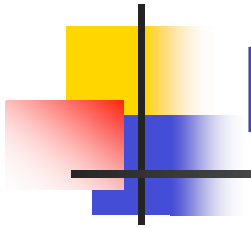
Internet Routing Table Analysis Update

Philip Smith

SANOG 14

15 - 23 July 2009

Chennai



Motivation

- 1998: No one was publishing any Internet routing table analysis
  - Only CIDR-Report reporting on top 20 contributors to routing table, and top 20 bad aggregators
- With support of APNIC, my weekly reporting report started 23rd February 1999:
  - Routing table size
  - CIDR-Report style reporting on a per-RIR basis
  - …and many other interesting features

Routing Report 20 July 2009

BGP routing table entries examined:	291223
Prefixes after maximum aggregation:	138126
Deaggregation factor:	2.11
Unique aggregates announced to Internet:	144846
Total ASes present in the Internet Routing Table:	31771
Prefixes per ASN:	9.17
Origin-only ASes present in the Internet Routing Table:	27636
Origin ASes announcing only one prefix:	13457
Transit ASes present in the Internet Routing Table:	4135
Transit-only ASes present in the Internet Routing Table:	96
Average AS path length visible in the Internet Routing Table:	3.6
Max AS path length visible:	39
Max AS path prepend of ASN (22394)	36
Prefixes from unregistered ASNs in the Routing Table:	418
Unregistered ASNs in the Routing Table:	122
Number of 32-bit ASNs allocated by the RIRs:	207
Prefixes from 32-bit ASNs in the Routing Table:	67
Special use prefixes present in the Routing Table:	0
Prefixes being announced from unallocated address space:	337
Number of addresses announced to Internet:	2071531008
Equivalent to 123 /8s, 121 /16s and 14 /24s	
Percentage of available address space announced:	55.9
Percentage of allocated address space announced:	64.7
Percentage of available address space allocated:	86.4
Percentage of address space in use by end-sites:	78.3
Total number of prefixes smaller than registry allocations:	144160

APNIC Region

Prefixes being announced by APNIC Region ASes:	69558
Total APNIC prefixes after maximum aggregation:	24774
APNIC Deaggregation factor:	2.81
Prefixes being announced from the APNIC address blocks:	68975
Unique aggregates announced from the APNIC address blocks:	31484
APNIC Region origin ASes present in the Internet Routing Table:	3708
APNIC Prefixes per ASN:	18.60
APNIC Region origin ASes announcing only one prefix:	1010
APNIC Region transit ASes present in the Internet Routing Table:	571
Average APNIC Region AS path length visible:	3.5
Max APNIC Region AS path length visible:	16
Number of APNIC addresses announced to Internet:	469442240
Equivalent to 27 /8s, 251 /16s and 30 /24s	
Percentage of available APNIC address space announced:	87.4

APNIC AS Blocks	4608-4864, 7467-7722, 9216-10239, 17408-18431
(pre-ERX allocations)	23552-24575, 37888-38911, 45056-46079
APNIC Address Blocks	58/8, 59/8, 60/8, 61/8, 110/8, 111/8, 112/8, 113/8, 114/8, 115/8, 116/8, 117/8, 118/8, 119/8, 120/8, 121/8, 122/8, 123/8, 124/8, 125/8, 126/8, 180/8, 183/8, 202/8, 203/8, 210/8, 211/8, 218/8, 219/8, 220/8, 221/8, 222/8

Global per AS prefix count summary

ASN	No of nets	/20 equiv	Max Agg	Description
6389	4244	3643	323	bellsouth.net, inc.
4323	1892	1048	385	Time Warner Telecom
1785	1710	714	139	PaeTec Communications, Inc.
4766	1701	6934	407	Korea Telecom (KIX)
17488	1542	137	103	Hathway IP Over Cable Interne
7018	1508	5909	1047	AT&T WorldNet Services
8151	1471	2882	231	UniNet S.A. de C.V.
20115	1415	1460	681	Charter Communications
2386	1268	670	919	AT&T Data Communications Serv
4755	1218	292	144	TATA Communications formerly
6478	1217	275	311	AT&T Worldnet Services
3356	1194	10961	444	Level 3 Communications, LLC
11492	1127	208	12	Cable One
9583	1126	87	559	Sify Limited
22773	1077	2604	66	Cox Communications, Inc.
18566	1062	296	10	Covad Communications
8452	1027	188	7	TEDATA
19262	1019	4091	237	Verizon Global Networks
4134	990	17290	375	CHINANET-BACKBONE
7011	988	240	570	Citizens Utilities

Number of prefixes announced by prefix length

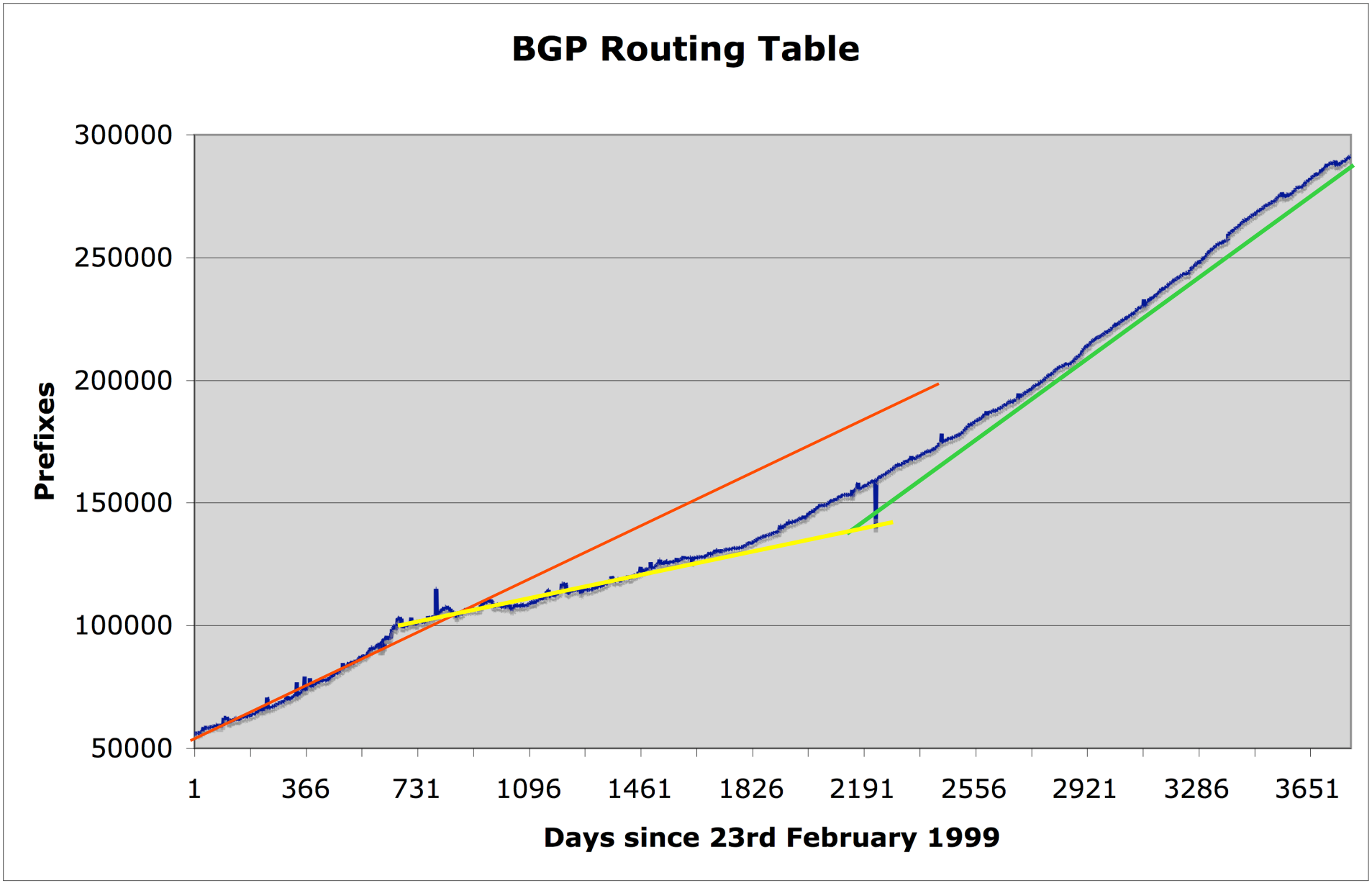
/1:0	/2:0	/3:0	/4:0	/5:0	/6:0
/7:0	/8:19	/9:10	/10:23	/11:58	/12:169
/13:350	/14:612	/15:1164	/16:10534	/17:4751	/18:8224
/19:17157	/20:20488	/21:20358	/22:26149	/23:26077	/24:152383
/25:930	/26:1030	/27:560	/28:153	/29:8	/30:7
/31:1	/32:8				

