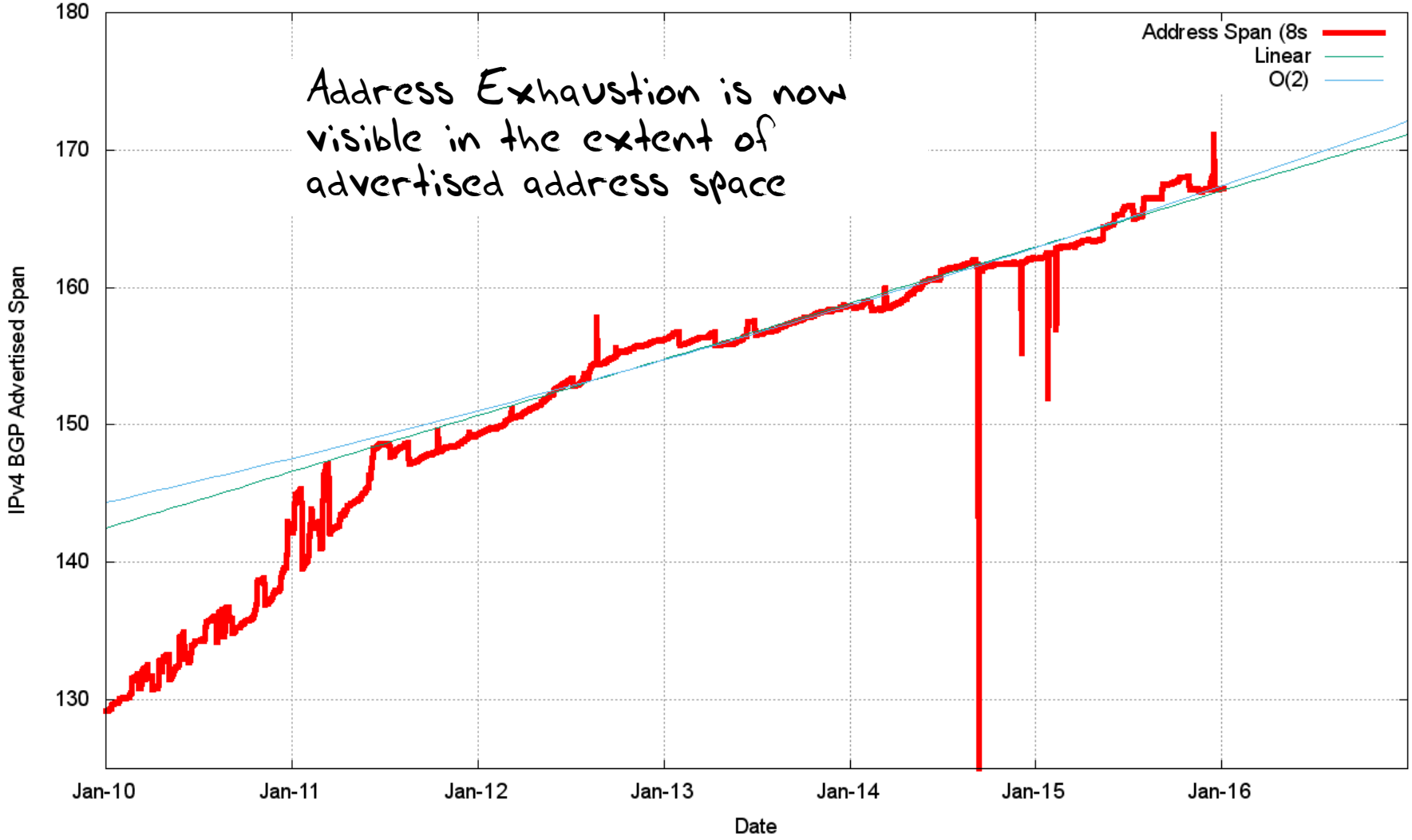
Routing Indicators for IPv4



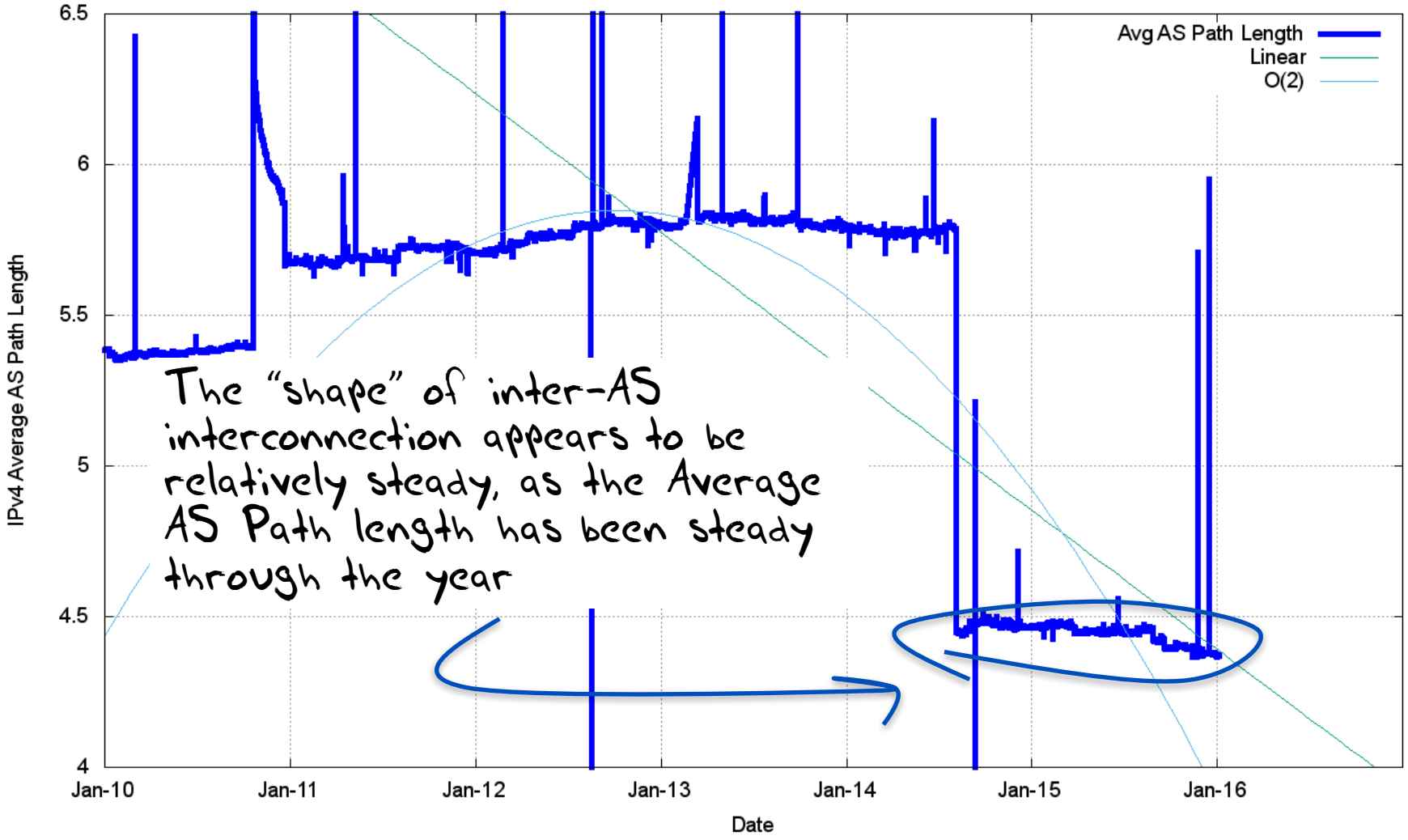
Routing Indicators for IPv4



Routing Indicators for IPv4



Routing Indicators for IPv4



What happened in 2015 in V4?

- From the look of the growth plots, its business as usual, despite the increasing pressure on IPv4 address availability
- The number of entries in the default-free zone is now heading to 600,000
- The pace of growth of the routing table is still relatively constant at ~50,000 new entries per year
  - IPv4 address exhaustion is not changing this!

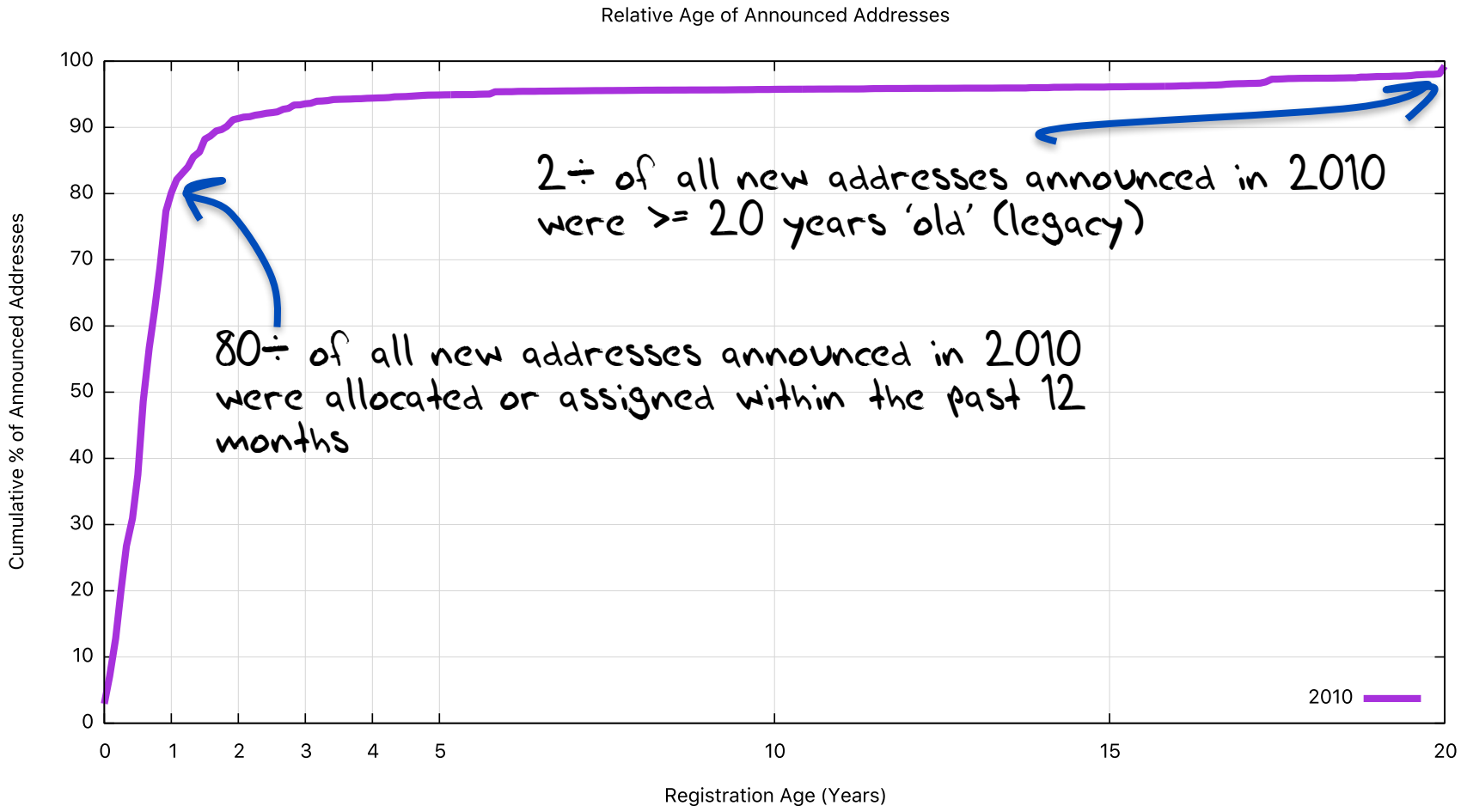
How can the IPv4 network continue to grow when we are running out of IPv4 addresses?

We are now recycling old addresses back into the routing system

Some of these addresses are transferred in ways that are recorded in the registry system, while others are being “leased” without any clear registration entry that describes the lessee

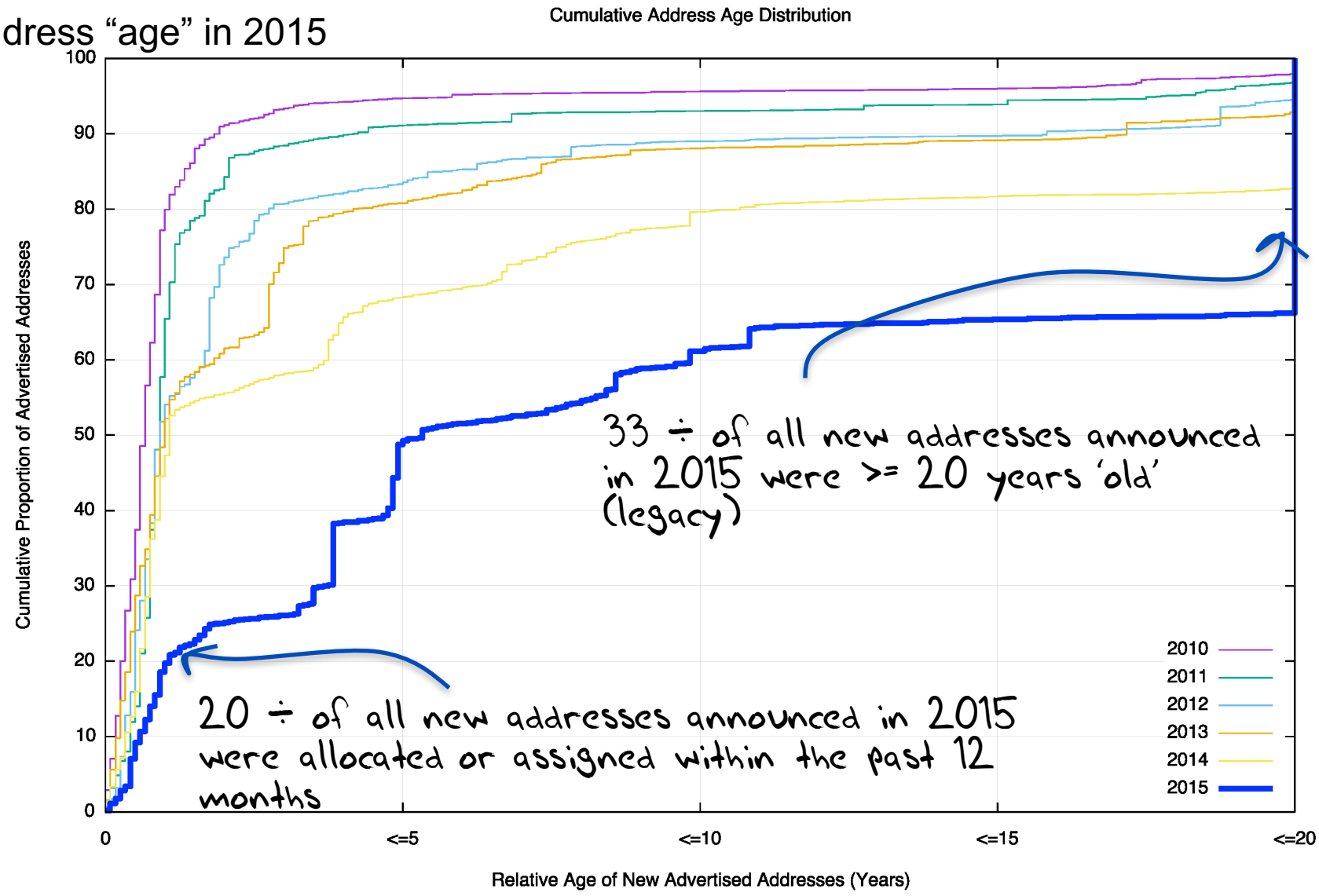
Address "Age"

