The Next Three Years

(IPv6, IPv4 run-out and 4-byte ASNs)

Philip Smith **<pfs@cisco.com>**

NZNOG 2010

Hamilton

“Internet history”

- The Internet has been growing since the start!
- From early ‘90s, two efforts to scale – short-term versus long-term

  More at “The Long and Windy ROAD”
  http://rms46.vlsm.org/1/42.html
- Immediate enhancements to allow continued growth

  CIDR, Supernetting, RIRs, DHCP, PPP, NAT,…
- Long term work on next generation of IP

  IPv4 to replace IPv6

  Development work since 1995

Current Situation

- General perception is that “IPv6 has not yet taken hold”
  - More discussions plus IPv4 run-out plans proposed
  - Private sector asks for RoI/Business case to “migrate”
- But reality is very different from perception!
  - Something needs to be done to sustain the Internet growth
  - IPv6 or NAT or both or something else?

Status in Internet Operational Community

- Service Providers get an IPv6 prefix from their regional Internet registries
  - Very straight forward process when compared with IPv4
- Much discussion amongst operators about transition:
  - NOG experiments of 2008 – http://www.civil-tongue.net/6and4/
  - What is really still missing from IPv6 – http://www.nanog.org/mtg-0710/presentations/Bush-v6-op-reality.pdf
  - Many presentations on IPv6 deployment experiences
- Many Service Providers have made their backbones IPv6 capable
  - As part of ongoing infrastructure upgrades

OS, Services, Applications, Content

- Operating Systems
  - MacOS X, Linux, BSD Family, many SYS V
  - Windows: XP SP2 (hidden away), Vista, 7
  - All use IPv6 first if available
- Applications
  - Browsers, E-mail clients, IM, bittorrent,…
- Services
  - DNS, Apache WebServer, E-mail gateways,…
- Content Availability
  - Needs to be on IPv4 ***and*** on IPv6

The On-going Debate (1)

- IPv6 Multihoming
  - Same toolset as IPv4 — long term non-scalable
  - ‘Ultimate Multihoming Solution’ no nearer discovery
    - LISP is making interesting progress though
- Early rigid IPv6 address allocation model
  - “One size fits all” barrier to deployment:
    - Only ISPs “should” get IPv6 space from RIRs
    - Enterprises “should” get IPv6 space from ISPs only
  - Routing table entries matter, not the nature of business
    - What is an ISP?

The On-going Debate (2)

- Not every IPv4 device is IPv6 capable

  Do we really need to replicate all IPv4 capability in IPv6 prior to considering deployment?

- “We have enough IPv4”