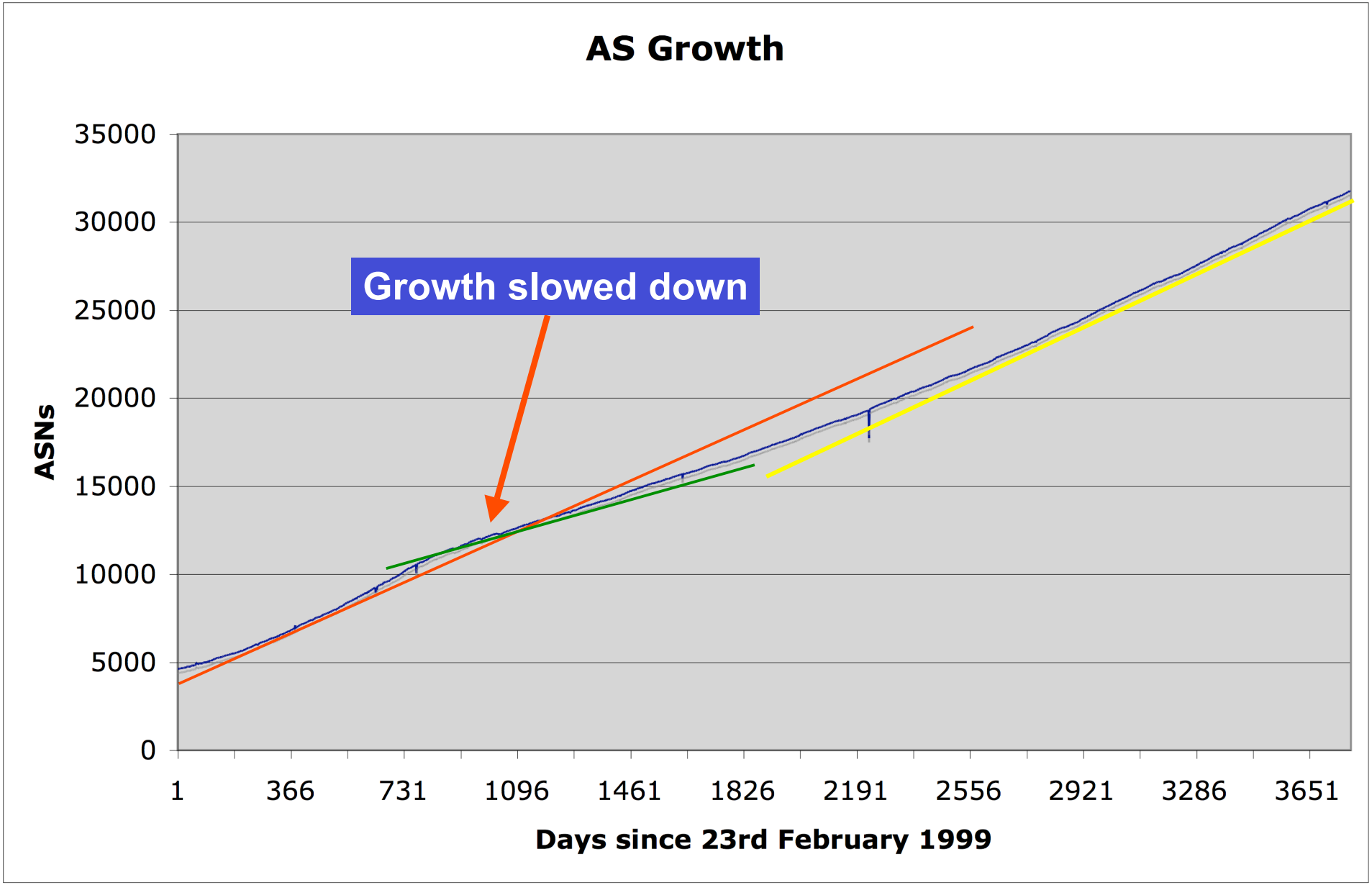
July 2009 ↑

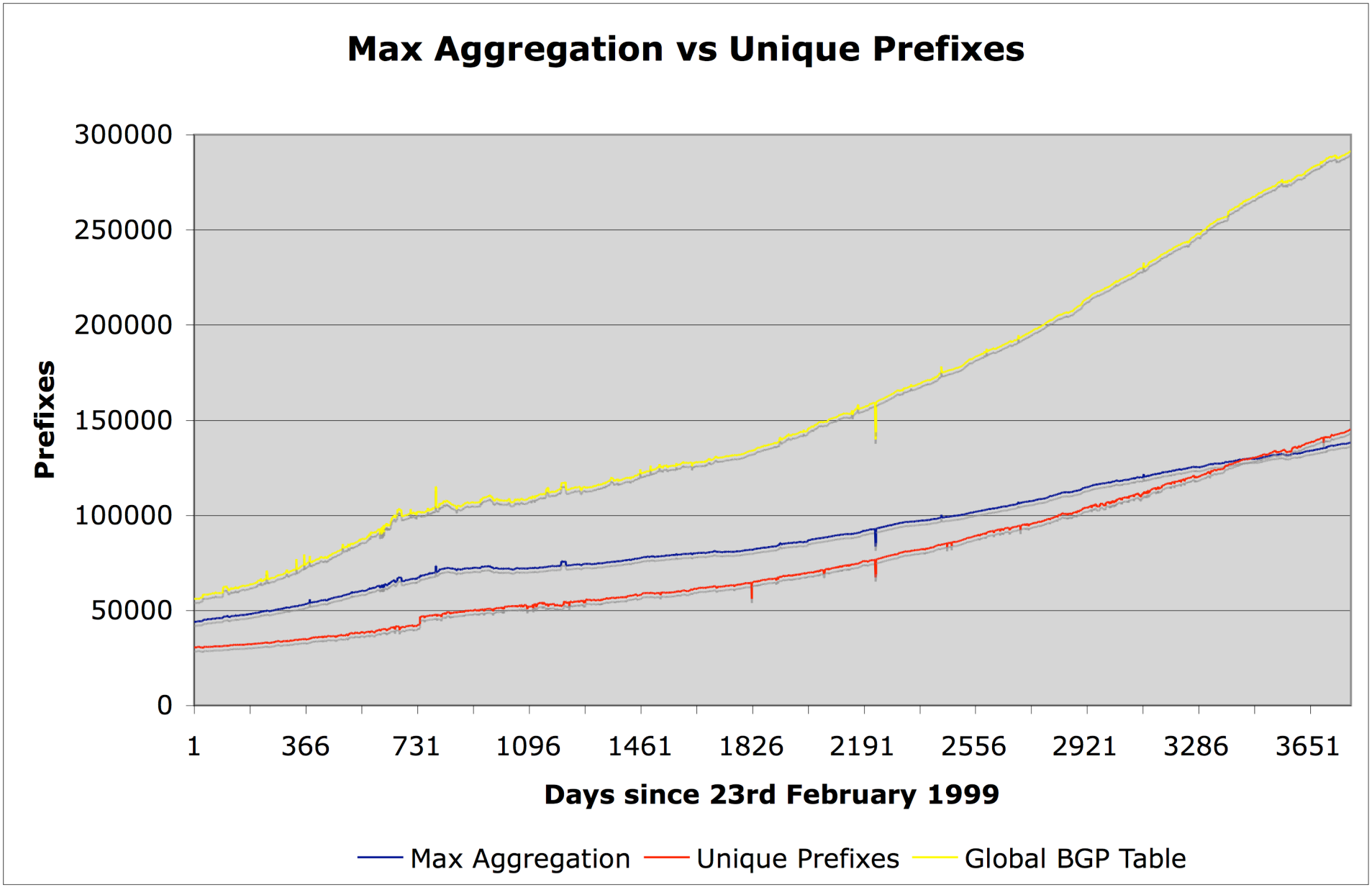