Address “age” in 2010



IPv4 Address Reuse

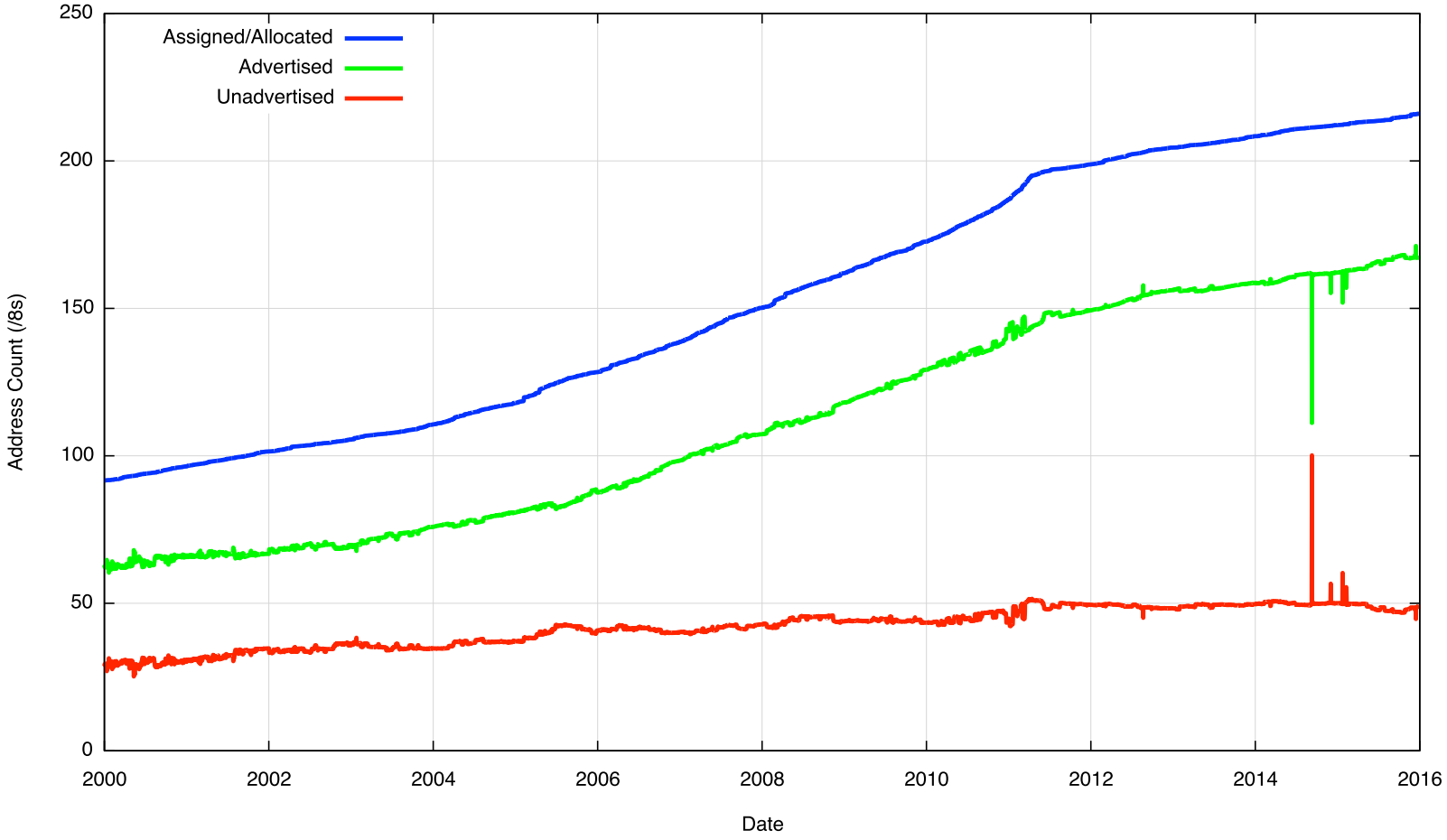
Address “age” in 2015



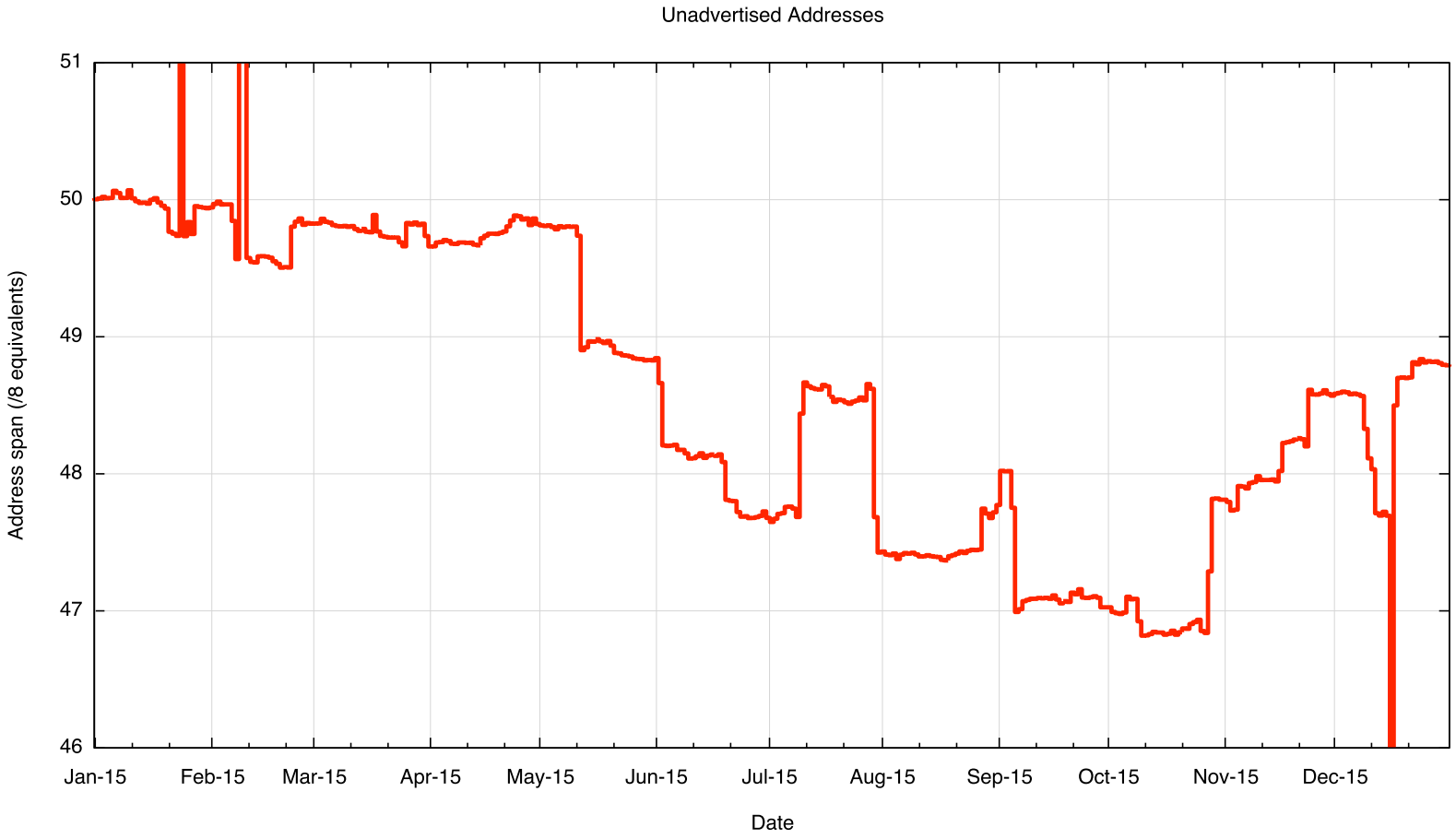
IPv4 in 2015 - Growth is Steady

- Overall IPv4 Internet growth in terms of BGP is at a rate of some **~47,000 entries p.a.**
- But we've run out of the unallocated address pools everywhere except Afrinic
- So what's driving this post-exhaustion growth?
  - Transfers?
  - Last /8 policies in RIPE and APNIC?
  - Leasing and address recovery?

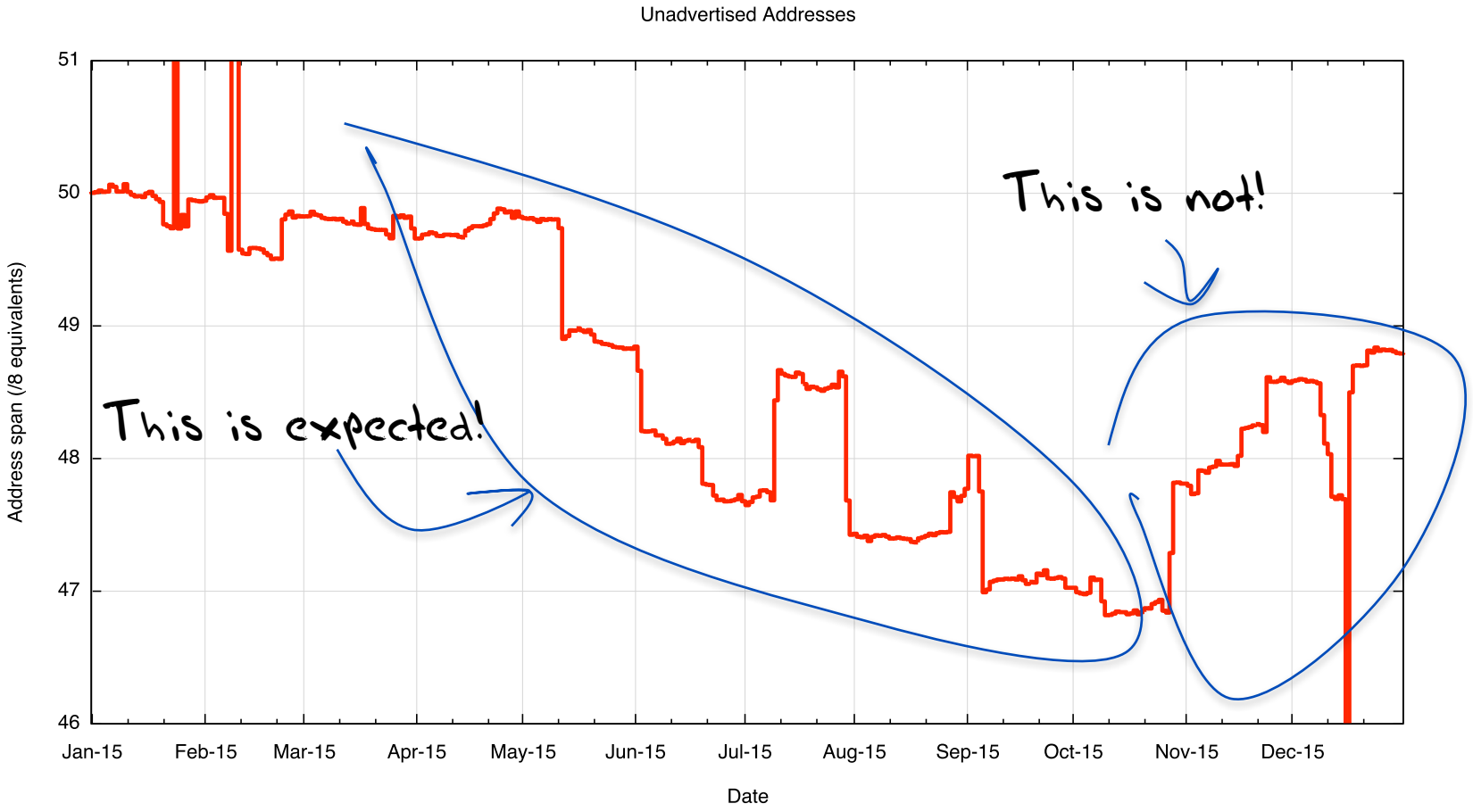
IPv4: Advertised vs Unadvertised Addresses