  Those with plenty denying those with little/nothing

- Migration versus Co-existence

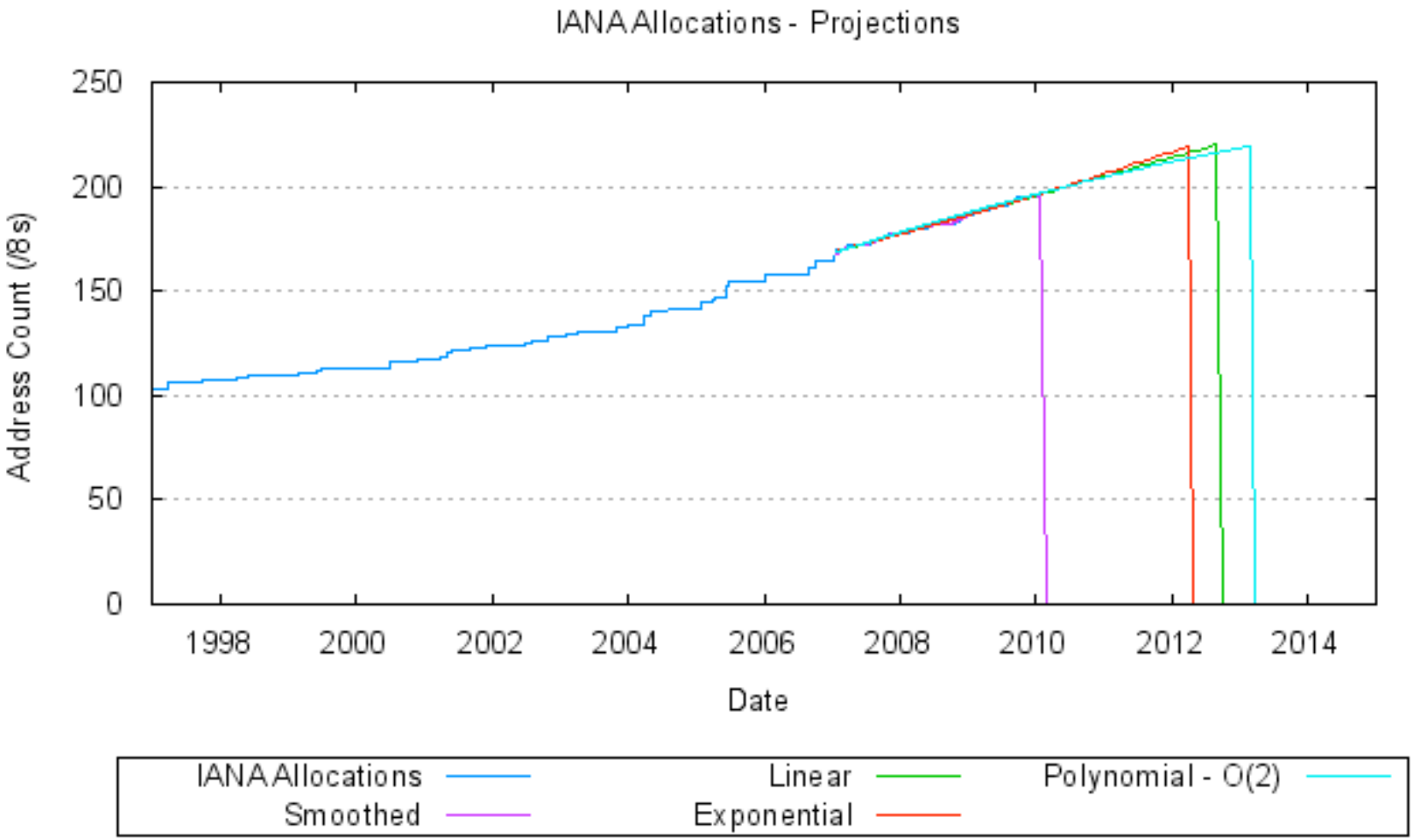
  Realistically IPv6 and IPv4 will co-exist for many years

  Dual-stack operating systems in network equipment makes this trivial

Why not use Network Address Translation?

- Private address space and Network address translation (NAT) could be used instead of IPv6
- But NAT has many serious issues:
  - Breaks the end-to-end model of IP
  - Breaks end-to-end network security
  - Serious consequences for Lawful Intercept
  - Non-NAT friendly applications means NAT has to be upgraded
  - Some applications don’t work through NATs
  - Layered NAT devices
  - Mandates that the network keeps the state of the connections
  - How to scale NAT performance for large networks??
  - Makes fast rerouting and multihoming difficult
  - How to offer content from behind a NAT?

Is IPv4 really running out?



http://www.potaroo.net/tools/ipv4/fig18.png

Is IPv4 really running out?

- Yes

  IANA IPv4 free pool runs out in September 2011

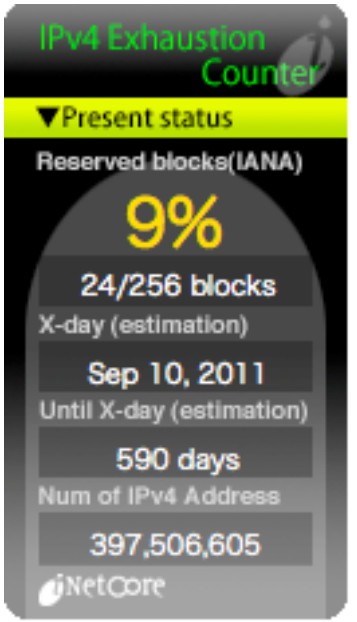
  RIR IPv4 free pool runs out approx one year later