July 2008 ↓

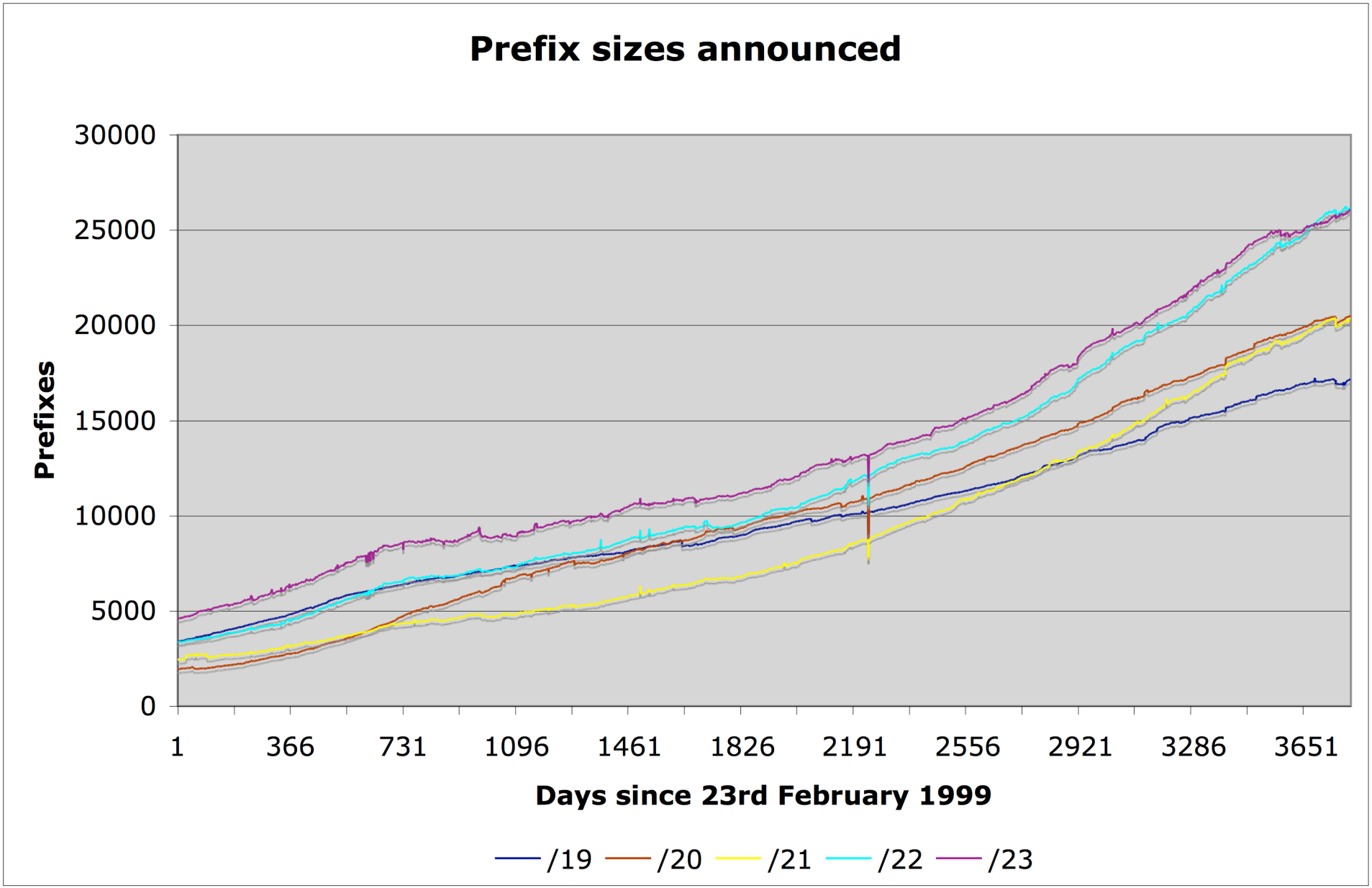
Number of prefixes announced by prefix length