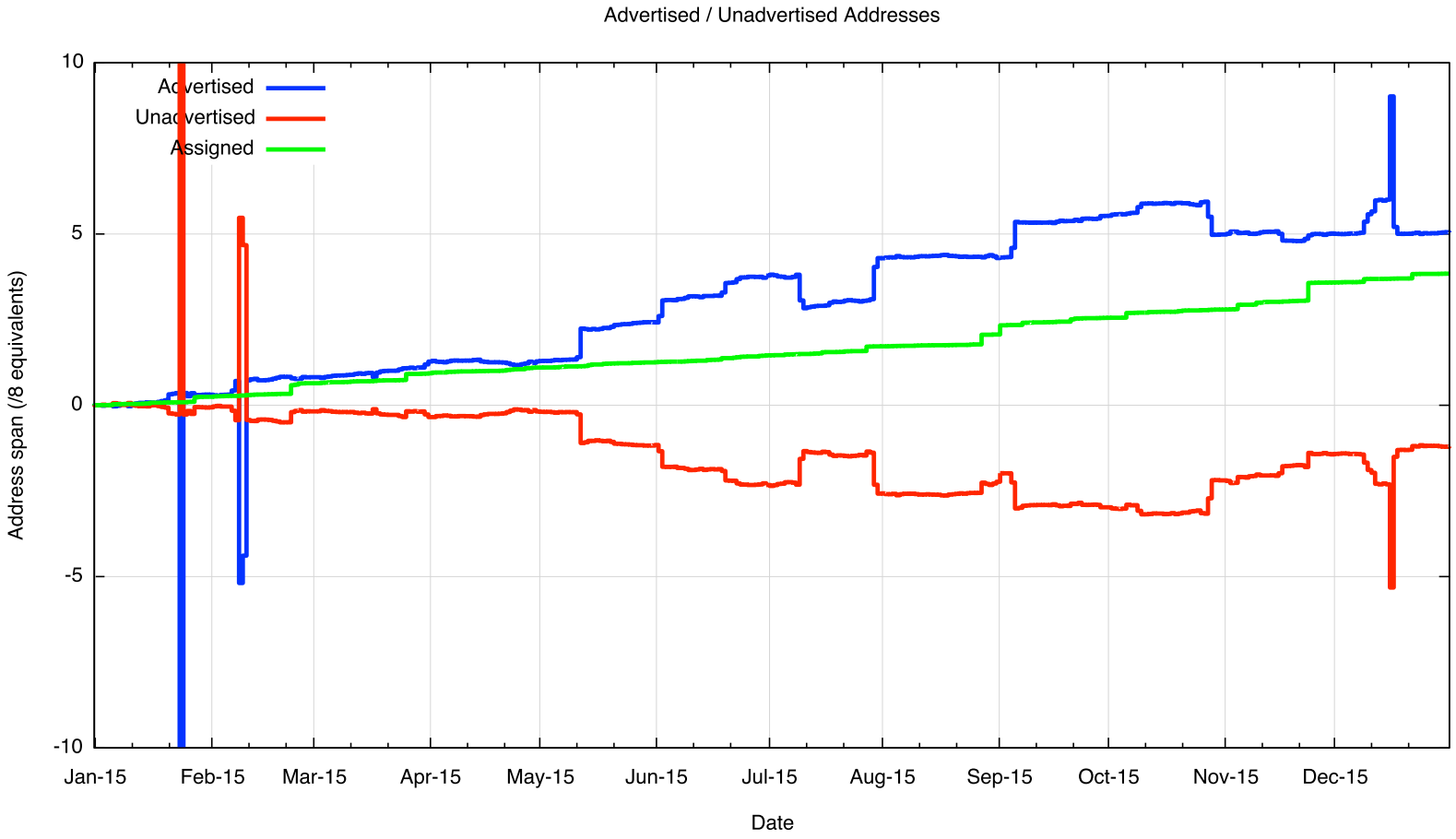
IPv4: Unadvertised Addresses



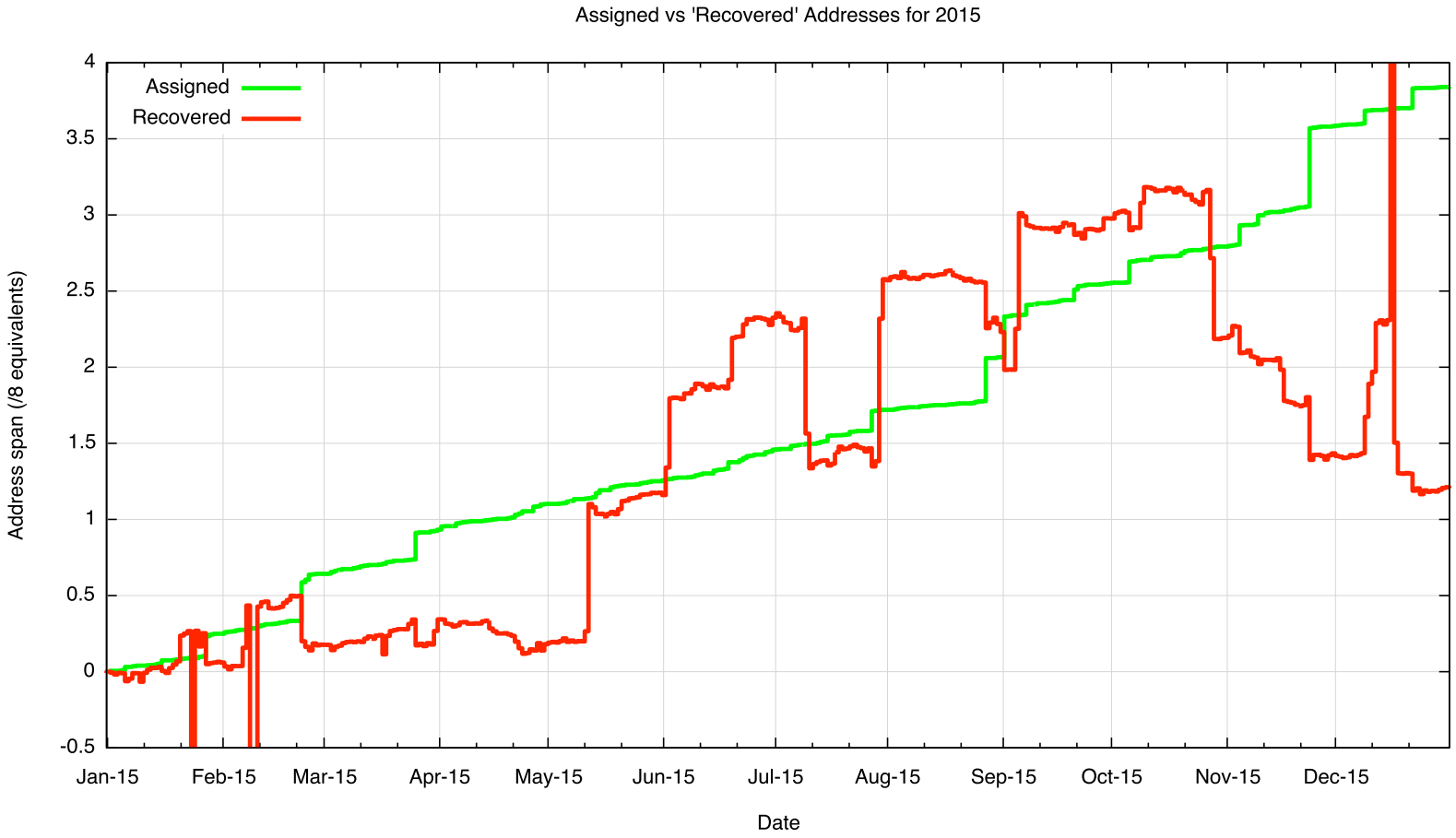
IPv4: Unadvertised Addresses



IPv4: Unadvertised Addresses



IPv4:Assigned vs Recovered



IPv4 in 2015

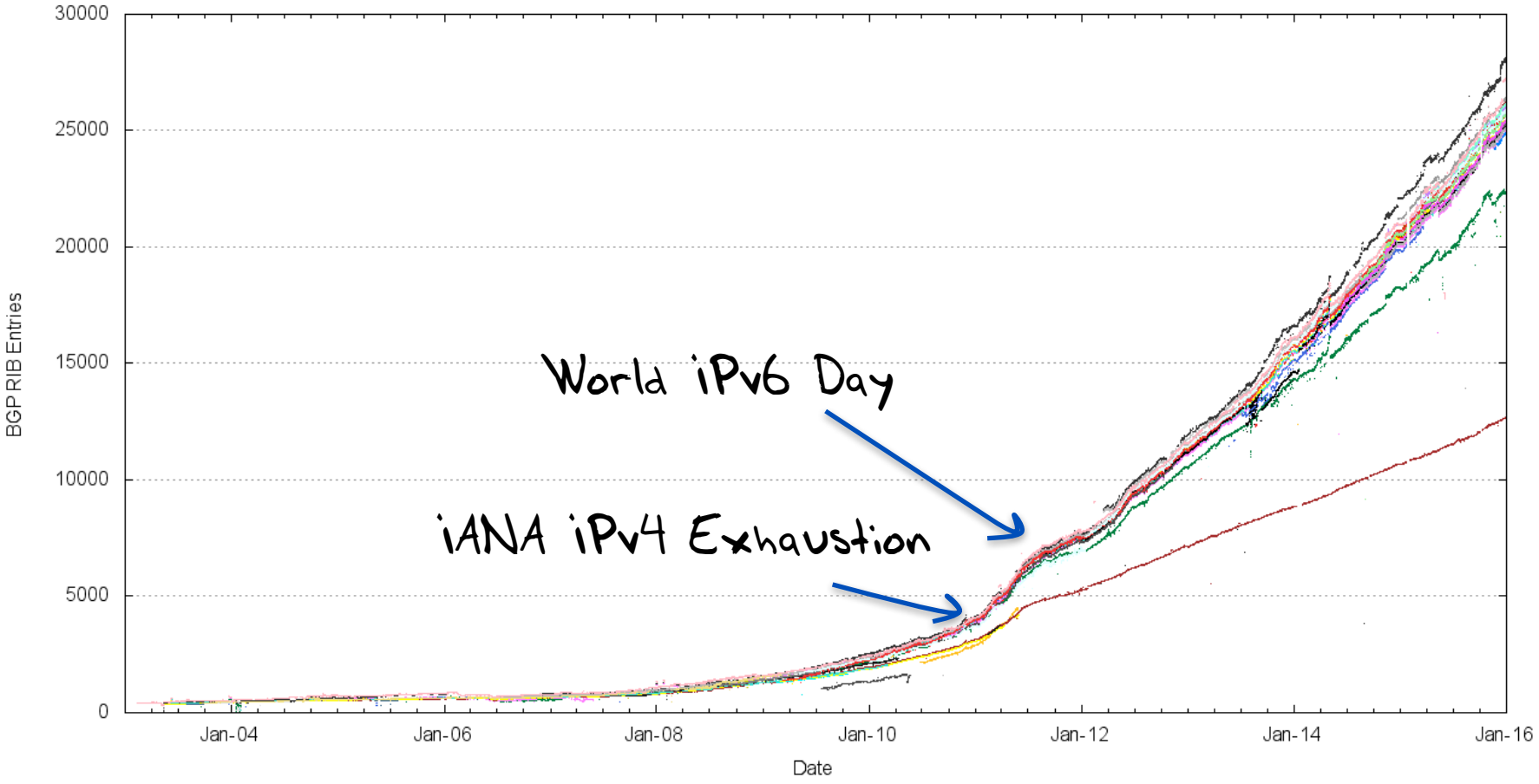
Approximately 4 /8s were assigned and advertised in 2015

- 2.3 /8s were assigned by ARIN
- 1 /8 assigned by AfriNIC

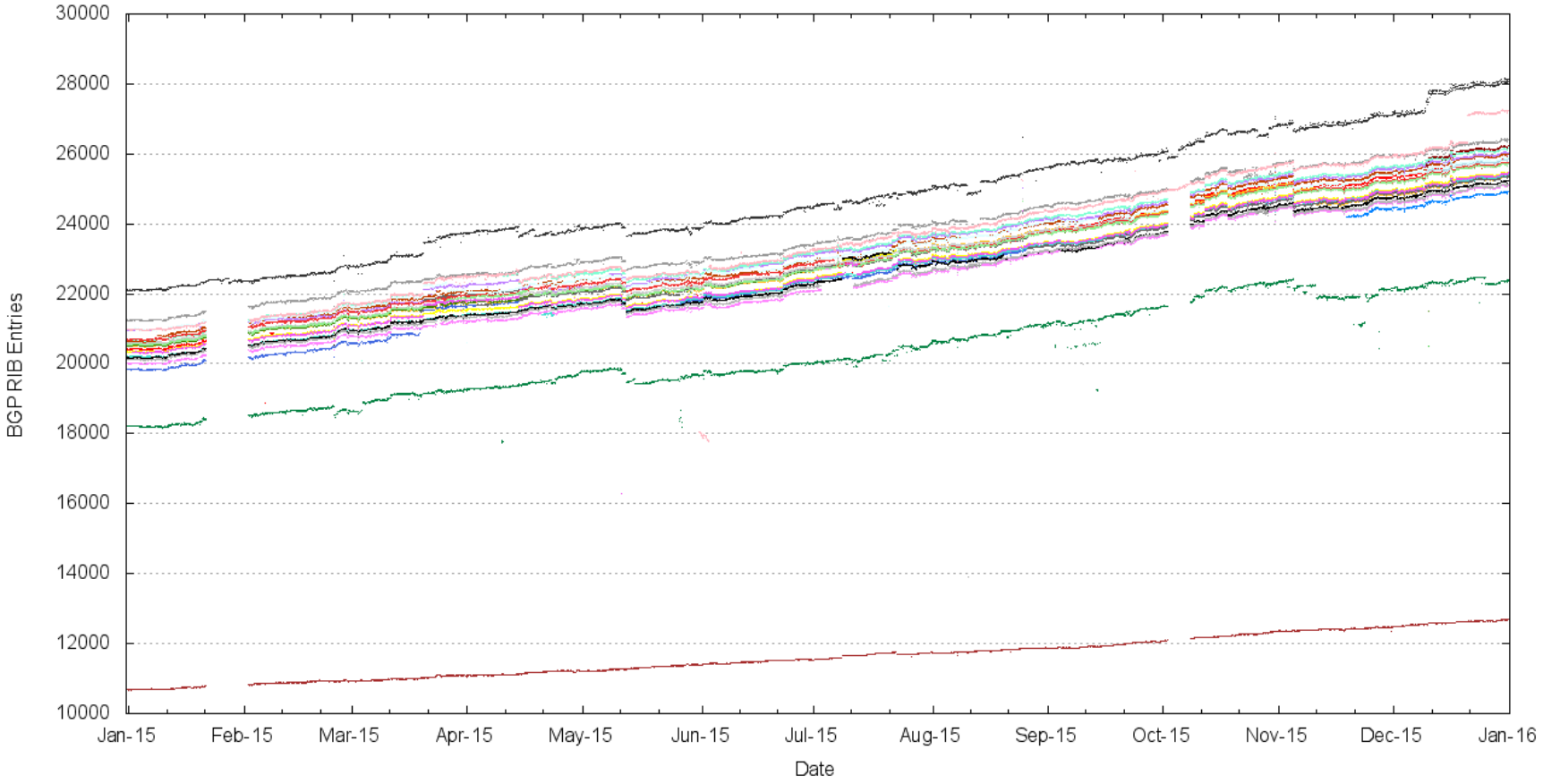
Up to 3 /8s were ‘recovered’ from the unallocated address pool and advertised during 2015