  http://www.potaroo.net/tools/ipv4/

- Small industry producing gadgets and widgets predicting IPv4 run-out

  http://inetcore.com/project/ipv4ec/index_en.html

  http://ipv6.he.net/statistics/



IPv4 run-out

- RIR Policy Development process in each RIR region is now handling proposals relating to IPv4 run-out
  - The Last /8
    - All RIRs will receive one /8 from the IANA free pool
  - IPv4 address transfer
    - Permits LIRs to transfer address space to each other rather than returning to their RIR
  - Soft landing
    - Reduce the allocation sizes for an LIR as IPv4 pool is depleted
  - IPv4 distribution for IPv6 transition
    - Reserving a range of IPv4 address to assist with IPv6 transition (for Large Scale NATs etc)

Issues Today

- Minimal content is available on IPv6
  - Notwithstanding ipv6.google.com
- Giving IPv6 to customers might confuse
  - Browsers,e-mail clients, etc are smart
  - But increased tech support if IPv6 version of content is ‘down’, but IPv4 version works
- Need to “prolong” IPv4 so there is time for all content to be available on IPv6

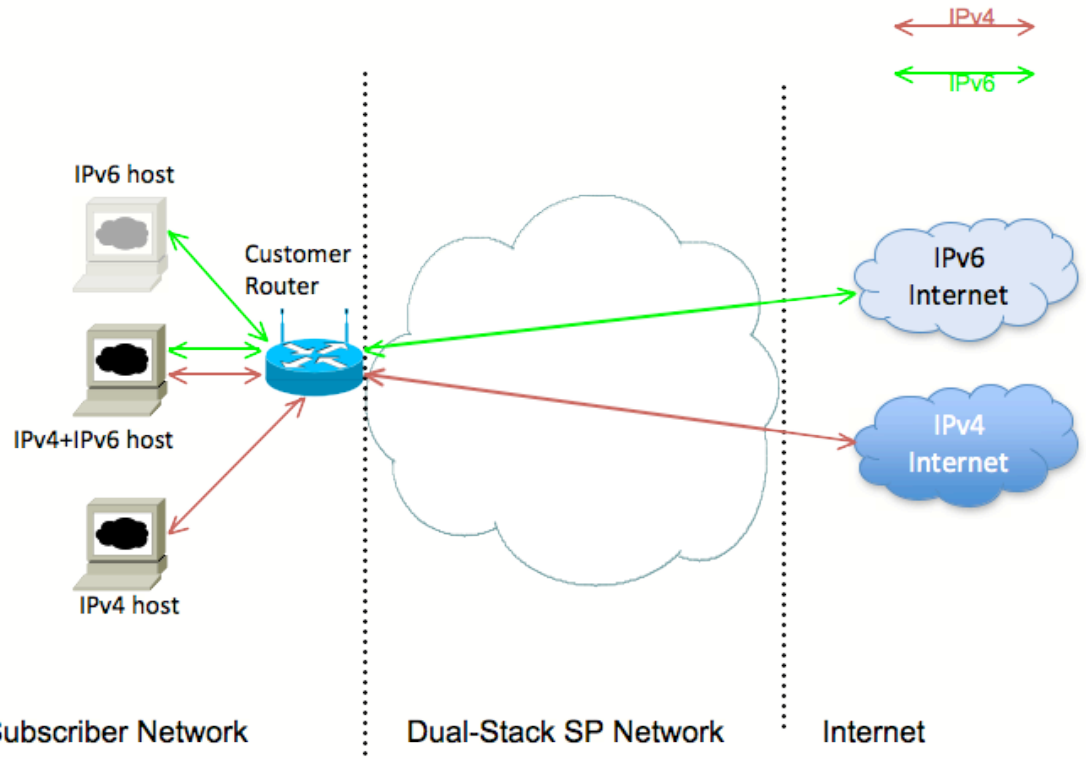
Strategies available

- Do nothing
  - Wait and see what competitors do
  - Business not growing, so don't care
- Extend life of IPv4
  - Push customers to NAT
  - Buy IPv4 address space on the marketplace
- Deploy IPv6
  - Dual stack infrastructure
  - IPv6 and NATed IPv4 for customers
  - Or various other combinations of IPv6, IPv4 and NAT