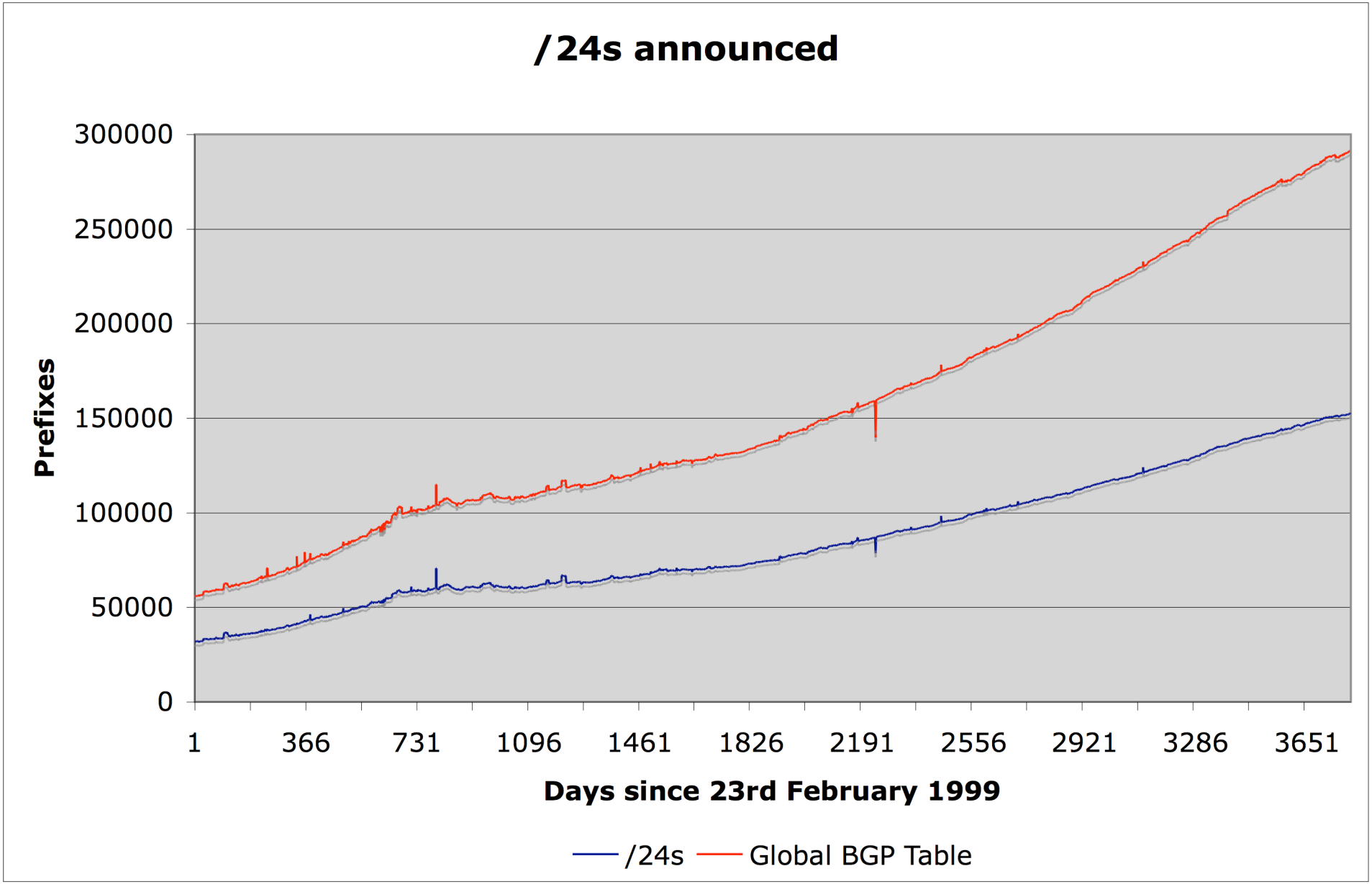
/1:0	/2:0	/3:0	/4:0	/5:0	/6:0
/7:0	/8:19	/9:9	/10:16	/11:45	/12:146
/13:288	/14:517	/15:1044	/16:10011	/17:4334	/18:7518
/19:15878	/20:18522	/21:18197	/22:22727	/23:23638	/24:137874
/25:843	/26:1029	/27:773	/28:81	/29:7	/30:1
/31:0	/32:6				