- But 2/8s of addresses were withdrawn in the last two months of the year

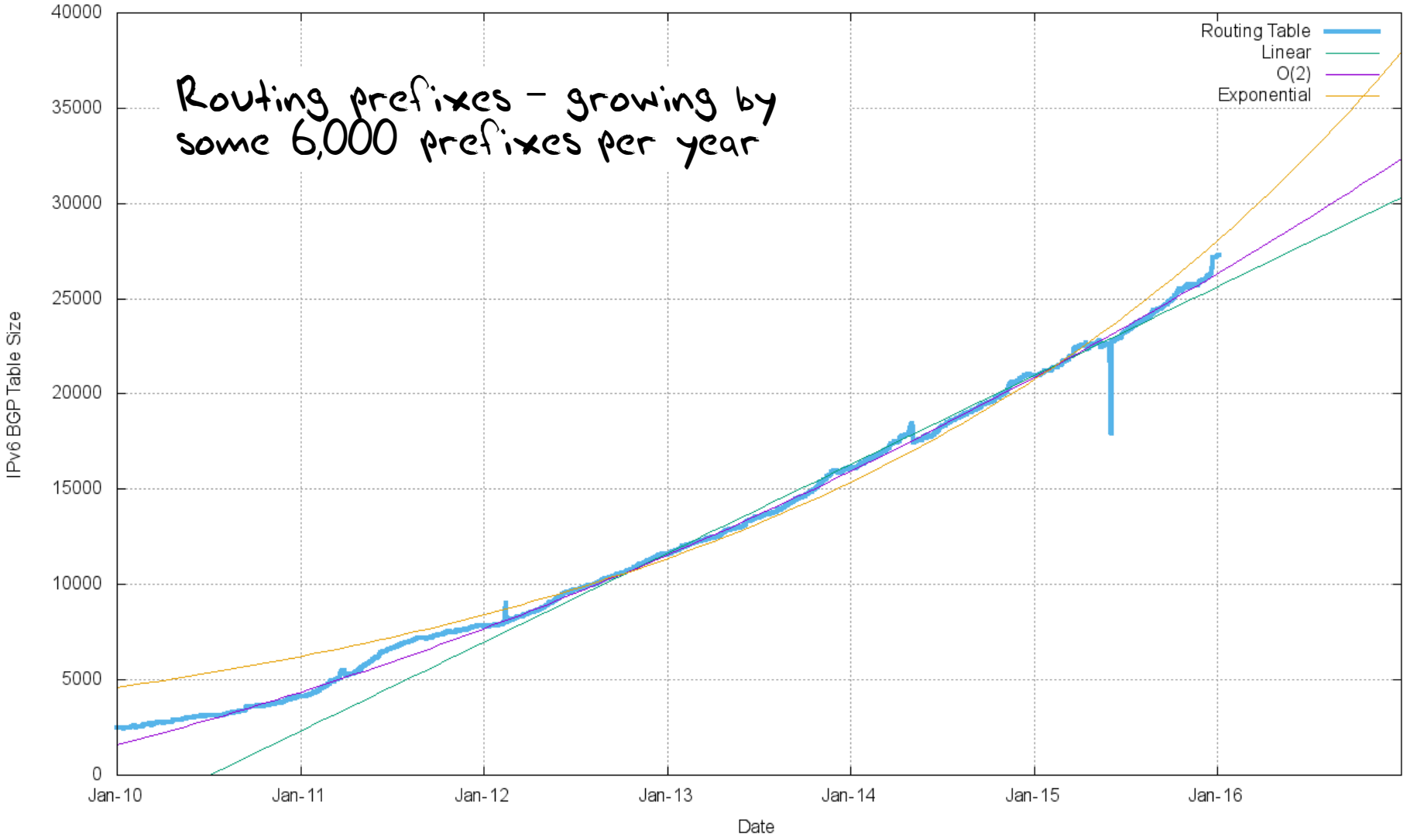
The Route Views view of IPv6



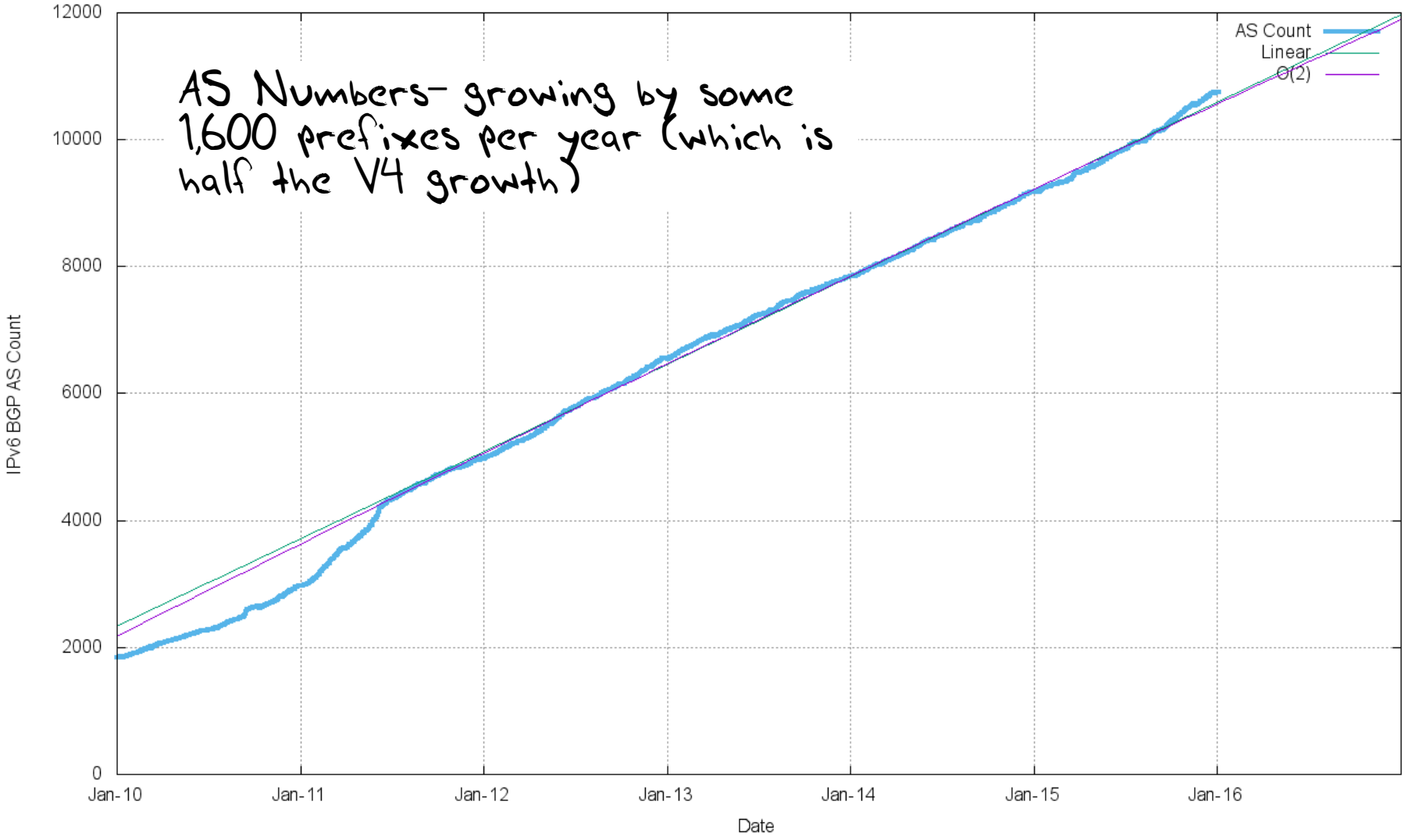
2015 for IPv6, as seen at Route Views



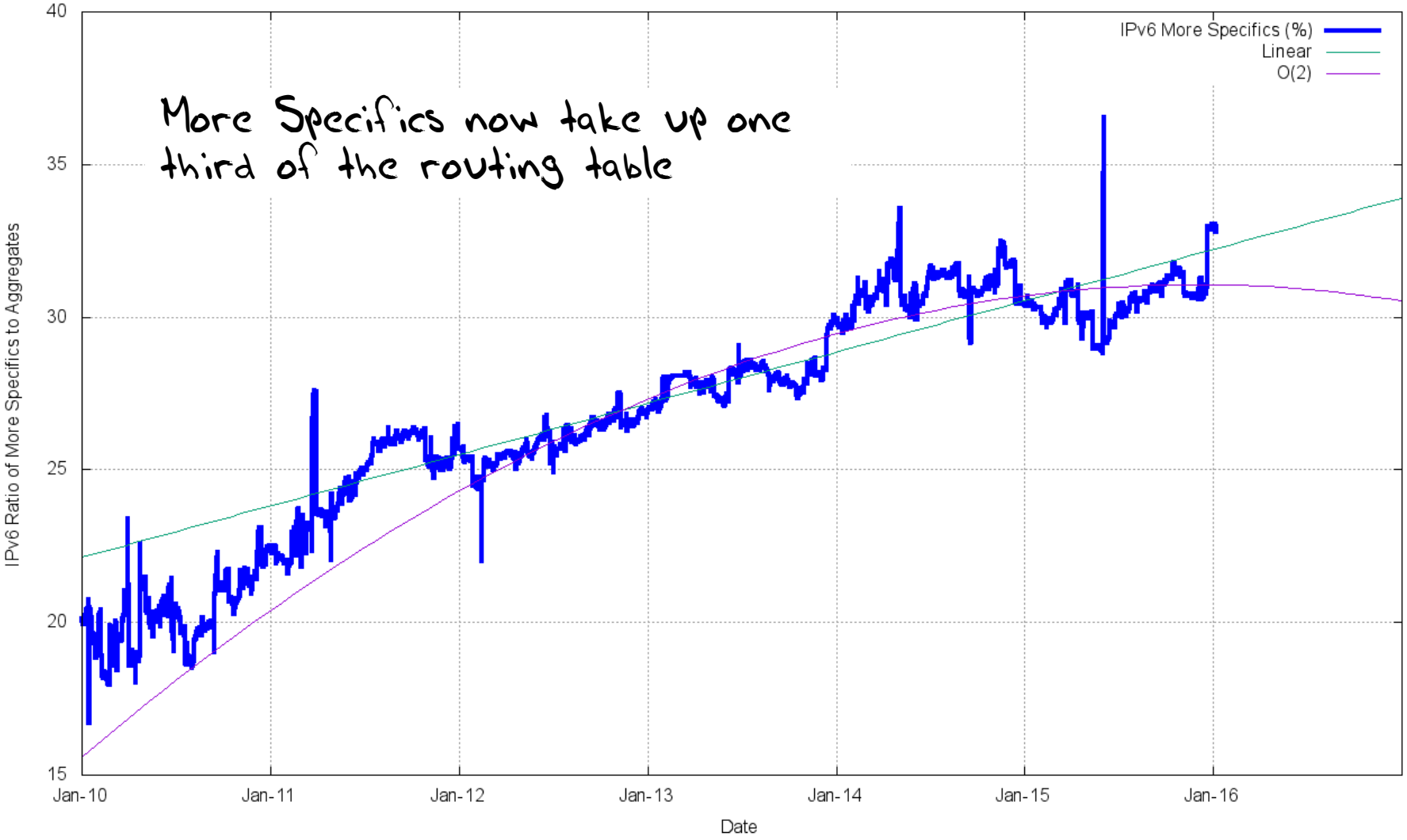
Routing Indicators for IPv6



Routing Indicators for IPv6

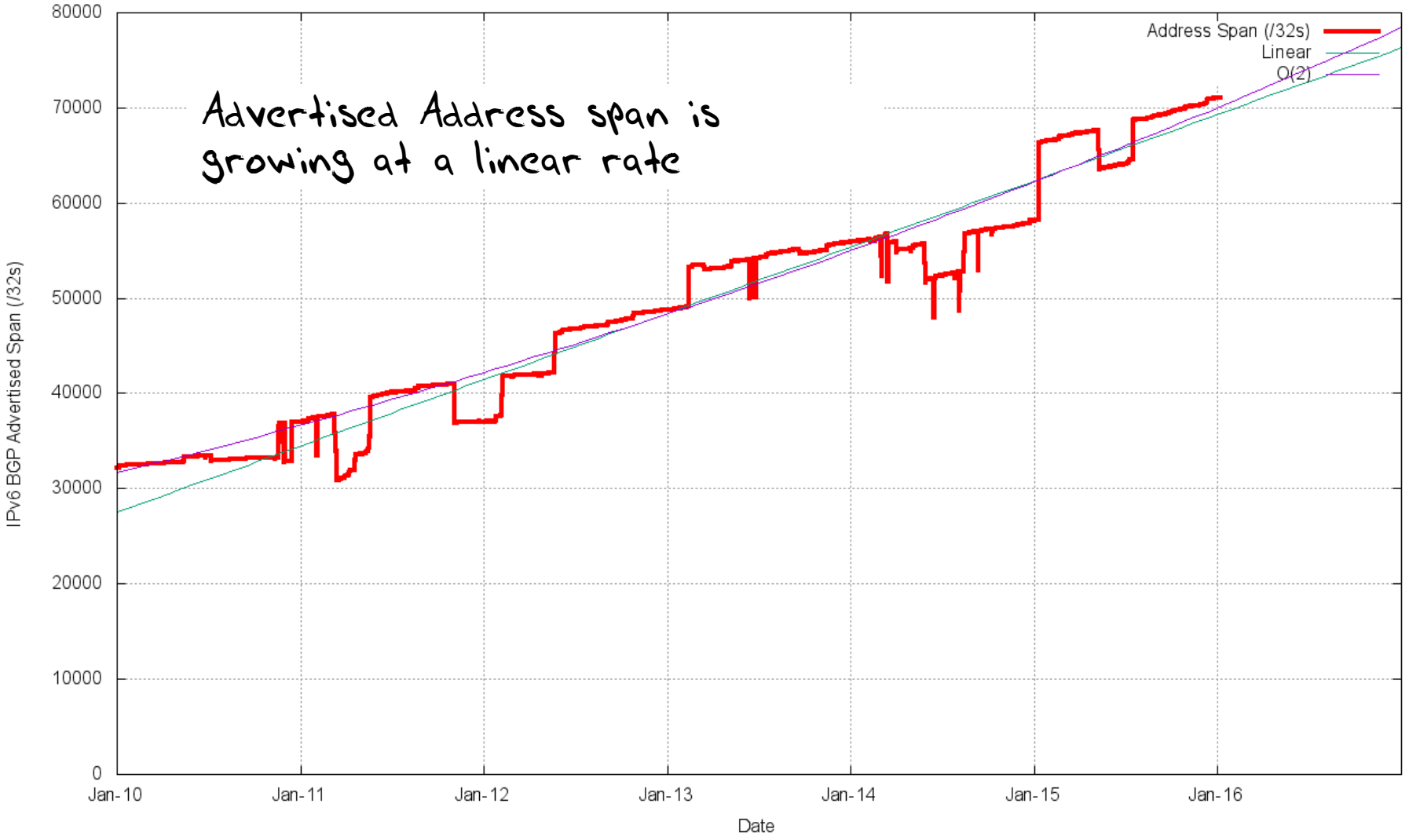


Routing Indicators for IPv6



Routing Indicators for IPv6

Routing Indicators for IPv6



Routing Indicators for IPv6

The "shape" of inter-AS interconnection appears to be steady, as the Average AS Path length has been held steady through the year

IPv6 in 2015

- Overall IPv6 Internet growth in terms of BGP is steady at some **6,000 route entries p.a.**

This is growth of BGP route objects is 1/7 of the growth rate of the IPv4 network – as compared to the AS growth rate which is 1/2 of the IPv4 AS number growth rate

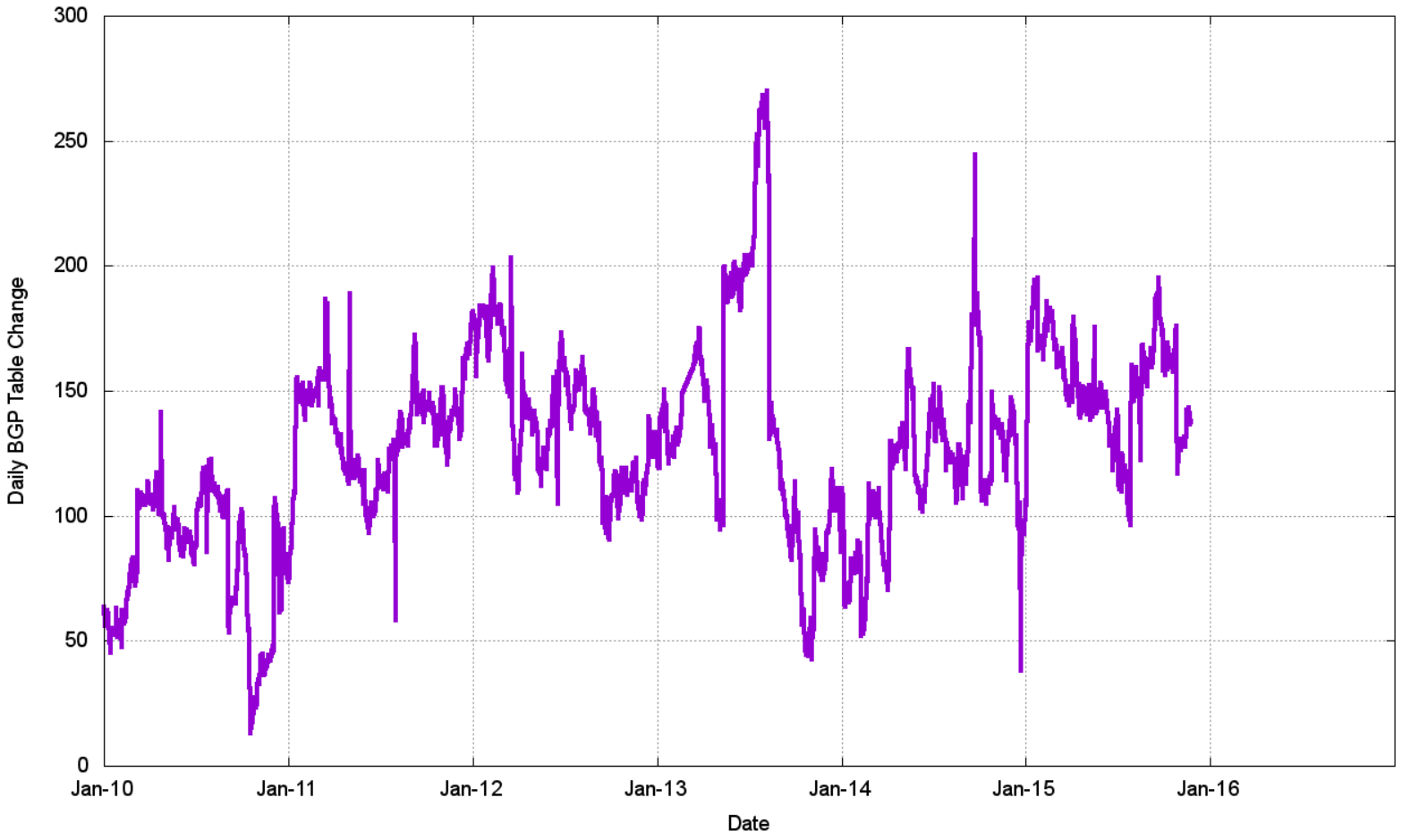
What to expect

BGP Size Projections

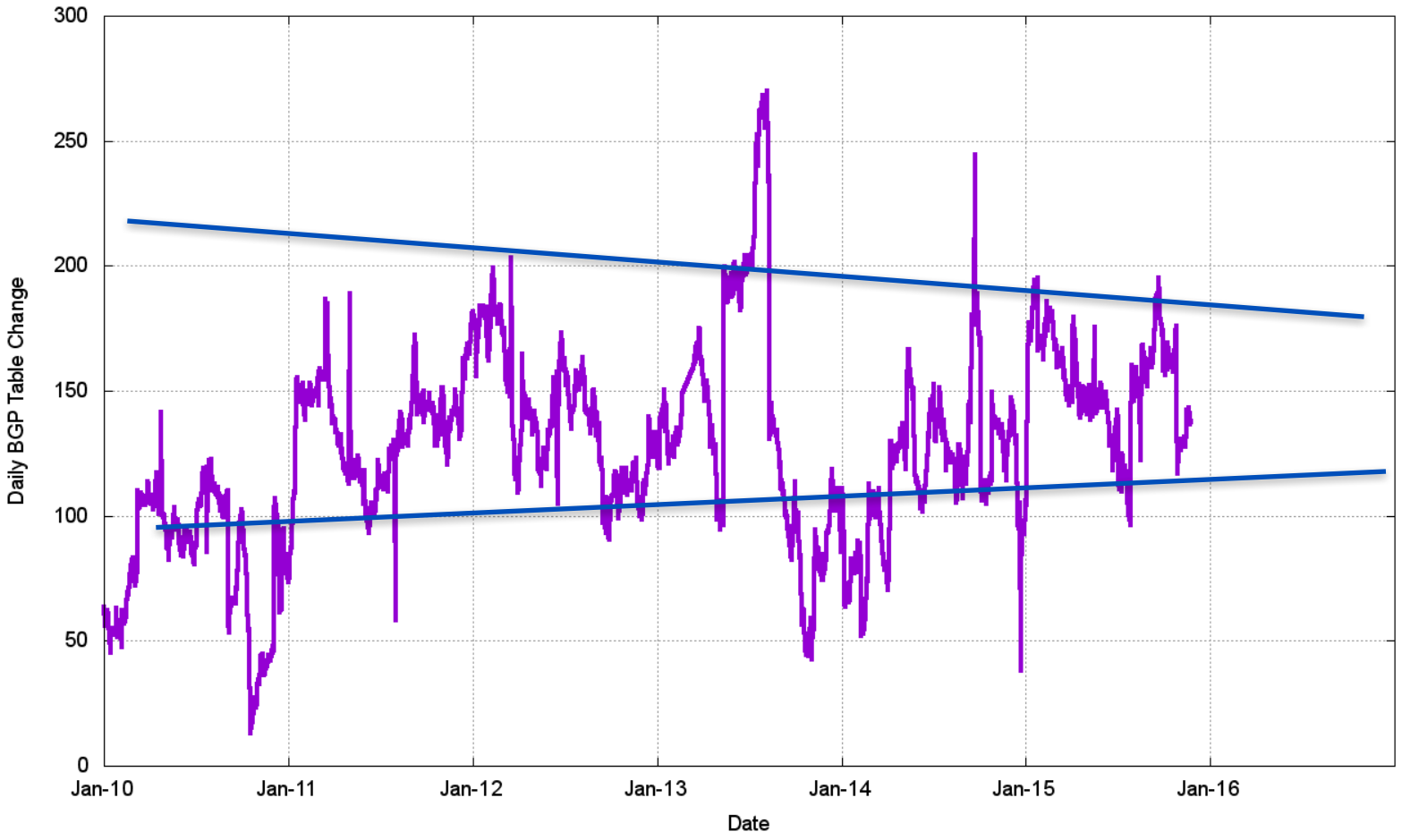
For the Internet this is a time of extreme uncertainty

- Registry IPv4 address run out
- Uncertainty over the impacts of any after-market in IPv4 on the routing table
- Uncertainty over IPv6 takeup leads to a mixed response to IPv6 so far, and no clear indicator of trigger points for change

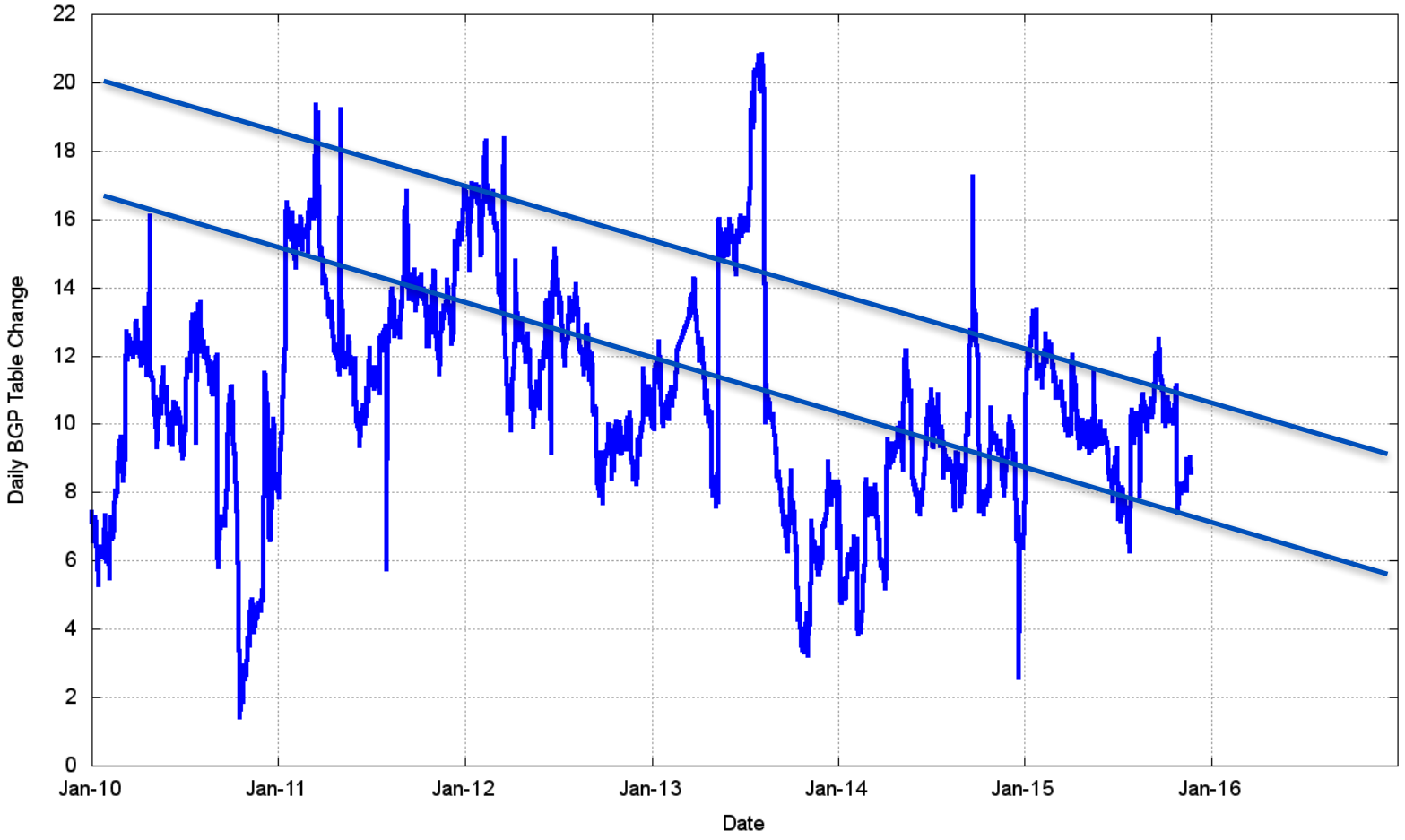
V4 - Daily Growth Rates



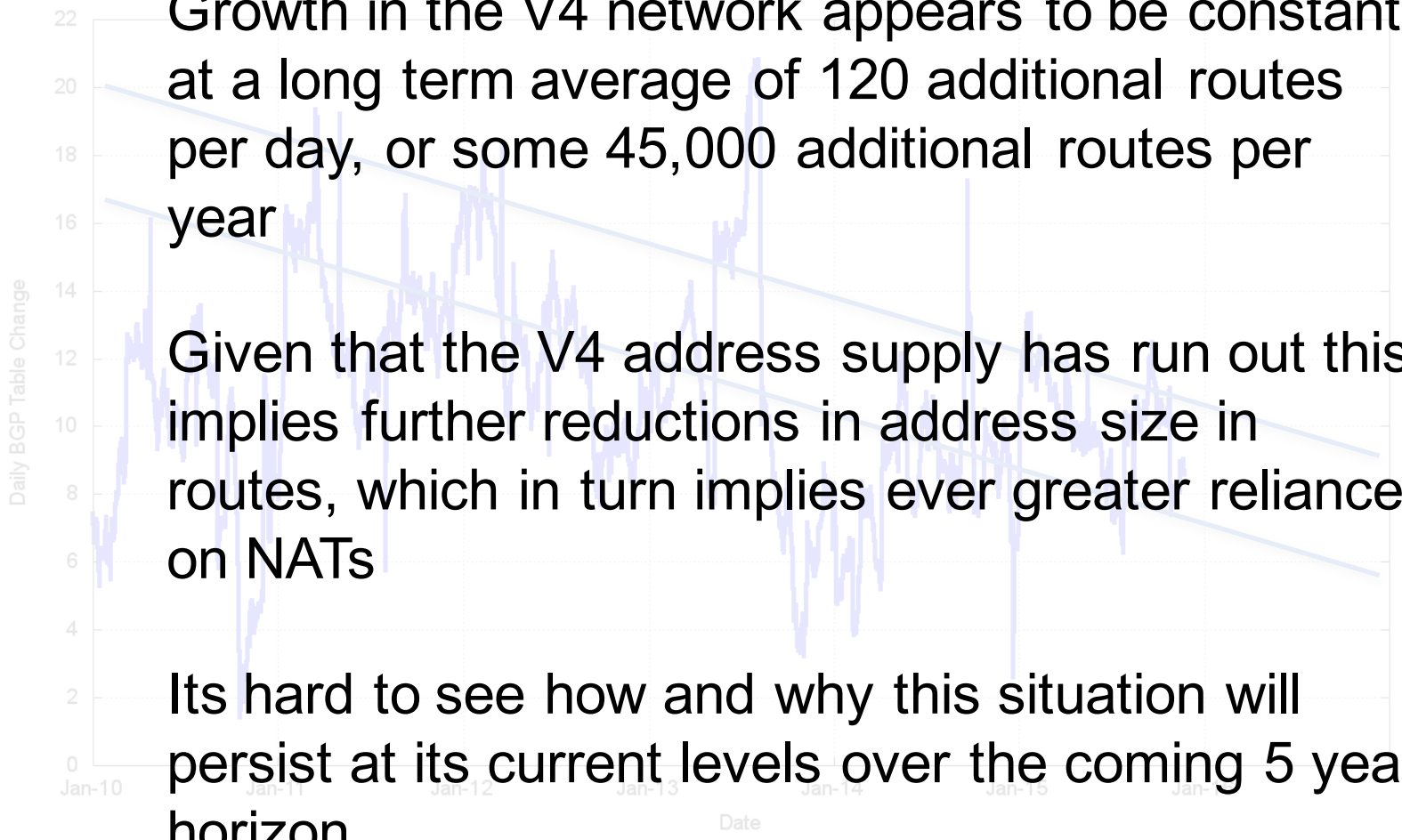
V4 - Daily Growth Rates



V4 - Relative Daily Growth Rates



V4 - Relative Daily Growth Rates



Growth in the V4 network appears to be constant at a long term average of 120 additional routes per day, or some 45,000 additional routes per year

Given that the V4 address supply has run out this implies further reductions in address size in routes, which in turn implies ever greater reliance on NATs

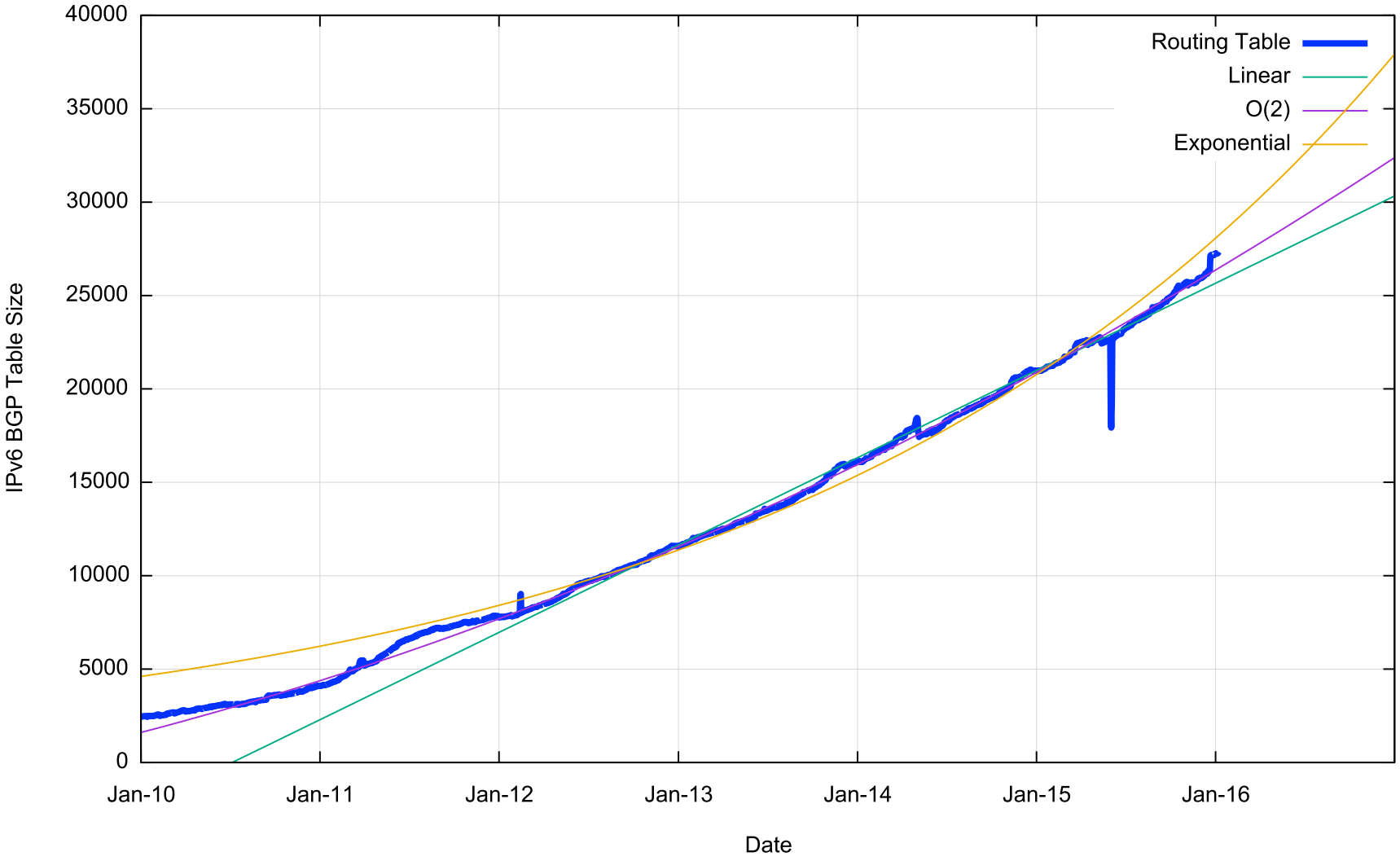
Its hard to see how and why this situation will persist at its current levels over the coming 5 year horizon

IPv4 BGP Table Size predictions

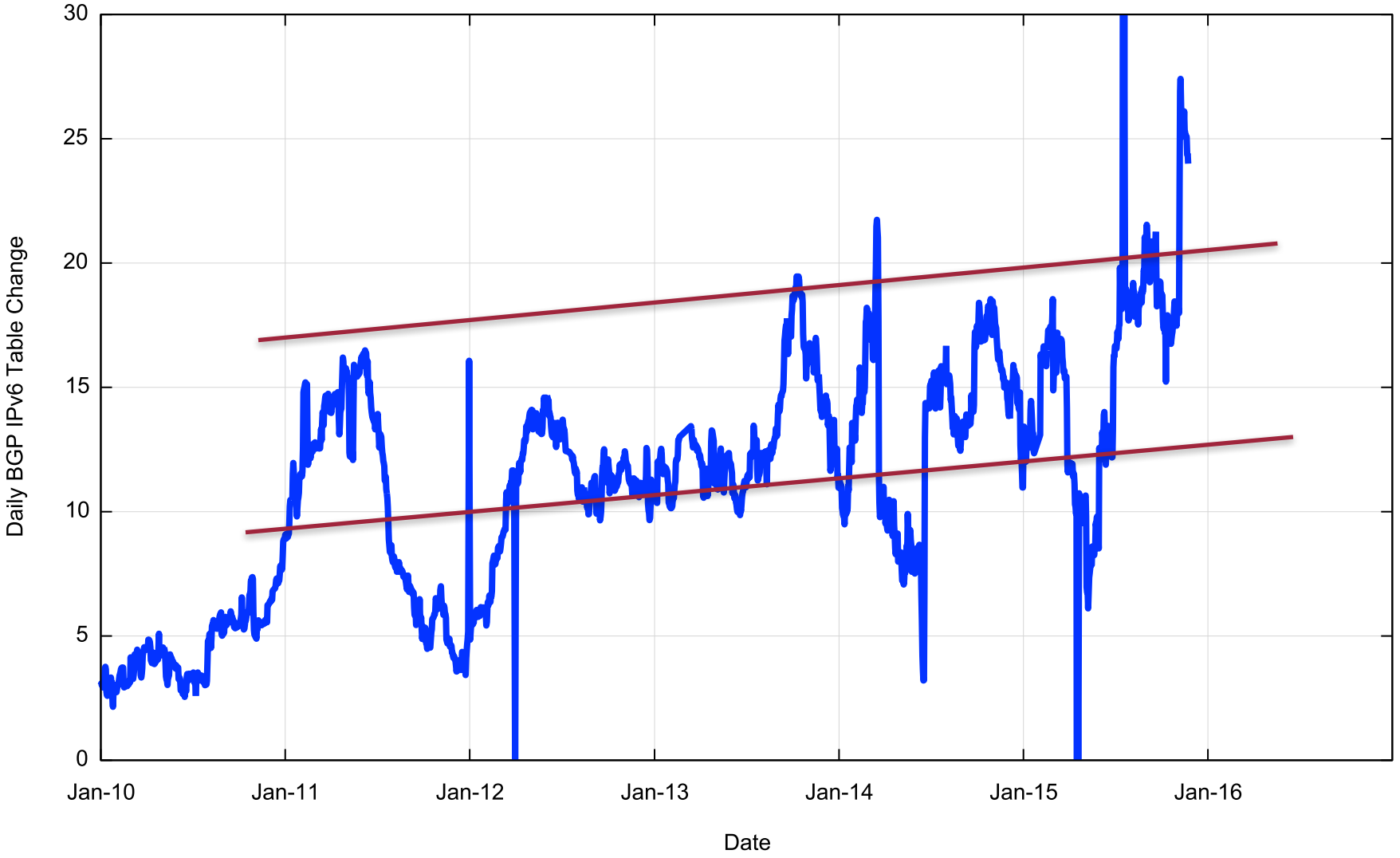
		2014 PREDICTION	2013 PREDICTION
Jan 2013	441,000		
2014	488,000		
2015	*530,000*		*540,000*
2016	***586,000***	*580,000*	*590,000*
2017	*628,000*	*620,000*	*640,000*
2018	*675,000*	*670,000*	*690,000*
2019	*722,000*	*710,000*	*740,000*
2020	*768,000*	*760,000*	
2021	*815,000*		

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

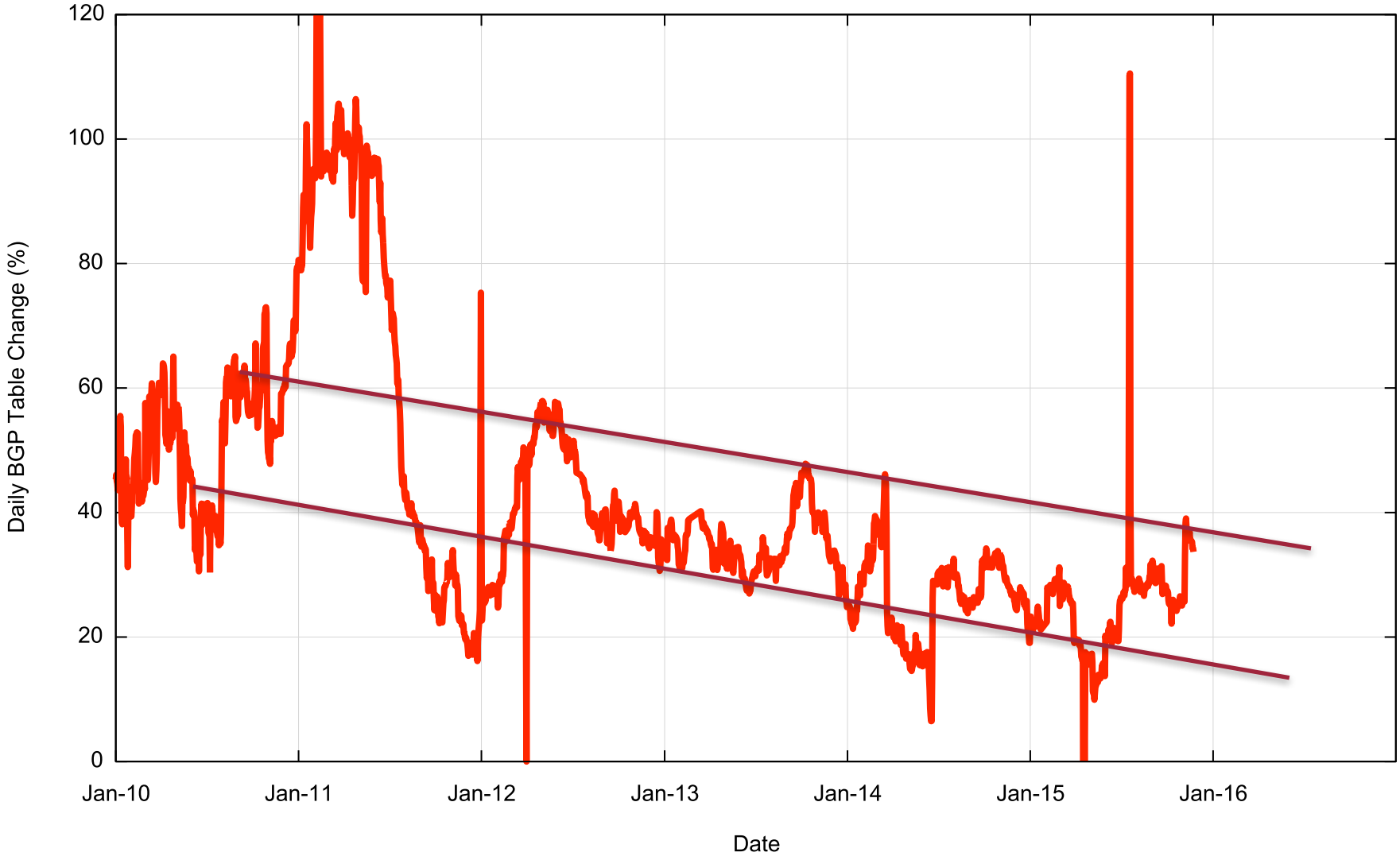
IPv6 Table Size



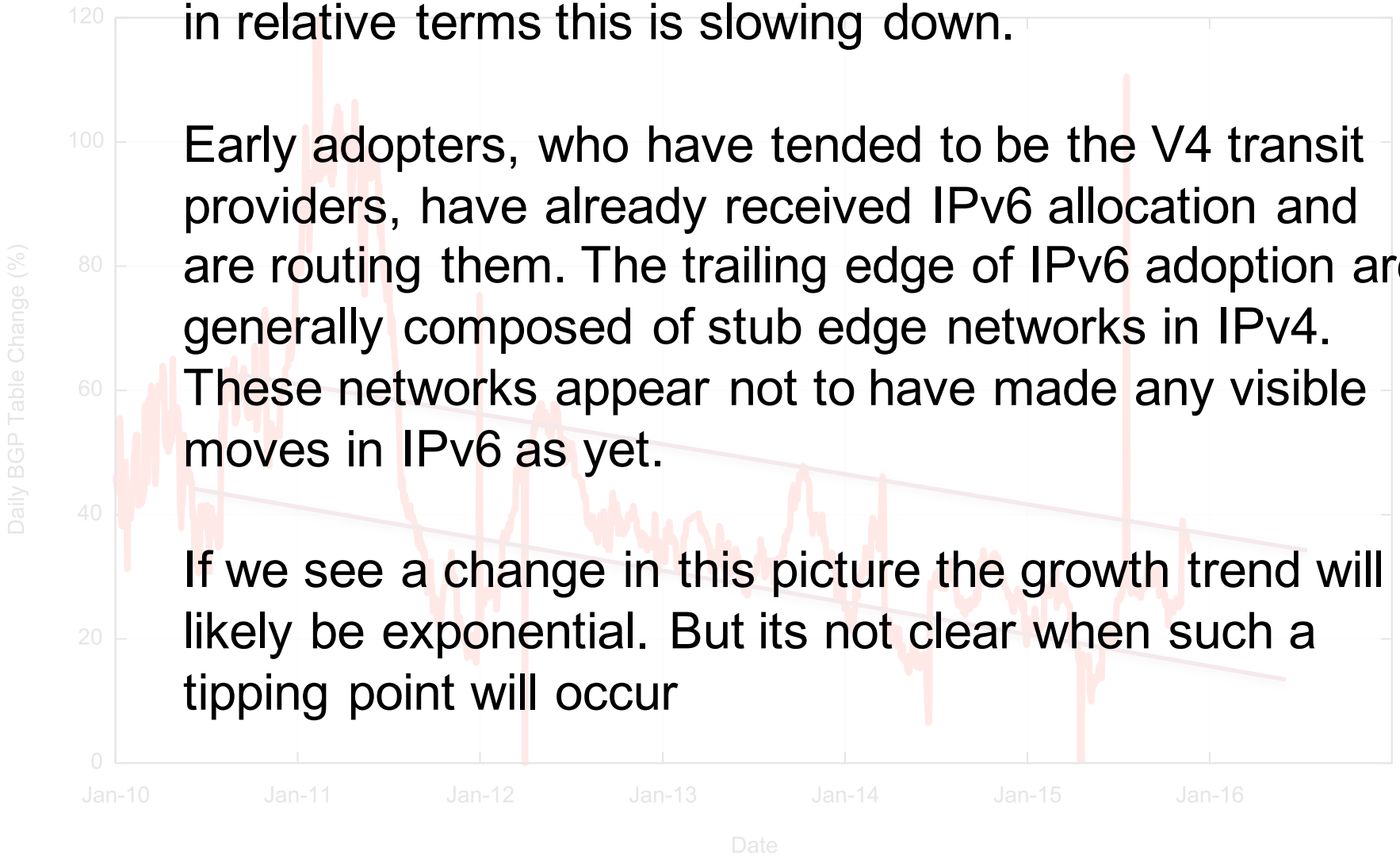
V6 - Daily Growth Rates



V6 - Relative Growth Rates



V6 - Relative Growth Rates



Growth in the V6 network appears to be increasing, but in relative terms this is slowing down.

Early adopters, who have tended to be the V4 transit providers, have already received IPv6 allocation and are routing them. The trailing edge of IPv6 adoption are generally composed of stub edge networks in IPv4. These networks appear not to have made any visible moves in IPv6 as yet.