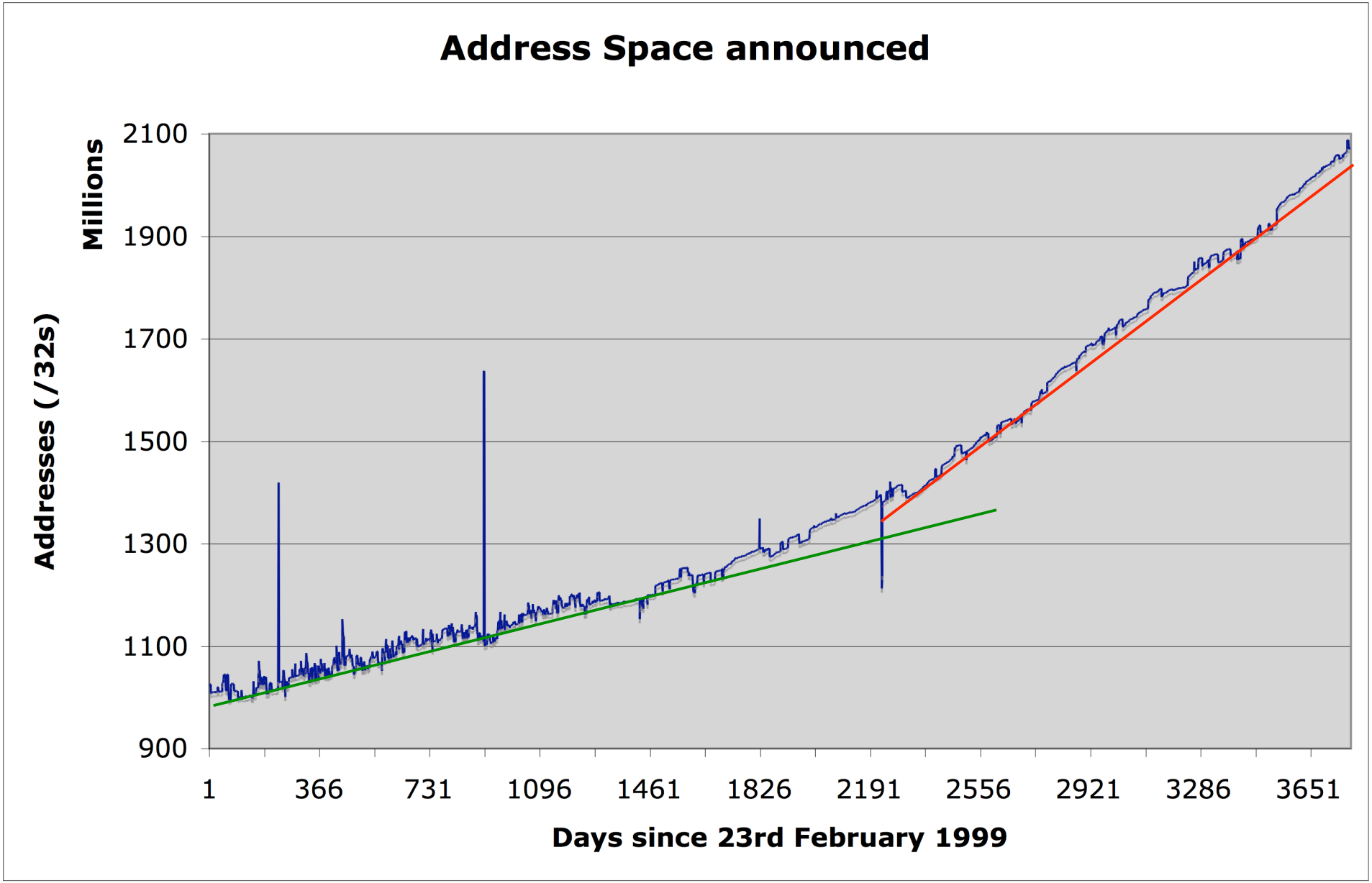

BGP Routing Table
300000
250000
200000
150000
100000
50000
Prefixes
1
366
731
1096
1461
1826
2191
2556
2921
3286
3651
Days since 23rd February 1999

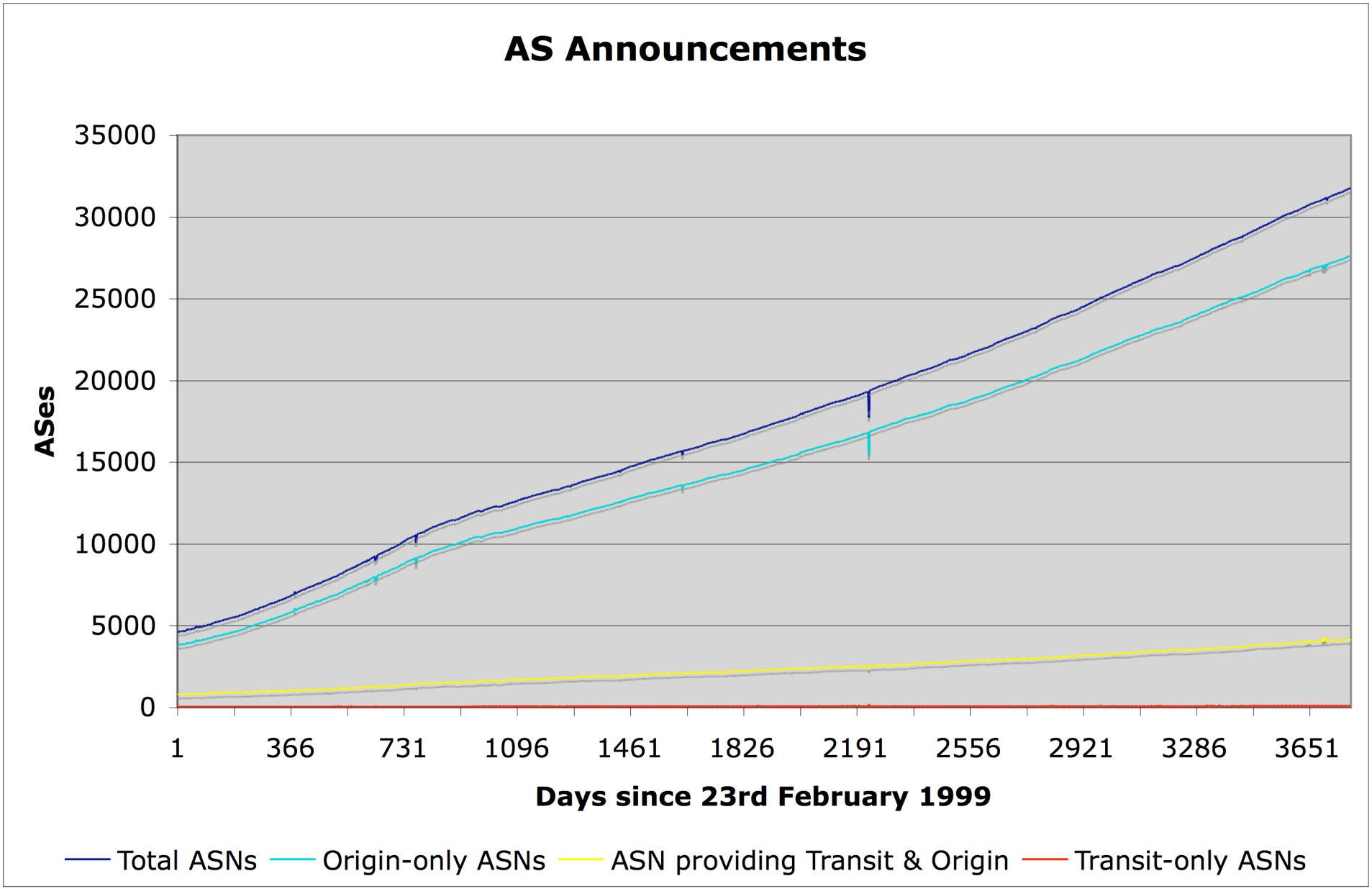

AS Growth
35000
30000
25000
20000
15000
10000
5000
0
ASNs
Growth slowed down
1
366
731
1096
1461
1826
2191
2556
2921
3286
3651
Days since 23rd February 1999