If we see a change in this picture the growth trend will likely be exponential. But its not clear when such a tipping point will occur

IPv6 BGP Table Size predictions

		Exponential Model	Linear Model
Jan 2014	16,100 entries		
2015	21,200		
2016	*27,000*		
2017		*38,000*	*30,000*
2018		*51,000*	*35,000*
2019		*70,000*	*40,000*
2020		*94,000*	*44,000*
2021		*127,000*	*49,000*

Range of potential outcomes

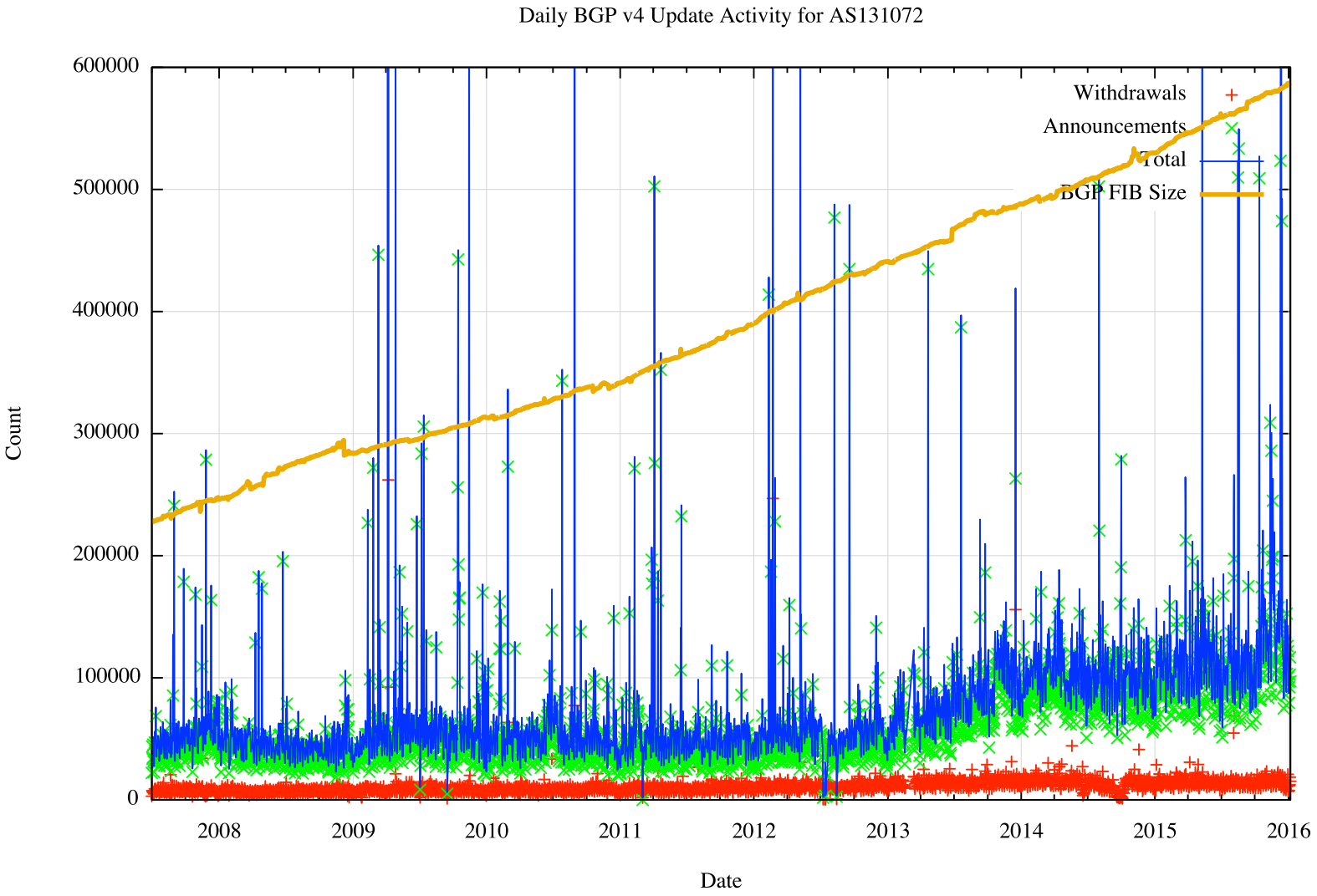
BGP Table Growth

- 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 at a faster relative rate, is still small in size in absolute terms
- As long as we are prepared to live within the technical constraints of the current routing paradigm, the Internet's use of BGP will continue to be viable for some time yet
- Nothing is melting in terms of the size of the routing table as yet

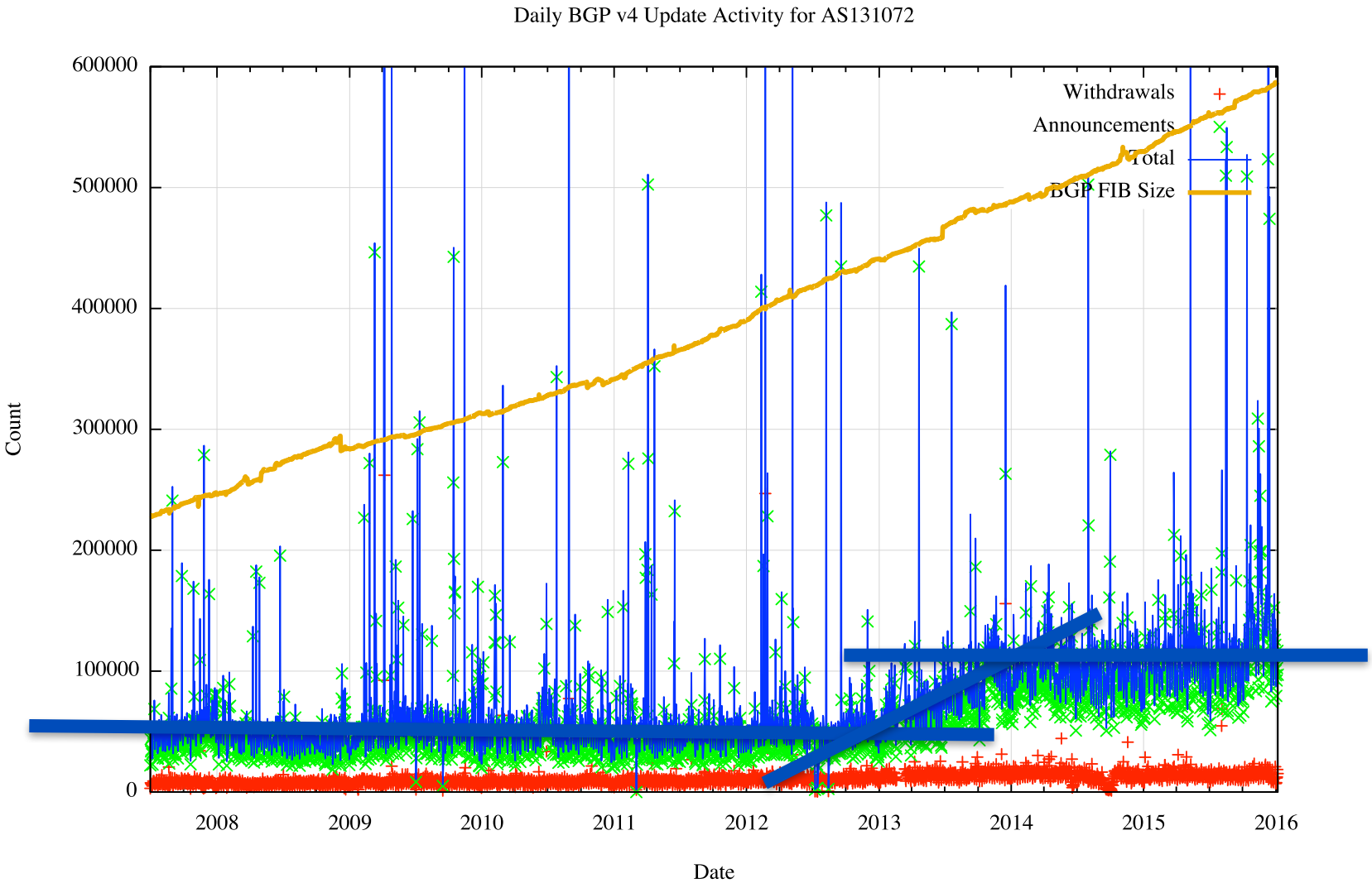
BGP Updates

- What about the level of updates in BGP?
- Let's look at the update load from a single eBGP feed in a DFZ context

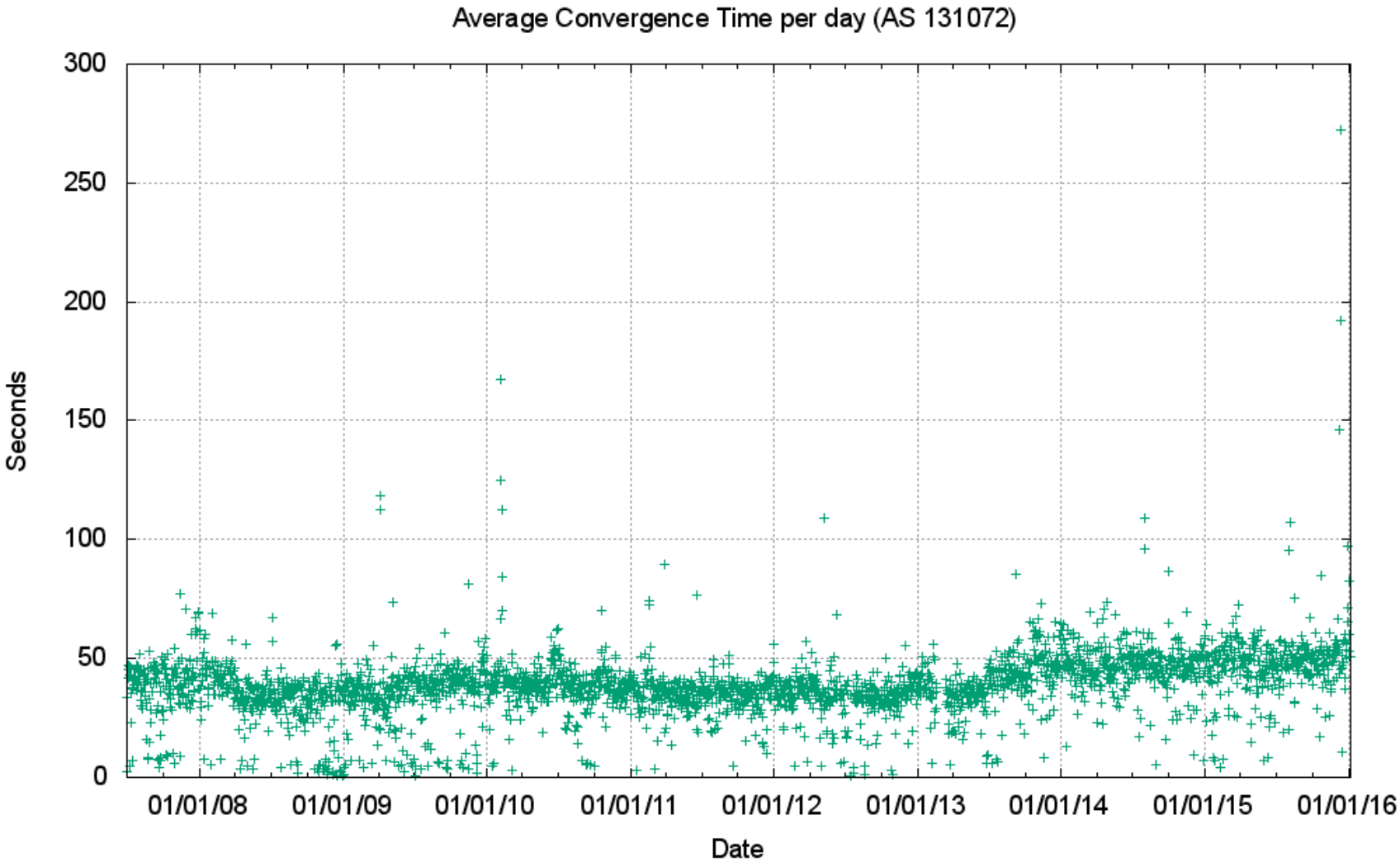
Announcements and Withdrawals



Announcements and Withdrawals



Convergence Performance



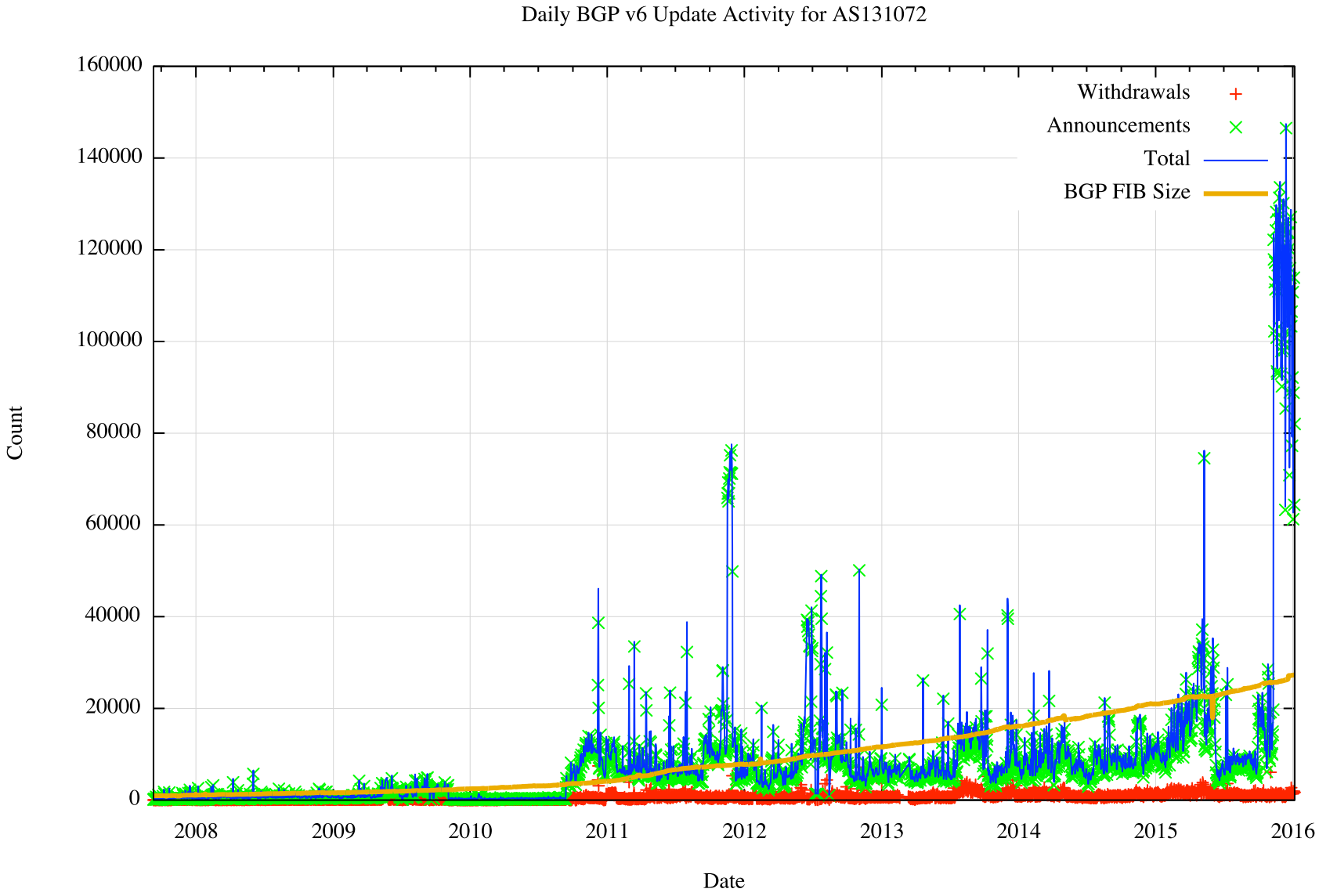
Updates in IPv4 BGP

Nothing in these figures is cause for any great level of concern …

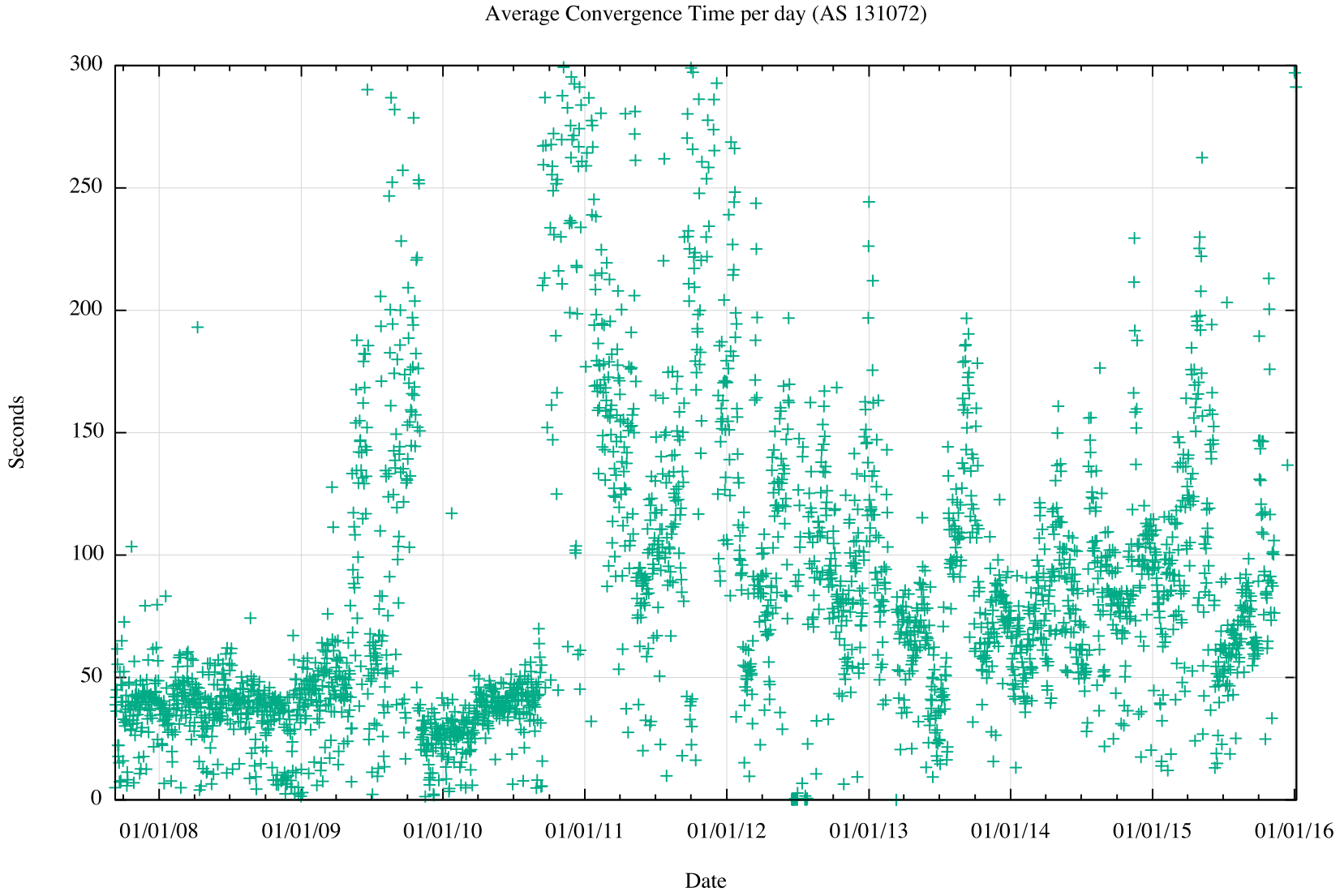
- The number of updates per instability event has been relatively constant, which for a distance vector routing protocol is weird, and completely unanticipated. Distance Vector routing protocols should get noisier as the population of protocol speakers increases, and the increase should be multiplicative.
- But this is not happening in the Internet
- Which is good, but why is this not happening?