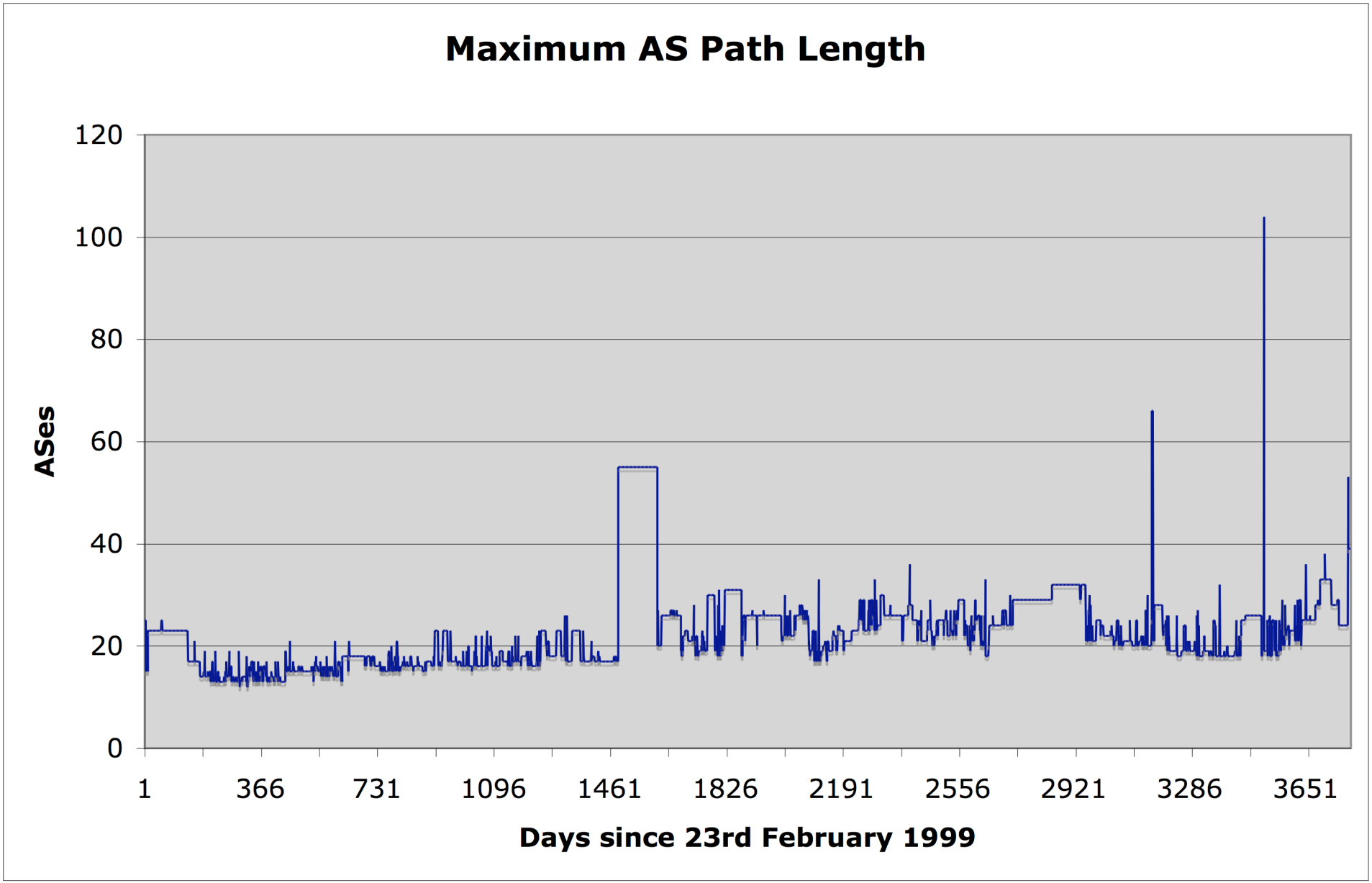

Max Aggregation vs Unique Prefixes
300000
250000
200000
150000
100000
50000
0
Prefixes
1
366
731
1096
1461
1826
2191
2556
2921
3286
3651
Days since 23rd February 1999
Max Aggregation
Unique Prefixes
Global BGP Table

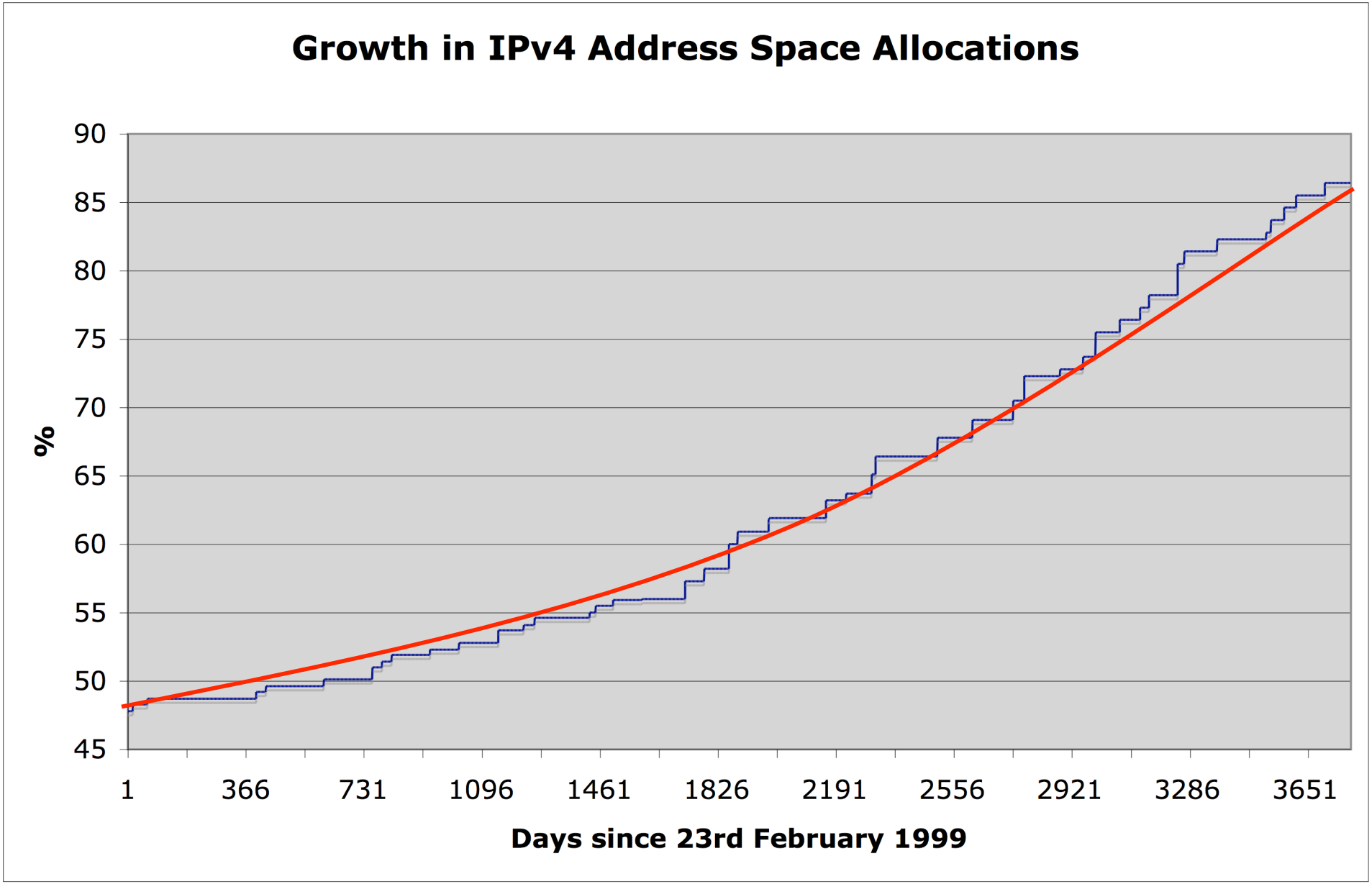

Prefix sizes announced
Prefixes
30000
25000
20000
15000
10000
5000
0
1
366
731
1096
1461
1826
2191
2556
2921
3286
3651
Days since 23rd February 1999
/19
/20
/21
/22
/23


/24s announced
300000
250000
200000
150000
100000
50000
0
Prefixes
1
366
731
1096
1461
1826
2191
2556
2921
3286
3651
Days since 23rd February 1999
/24s
Global BGP Table

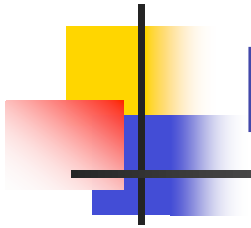

Address Space announced
Addresses (/32s)
Millions
2100
1900
1700
1500
1300
1100
900
1
366
731
1096
1461
1826
2191
2556
2921
3286
3651
Days since 23rd February 1999


AS Announcements
35000
30000
25000
20000
15000
10000
5000
0
ASes
1
366
731
1096
1461
1826
2191
2556
2921
3286
3651
Days since 23rd February 1999
Total ASNs
Origin-only ASNs
ASN providing Transit & Origin
Transit-only ASNs

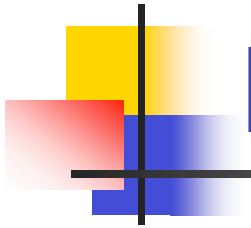

Maximum AS Path Length
ASes
120
100
80
60
40
20
0
1
366
731
1096
1461
1826
2191
2556
2921
3286
3651
Days since 23rd February 1999


Growth in IPv4 Address Space Allocations
90
85
80
75
70
65
60
55
50
45
%
1
366
731
1096
1461
1826
2191
2556
2921
3286
3651
Days since 23rd February 1999

Route Aggregation Recommendations

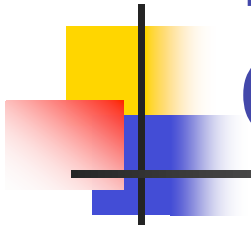


- RIPE Document — RIPE-399
  - http://www.ripe.net/ripe/docs/ripe-399.html
- Discusses:
  - History of aggregation
  - Causes of de-aggregation
  - Impacts on global routing system
  - Available Solutions
  - Recommendations for ISPs



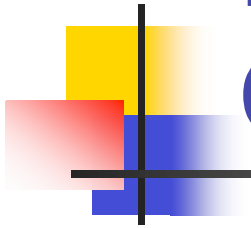
History:

- Classful to classless migration
  - Clean-up efforts in 192/8
- CIDR Report
  - Started by Tony Bates to encourage adoption of CIDR & aggregation
  - Mostly ignored through late 90s
  - Now part of extensive BGP table analysis by Geoff Huston
- Introduction of Regional Internet Registry system and PA address space



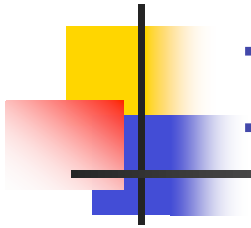
Deaggregation: Claimed causes (1):

- Routing System Security
  - "Announcing /24s means that no one else can DOS the network"
- Reduction of DOS attacks & miscreant activities
  - "Announcing only address space in use as rest attracts 'noise'"
- Commercial Reasons
  - "Mind your own business"



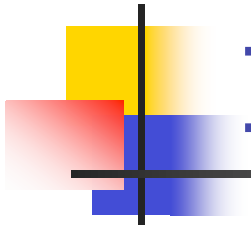
Deaggregation: Claimed causes (2):

- Leakage of iBGP outside of local AS
  - eBGP is NOT iBGP – how many ISPs know this?
- Traffic Engineering for Multihoming
  - Spraying out /24s hoping it will work
  - Rather than do any **real engineering**
- Legacy Assignments
  - "All those pre-RIR assignments are to blame"
  - In reality it is both RIR and legacy assignments



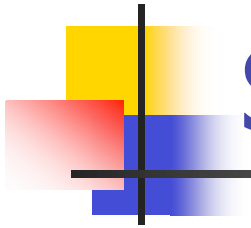
Impacts (1):

- Router memory
  - Shortens router life time as vendors underestimate memory growth requirements
  - Depreciation life-cycle shortened
  - Increased costs for ISP and customers
- Router processing power
  - Processors are underpowered as vendors underestimate CPU requirement
  - Depreciation life-cycle shortened
  - Increased costs for ISP and customers



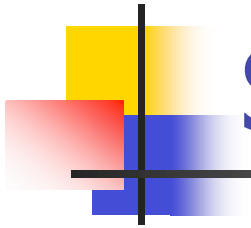
Impacts (2):

- Routing System convergence
  - Larger routing table → slowed convergence
  - Can be improved by faster control plane processors — see earlier
- Network Performance & Stability
  - Slowed convergence → slowed recovery from failure
  - Slowed recovery → longer downtime
  - Longer downtime → unhappy customers



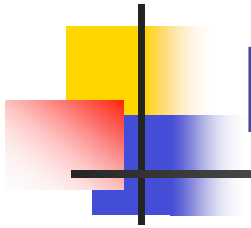
Solutions (1):

- CIDR Report
  - Global aggregation efforts
  - Running since 1994
- Routing Table Report
  - Per RIR region aggregation efforts
  - Running since 1999
- Filtering recommendations
  - BCP38, training, tutorials, Team Cymru,…
- “CIDR Police”



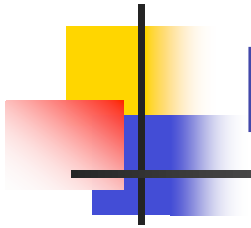
Solutions (2):

- BGP Features:
  - NO_EXPORT Community
  - NOPEER Community
    - RFC3765 — but no one has implemented it
  - AS_PATHLIMIT attribute
    - Still working through IETF IDR Working Group
  - Provider Specific Communities
    - Some ISPs use them; most do not



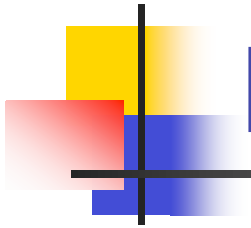
RIPE-399 Recommendations:

- Announcement of initial allocation as a single entity
- Subsequent allocations aggregated if they are contiguous and bit-wise aligned
- Prudent subdivision of aggregates for Multihoming
- Use BGP enhancements already discussed
- (All of this applies to IPv6 too)



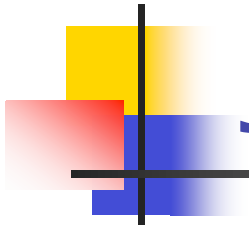
Looking at Deaggregation

- CIDR Report
  - www.cidr-report.org
  - Encourages aggregation following CIDRisation of Internet
  - Today: extensive suite of reports and tools covering state of BGP table
- Routing Report
  - BGP table status on per RIR basis
  - Original CIDR Report and a whole lot more