Likely contributors to this outcome are the damping effect of widespread use of the MRAI interval by eBGP speakers, and the topology factor, as seen in the relatively constant V4 AS Path Length

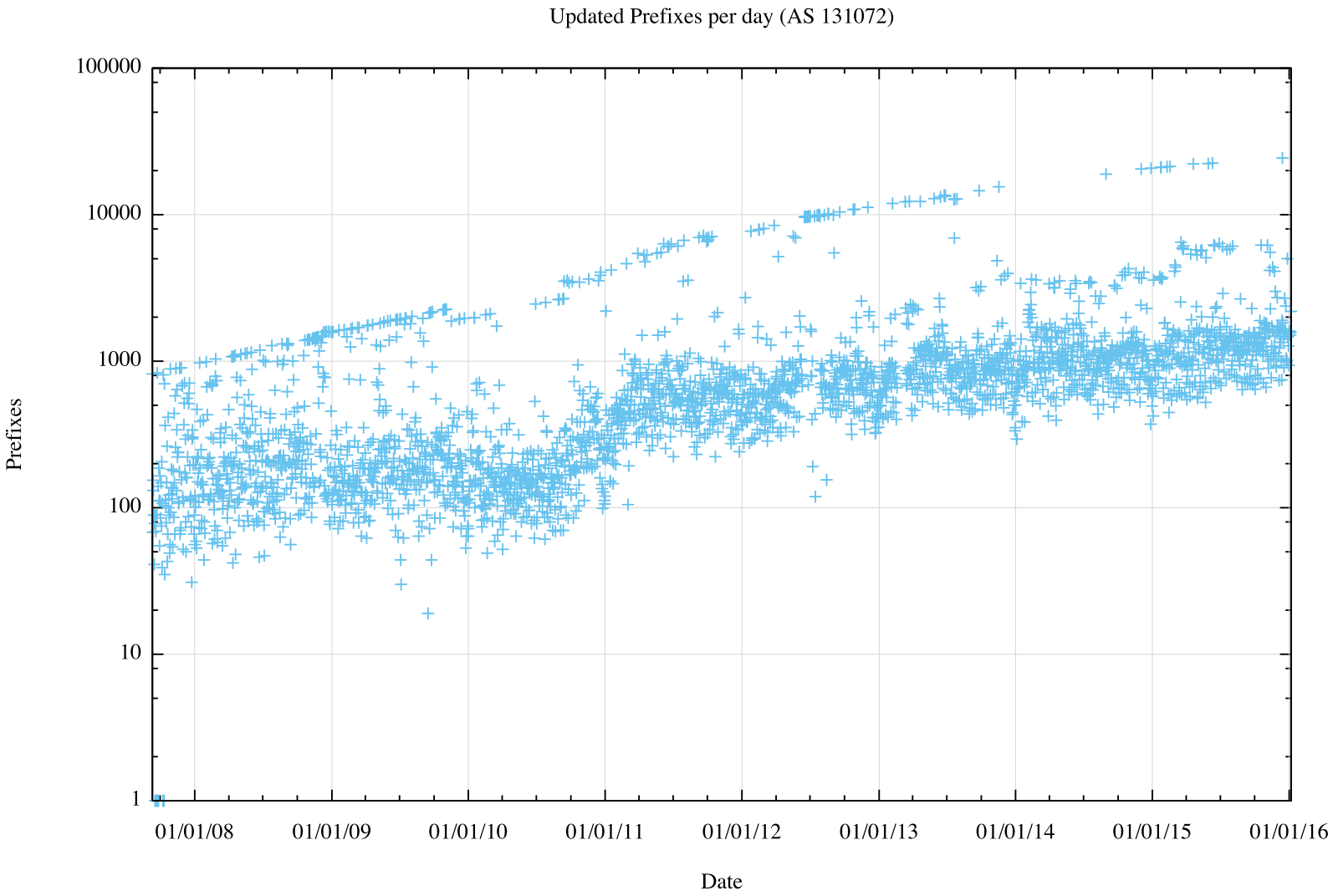
V6 Announcements and Withdrawals



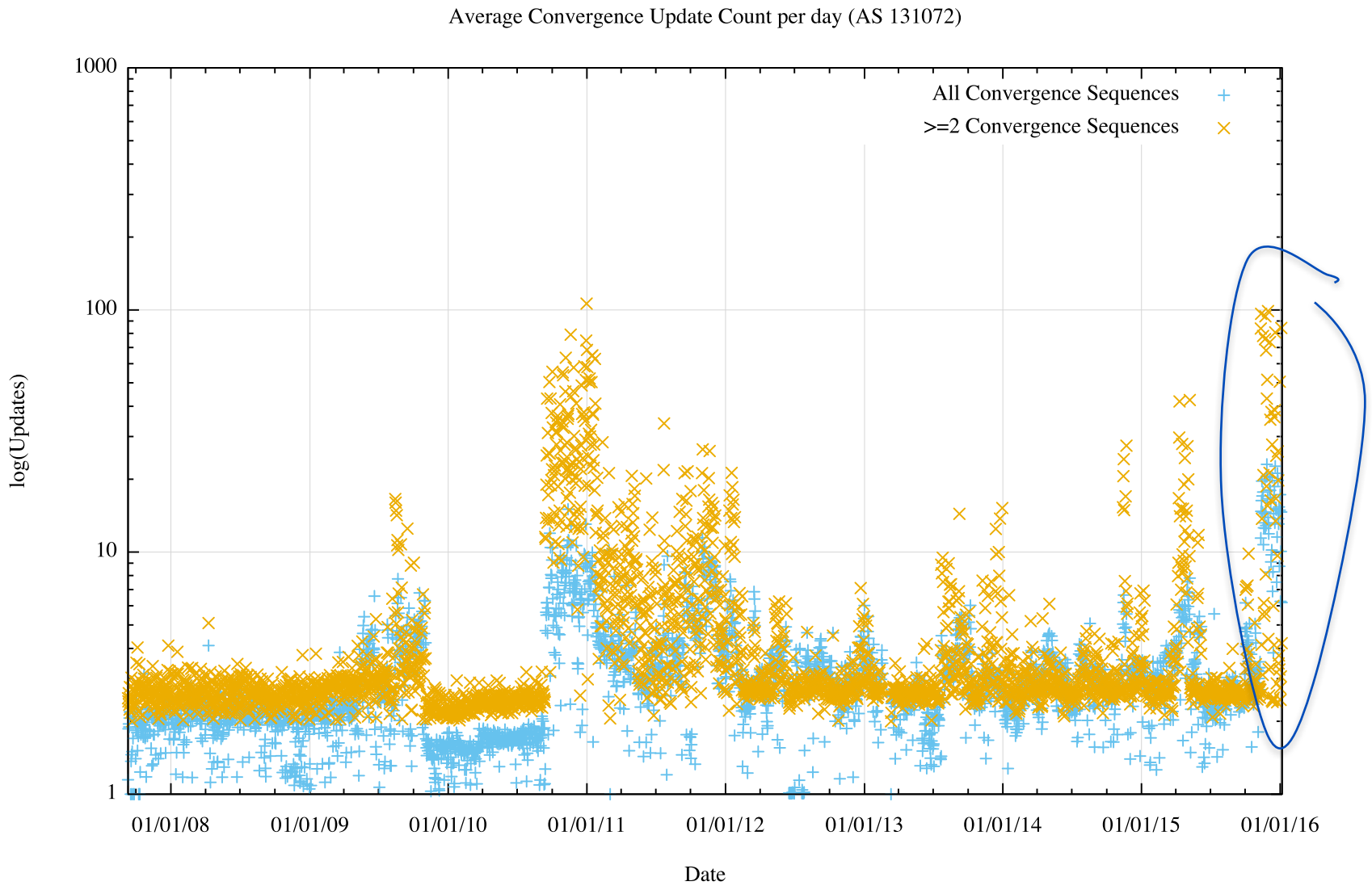
V6 Convergence Performance



V6 Updated prefixes per day



V6 Updates per event



Updates in IPv6 BGP

IPv6 routing behaviour is diverging from IPv4 behaviour

The instability is greater

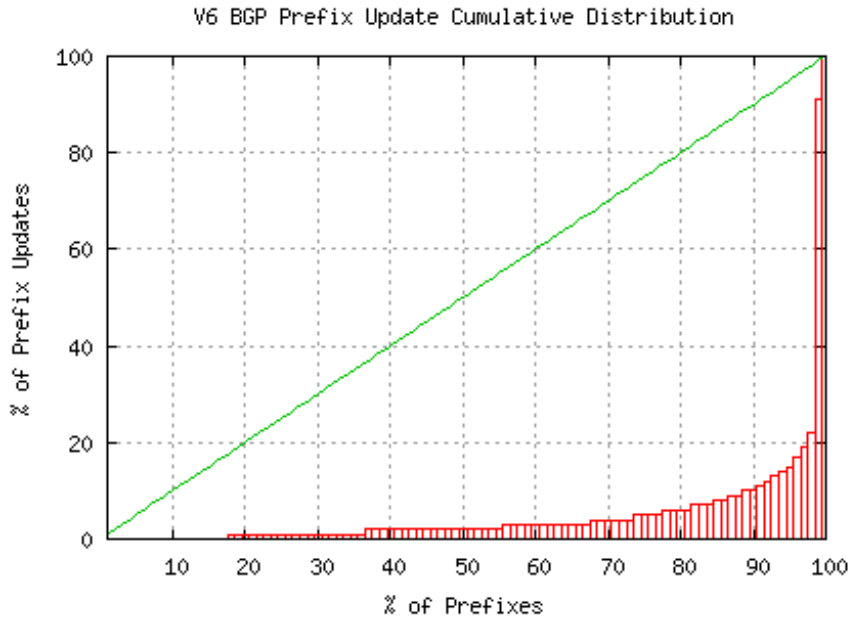
Its not the number of unstable prefixes, but the number of updates and elapsed time for the network to re-converge for each instability event

It this were to happen in the V4 network at the same relative scale it would be a major stability problem!

So what is going on and why has this happened?

Updates in IPv6

BGP Route Updates are very unequally distributed across the prefix set – they appear to affect a very small number of prefixes which stand out well above the average



Updates in IPv6

RANK	PREFIX	UPDs	%	Origin AS -- AS NAME
1	2c0f:fe90::/32	48777	8.99%	36943 -- webafrica,ZA
2	2407:8c00:ffe1::/48	39060	7.20%	131317 -- TTSLMEIS-IN TTSL-ISP DIVISION,IN
3	2804:14d:5a82::/48	34537	6.37%	28573 -- CLARO S.A.,BR
4	2605:cf00::/32	20020	3.69%	46525 -- RURALWAVE-LTD - Rural Wave,CA
5	2403:8600:ea89::/48	19448	3.58%	131317 -- TTSLMEIS-IN TTSL-ISP DIVISION,IN
6	2001:470:e9::/48	16613	3.06%	25104 -- WEBCODE WebCode Ltd.,BG
7	2804:39c:7000::/36	12787	2.36%	28329 -- G8 NETWORKS LTDA,BR
9	2001:7fb:ff02::/48	11097	2.05%	12654 -- RIPE-NCC-RIS-AS Reseaux IP Europeens Network Coordination Centre (RIPE NCC),EU
10	2607:f870:1::/48	9530	1.76%	11992 -- CENTENNIAL-PR - Centennial de Puerto Rico,PR
11	2001:67c:370::/48	9383	1.73%	56554 -- IETF-MEETING Internet Society,CH
12	2001:67c:1230::/46	9274	1.71%	56554 -- IETF-MEETING Internet Society,CH
14	2001:df8::/32	8911	1.64%	56554 -- IETF-MEETING Internet Society,CH
15	2a02:28c8::/32	8208	1.51%	42353 -- SIMWOOD Simwood eSMS Limited,GB
16	2620:27:f::/48	6064	1.12%	10846 -- DEERE - Deere & Company,US
17	2804:14d:8085::/48	5703	1.05%	28573 -- CLARO S.A.,BR
19	2a03:4600::/32	5135	0.95%	44334 -- RTLNET-ASN RTLNET,FR
20	2804:14d:8080::/48	4992	0.92%	28573 -- CLARO S.A.,BR

The busiest 48 prefixes accounted for 2/3 of all prefix updates

In Summary:

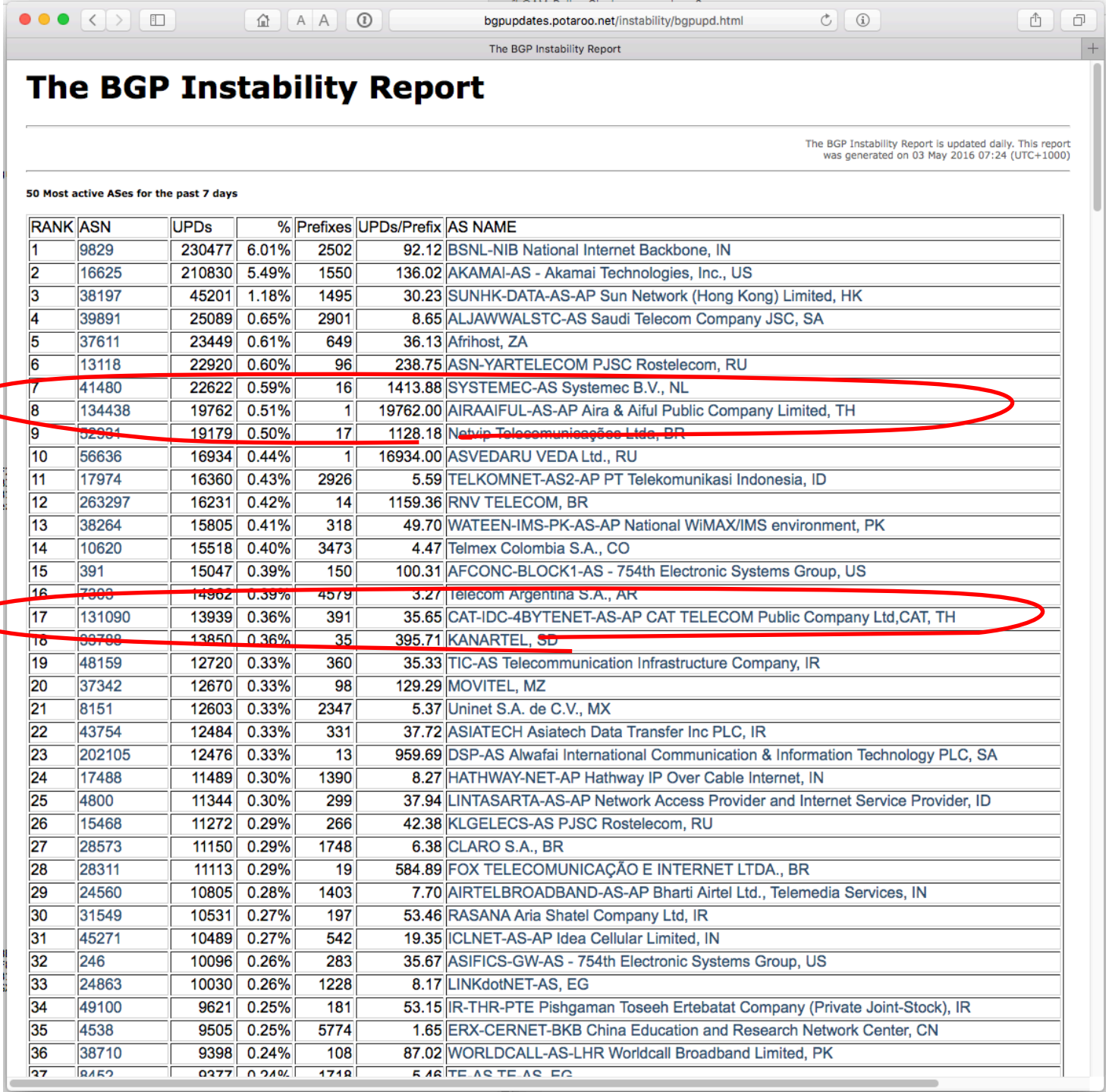
- The routing system continues to scale
  - Update rates are steady
  - Convergence times are steady
  - Instability is highly concentrated (a small number of prefixes generate the overwhelming majority of update traffic)
  - Which is great!