Deaggregation Factor

- Routing Report
  - One summary takes BGP table and aggregates prefixes by origin AS
    - Called "Max Aggregation" in report
  - Global and per RIR basis
    - http://thyme.apnic.net/current/
- New Deaggregation Factor:
  - Measure of Routing Table size/Aggregated Size
  - Global value has been increasing slowly and steadily since "records began"



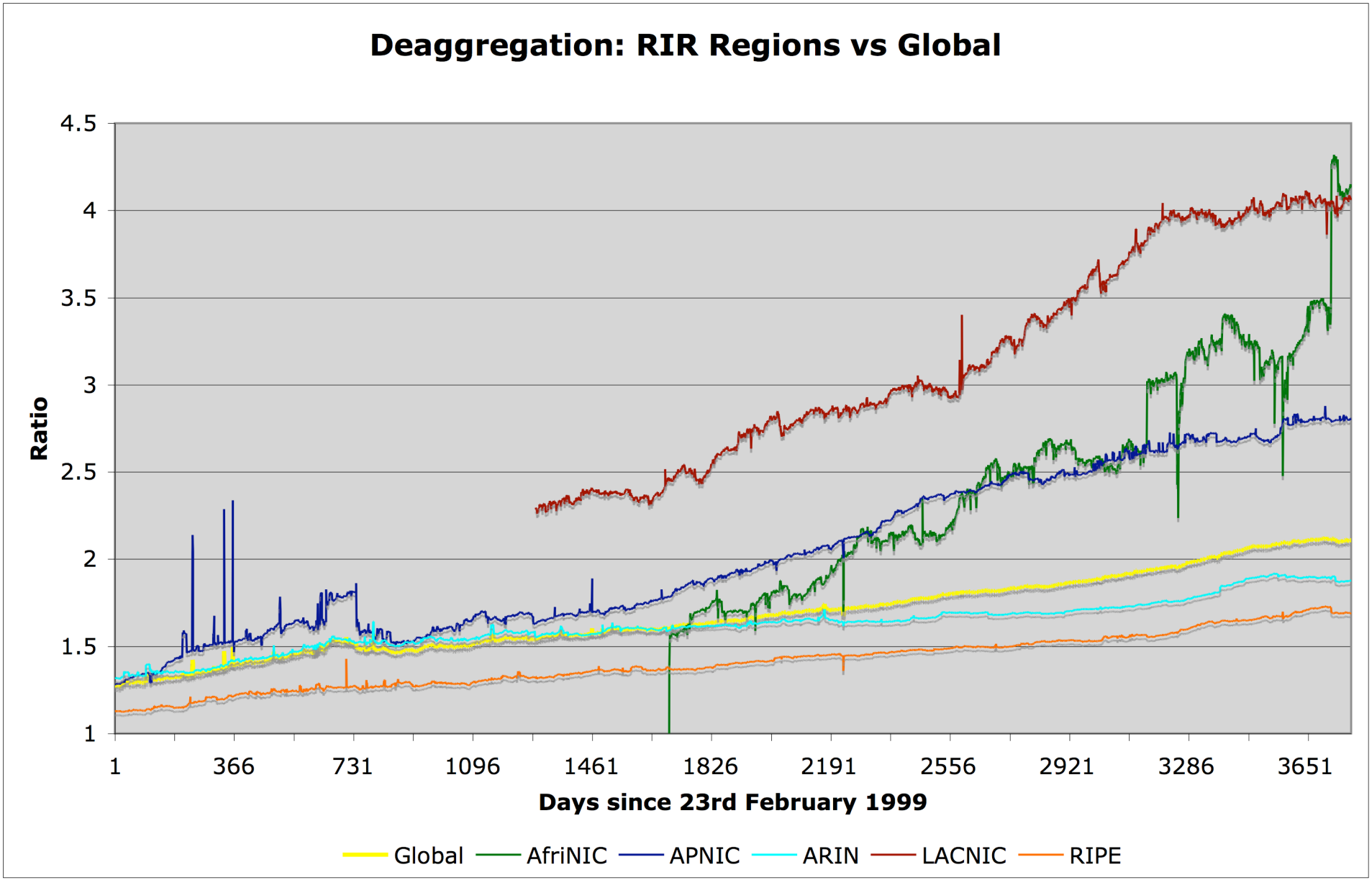
July 2009

Total Prefixes

- Global BGP Table
  - 291k prefixes
- Europe & Middle East
  - 67k prefixes
- North America
  - 124k prefixes
- Asia & Pacific
  - 70k prefixes
- Africa
  - 6k prefixes
- Latin America & Caribbean
  - 25k prefixes

Deaggregation Factor

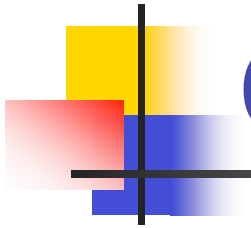
- Global Average
  - 2.11
- Europe & Middle East
  - 1.69
- North America
  - 1.88
- Asia & Pacific
  - 2.81
- Africa
  - 4.13
- Latin America & Caribbean
  - 4.07


Deaggregation: RIR Regions vs Global
Ratio
4.5
4
3.5
3
2.5
2
1.5
1
1
366
731
1096
1461
1826
2191
2556
2921
3286
3651
Days since 23rd February 1999
Global
AfriNIC
APNIC
ARIN
LACNIC
RIPE

Asia Pacific Aggregation Savings Summary

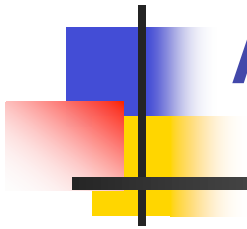
ASN	No of Nets	Savings	Description
17488	1542	1439	Hathway IP Over Cable Interne
4766	1701	1294	Korea Telecom (KIX)
4755	1218	1074	TATA Communications formerly
9829	800	786	BSNL National Internet Backbo
18101	749	717	Reliance Infocom Ltd Internet
7545	812	710	TPG Internet Pty Ltd
17908	697	650	Tata Communications
4134	990	615	CHINANET-BACKBONE
17974	698	604	PT TELEKOMUNIKASI INDONESIA
9498	630	583	BHARTI BT INTERNET LTD.
9583	1126	567	Sify Limited
24560	729	561	Bharti Airtel Ltd.
17676	564	503	Softbank BB Corp.
4808	666	498	CNCGROUP IP network: China169
4780	512	442	Digital United Inc.
9443	492	412	Primus Telecommunications
9808	406	397	Guangdong Mobile Communicatio
4802	517	348	Wantree Development
7643	349	341	VNPT
10091	349	338	SCV Broadband Access Provider

http://thyme.apnic.net/current/data-CIDRnet-APNIC



Observations

- Range of operational "practices" between RIR regions
  - Deaggregation by newer ISPs & developing regions is growing rapidly
  - Is harming the **entire** Internet
- RIPE-399 is only a recommendation
  - Hopefully all the RIRs will include pointers with each address allocation
  - Hopefully more ISPs will pay attention to it
  - Training is there — most ISPs choose to ignore it



Internet Routing Table Analysis Update

Questions?