But

- We have no idea why it is so stable!
- Nor do we understand what will kick the system into instability

What can you do?

What can you do?

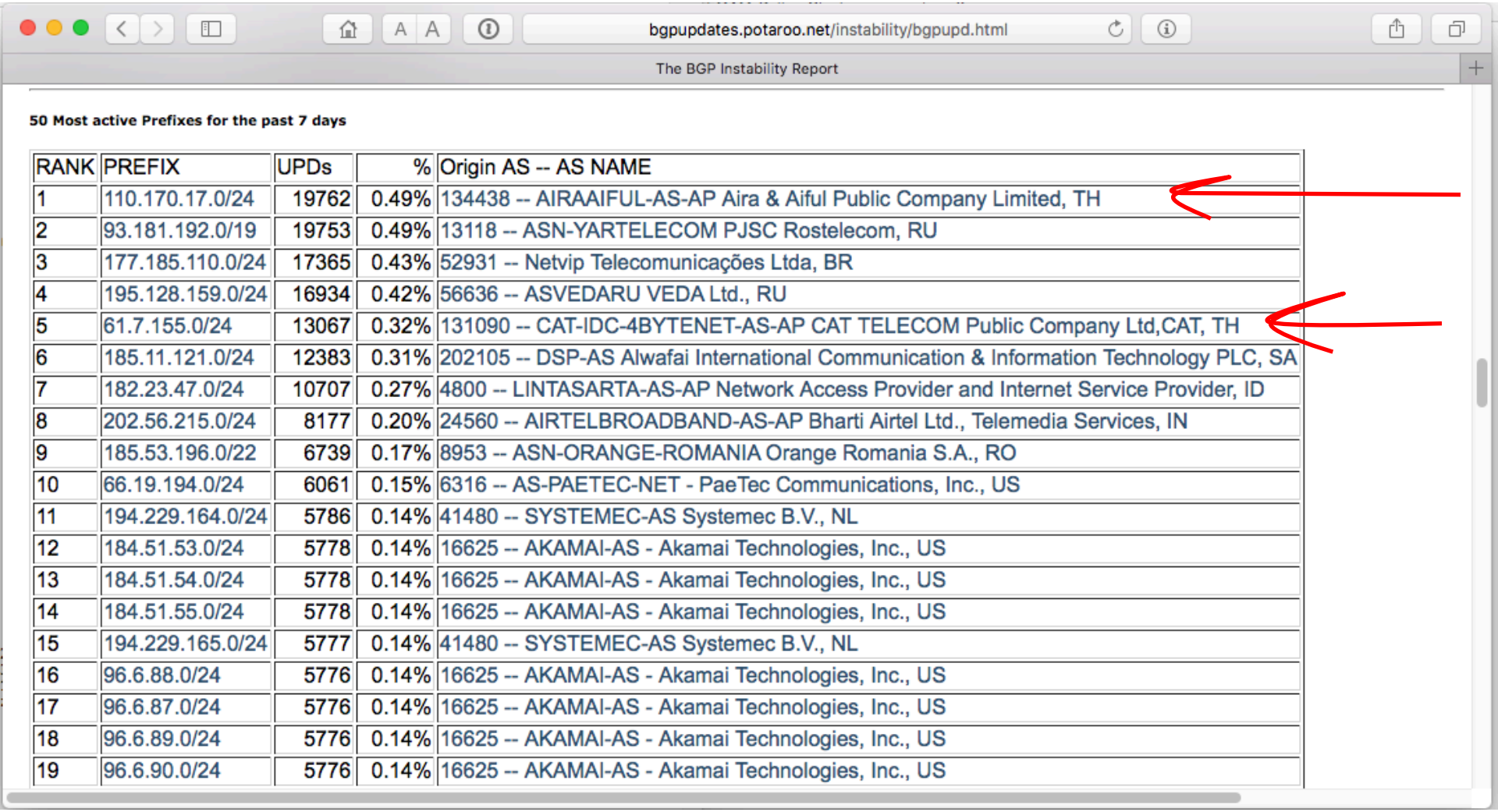


The BGP Instability Report

The BGP Instability Report is updated daily. This report was generated on 03 May 2016 07:24 (UTC+1000)

50 Most active ASes for the past 7 days

RANK	ASN	UPDs	%	Prefixes	UPDs/Prefix	AS NAME
1	9829	230477	6.01%	2502	92.12	BSNL-NIB National Internet Backbone, IN
2	16625	210830	5.49%	1550	136.02	AKAMAI-AS - Akamai Technologies, Inc., US
3	38197	45201	1.18%	1495	30.23	SUNHK-DATA-AS-AP Sun Network (Hong Kong) Limited, HK
4	39891	25089	0.65%	2901	8.65	ALJAWWALSTC-AS Saudi Telecom Company JSC, SA
5	37611	23449	0.61%	649	36.13	Afrihost, ZA
6	13118	22920	0.60%	96	238.75	ASN-YARTELECOM PJSC Rostelecom, RU
7	41480	22622	0.59%	16	1413.88	SYSTEMEC-AS Systemec B.V., NL
8	134438	19762	0.51%	1	19762.00	AIRAAIFUL-AS-AP Aira & Aiful Public Company Limited, TH
9	52931	19179	0.50%	17	1128.18	Netvip Telecomunicações Ltda, BR
10	56636	16934	0.44%	1	16934.00	ASVEDARU VEDA Ltd., RU
11	17974	16360	0.43%	2926	5.59	TELKOMNET-AS2-AP PT Telekomunikasi Indonesia, ID
12	263297	16231	0.42%	14	1159.36	RNV TELECOM, BR
13	38264	15805	0.41%	318	49.70	WATEEN-IMS-PK-AS-AP National WiMAX/IMS environment, PK
14	10620	15518	0.40%	3473	4.47	Telmex Colombia S.A., CO
15	391	15047	0.39%	150	100.31	AFCONC-BLOCK1-AS - 754th Electronic Systems Group, US
16	7303	14962	0.39%	4579	3.27	Telecom Argentina S.A., AR
17	131090	13939	0.36%	391	35.65	CAT-IDC-4BYTENET-AS-AP CAT TELECOM Public Company Ltd,CAT, TH
18	33788	13850	0.36%	35	395.71	KANARTEL, SD
19	48159	12720	0.33%	360	35.33	TIC-AS Telecommunication Infrastructure Company, IR
20	37342	12670	0.33%	98	129.29	MOVITEL, MZ
21	8151	12603	0.33%	2347	5.37	Uninet S.A. de C.V., MX
22	43754	12484	0.33%	331	37.72	ASIATECH Asiatech Data Transfer Inc PLC, IR
23	202105	12476	0.33%	13	959.69	DSP-AS Alwafai International Communication & Information Technology PLC, SA
24	17488	11489	0.30%	1390	8.27	HATHWAY-NET-AP Hathway IP Over Cable Internet, IN
25	4800	11344	0.30%	299	37.94	LINTASARTA-AS-AP Network Access Provider and Internet Service Provider, ID
26	15468	11272	0.29%	266	42.38	KLGELECS-AS PJSC Rostelecom, RU
27	28573	11150	0.29%	1748	6.38	CLARO S.A., BR
28	28311	11113	0.29%	19	584.89	FOX TELECOMUNICAÇÃO E INTERNET LTDA., BR
29	24560	10805	0.28%	1403	7.70	AIRTELBROADBAND-AS-AP Bharti Airtel Ltd., Telemedia Services, IN
30	31549	10531	0.27%	197	53.46	RASANA Aria Shatel Company Ltd, IR
31	45271	10489	0.27%	542	19.35	ICLNET-AS-AP Idea Cellular Limited, IN
32	246	10096	0.26%	283	35.67	ASIFICS-GW-AS - 754th Electronic Systems Group, US
33	24863	10030	0.26%	1228	8.17	LINKdotNET-AS, EG
34	49100	9621	0.25%	181	53.15	IR-THR-PTE Pishgaman Toseeh Ertebatat Company (Private Joint-Stock), IR
35	4538	9505	0.25%	5774	1.65	ERX-CERNET-BKB China Education and Research Network Center, CN
36	38710	9398	0.24%	108	87.02	WORLDCALL-AS-LHR Worldcall Broadband Limited, PK
37	8452	9377	0.24%	1718	5.46	TE-AS TE-AS, EG



50 Most active Prefixes for the past 7 days

RANK	PREFIX	UPDs	%	Origin AS -- AS NAME
1	110.170.17.0/24	19762	0.49%	134438 -- AIRAAIFUL-AS-AP Aira & Aiful Public Company Limited, TH
2	93.181.192.0/19	19753	0.49%	13118 -- ASN-YARTELECOM PJSC Rostelecom, RU
3	177.185.110.0/24	17365	0.43%	52931 -- Netvip Telecomunicações Ltda, BR
4	195.128.159.0/24	16934	0.42%	56636 -- ASVEDARU VEDA Ltd., RU
5	61.7.155.0/24	13067	0.32%	131090 -- CAT-IDC-4BYTENET-AS-AP CAT TELECOM Public Company Ltd,CAT, TH
6	185.11.121.0/24	12383	0.31%	202105 -- DSP-AS Alwafai International Communication & Information Technology PLC, SA
7	182.23.47.0/24	10707	0.27%	4800 -- LINTASARTA-AS-AP Network Access Provider and Internet Service Provider, ID
8	202.56.215.0/24	8177	0.20%	24560 -- AIRTELBROADBAND-AS-AP Bharti Airtel Ltd., Telemedia Services, IN
9	185.53.196.0/22	6739	0.17%	8953 -- ASN-ORANGE-ROMANIA Orange Romania S.A., RO
10	66.19.194.0/24	6061	0.15%	6316 -- AS-PAETEC-NET - PaeTec Communications, Inc., US
11	194.229.164.0/24	5786	0.14%	41480 -- SYSTEMEC-AS Systemec B.V., NL
12	184.51.53.0/24	5778	0.14%	16625 -- AKAMAI-AS - Akamai Technologies, Inc., US
13	184.51.54.0/24	5778	0.14%	16625 -- AKAMAI-AS - Akamai Technologies, Inc., US
14	184.51.55.0/24	5778	0.14%	16625 -- AKAMAI-AS - Akamai Technologies, Inc., US
15	194.229.165.0/24	5777	0.14%	41480 -- SYSTEMEC-AS Systemec B.V., NL
16	96.6.88.0/24	5776	0.14%	16625 -- AKAMAI-AS - Akamai Technologies, Inc., US
17	96.6.87.0/24	5776	0.14%	16625 -- AKAMAI-AS - Akamai Technologies, Inc., US
18	96.6.89.0/24	5776	0.14%	16625 -- AKAMAI-AS - Akamai Technologies, Inc., US
19	96.6.90.0/24	5776	0.14%	16625 -- AKAMAI-AS - Akamai Technologies, Inc., US

The BGP IPv6 Instability Report

This report is updated daily. The current report was generated on 4 May 2016 01:10 (UTC+1000)

50 Most active ASes for the past 31 days

RANK	ASN	UPDs	%	Prefixes	UPDs/Prefix	AS NAME
1	28573	130833	14.55%	380	344.30	CLARO S.A., BR
2	2611	63209	7.03%	24	2633.71	BELNET BELNET, BE
3	19210	34618	3.85%	3	11539.33	SYNCHRONOSS - SYNCHRONOSS TECHNOLOGIES, US
4	131317	31003	3.45%	5	6200.60	TTSLMEIS-IN TTSL-ISP DIVISION, IN
5	37002	22749	2.53%	10	2274.90	Reunicable, RE
6	12654	21575	2.40%	31	695.97	RIPE-NCC-RIS-AS Reseaux IP Europeens Network Coordination Centre (RIPE NCC), EU
7	55441	17701	1.97%	13	1361.62	TATA-DOCOMO-AS-AP D 26/2 TTC INDUSTRIAL AREA MIDC SANPADA, IN
8	263292	17584	1.96%	1	17584.00	L E M TELECOMUNICAÇÕES LTDA -ME, BR
9	28135	17020	1.89%	8	2127.50	ASSOCIAÇÃO NACIONAL PARA INCLUSÃO DIGITAL - ANID, BR
10	9835	16712	1.86%	11	1519.27	GITS-TH-AS-AP Government Information Technology Services, TH
11	1837	14262	1.59%	4	3565.50	NCIA NATO Communications and Information Agency, NL
12	200535	13836	1.54%	17	813.88	AS-LEANOR Leanor Communications Ltd., GB
13	36943	11146	1.24%	1	11146.00	webafrica, ZA
14	41741	10393	1.16%	1	10393.00	BBS-AS BBS AS, NO
15	20940	10298	1.15%	334	30.83	AKAMAI-ASN1 Akamai International B.V., US
16	32629	10054	1.12%	2	5027.00	CITY-OF-CHARLOTTE-ASN - City of Charlotte, US
17	16326	7482	0.83%	1	7482.00	SMALS-MVM SmalS-MvM, BE
18	31511	6877	0.77%	1	6877.00	B-RAIL-BE-AS Infrabel NV van publiek recht, BE
19	203338	6750	0.75%	1	6750.00	DDBE-INTERNET SA DIMENSION DATA BELGIUM, BE
20	3382	6528	0.73%	1	6528.00	ERX-JUITA-UINET University of Indonesia, ID
21	56538	6224	0.69%	3	2074.67	ASN-NETACCESS Hotwire, NO
22	11992	5906	0.66%	2	2953.00	CENTENNIAL-PR - Centennial de Puerto Rico, PR
23	204055	5734	0.64%	1	5734.00	CENTRALNIC-ANYCAST-H CentralNic Ltd, GB
24	201303	5733	0.64%	1	5733.00	CENTRALNIC-ANYCAST-F CentralNic Ltd, GB
25	199330	5494	0.61%	1	5494.00	CENTRALNIC-ANYCAST-A CentralNic Ltd, GB
26	17451	5469	0.61%	65	84.14	BIZNET-AS-AP BIZNET NETWORKS, ID
27	39288	5244	0.58%	1	5244.00	CZUBYPL-NET Nette sp. z o.o., PL
28	46047	4749	0.53%	1	4749.00	POLSRI-AS-ID Politeknik Negeri Sriwijaya, ID
29	9829	4465	0.50%	40	111.62	BSNL-NIB National Internet Backbone, IN
30	56237	4093	0.46%	1	4093.00	UNILA-AS-ID Universitas Lampung, ID

RANK	PREFIX	UPDs	%	Origin AS -- AS NAME
1	2403:8600:ea89::/48	37379	3.92%	131317 -- TTSLMEIS-IN TTSL-ISP DIVISION, IN 55441 -- TATA-DOCOMO-AS-AP D 26/2 TTC INDUSTRIAL AREA MIDC SANPADA, IN
2	2620:11b:4000::/48	34612	3.63%	19210 -- SYNCHRONOSS - SYNCHRONOSS TECHNOLOGIES, US
3	2804:14d::/40	31325	3.28%	28573 -- CLARO S.A., BR
4	2804:14d:8085::/48	31321	3.28%	28573 -- CLARO S.A., BR
5	2804:14d:8080::/48	30244	3.17%	28573 -- CLARO S.A., BR
6	2c0f:f918:106::/48	22704	2.38%	37002 -- Reunicable, RE
7	2804:1690::/32	17584	1.84%	263292 -- L E M TELECOMUNICAÇÕES LTDA -ME, BR
8	2804:88:a000::/44	16826	1.76%	28135 -- ASSOCIAÇÃO NACIONAL PARA INCLUSÃO DIGITAL - ANID, BR
9	2804:14d:8000::/40	12992	1.36%	28573 -- CLARO S.A., BR
10	2407:8c00:ffe1::/48	11156	1.17%	131317 -- TTSLMEIS-IN TTSL-ISP DIVISION, IN
11	2c0f:fe90::/32	11146	1.17%	36943 -- webafrica, ZA
12	2a03:6180::/32	10393	1.09%	41741 -- BBS-AS BBS AS, NO
13	2620:0:2f0::/48	10051	1.05%	32629 -- CITY-OF-CHARLOTTE-ASN - City of Charlotte, US
14	2001:7fb:ff02::/48	9249	0.97%	12654 -- RIPE-NCC-RIS-AS Reseaux IP Europeens Network Coordination Centre (RIPE NCC), EU
15	2a02:2c40::/32	9022	0.95%	2611 -- BELNET BELNET, BE
16	2401:9d00:1112::/48	8408	0.88%	9835 -- GITS-TH-AS-AP Government Information Technology Services, TH
17	2401:9d00:1111::/48	8232	0.86%	9835 -- GITS-TH-AS-AP Government Information Technology Services, TH
18	2804:14c:8583::/48	7918	0.83%	28573 -- CLARO S.A., BR
19	2a01:690::/29	7482	0.78%	16326 -- SMALS-MVM SmalS-MvM, BE
20	2001:67c:1a17::/48	7110	0.75%	1837 -- NCIA NATO Communications and Information Agency, NL
21	2001:67c:1a10::/45	7108	0.75%	1837 -- NCIA NATO Communications and Information Agency, NL
22	2a04:4720::/30	6877	0.72%	31511 -- B-RAIL-BE-AS Infrabel NV van publiek recht, BE
23	2a07:d40::/29	6750	0.71%	203338 -- DDBE-INTERNET SA DIMENSION DATA BELGIUM, BE
24	2a02:26f0:70::/48	6675	0.70%	20940 -- AKAMAI-ASN1 Akamai International B.V., US
25	2001:6a8::/32	6650	0.70%	2611 -- BELNET BELNET, BE
26	2a00:b2c0:fade::/48	6613	0.69%	2611 -- BELNET BELNET, BE
27	2a04:b5c0::/29	6611	0.69%	2611 -- BELNET BELNET, BE
28	2403:da00::/32	6528	0.68%	3382 -- ERX-JUITA-UINET University of Indonesia, ID
29	2607:f870:1::/48	5899	0.62%	11992 -- CENTENNIAL-PR - Centennial de Puerto Rico, PR

That's it!