Prolonging IPv4 to help with IPv6

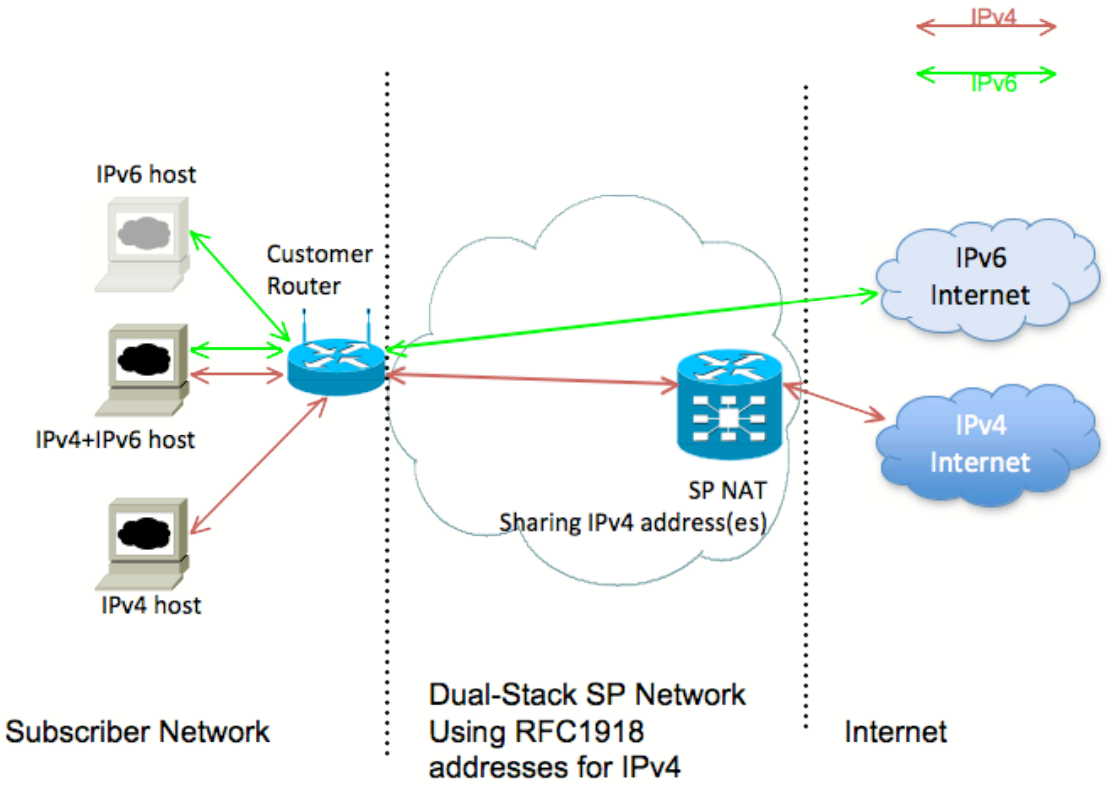
- Large variety of proposals to “make IPv4 last longer” to help with IPv6 deployment
- All involve Large Scale NAT (LSN)
  - NAT444/SP NAT
    - NAT to customer, NAT’ed core.
  - Dual Stack Lite
    - Private IPv4 to IPv6 to Public IPv4
    - Activity of IETF Softwires Working Group
  - NAT64 & IVI
    - Translation between IPv6 and IPv4
    - Activity of IETF Behave Working Group

Dual Stack Network



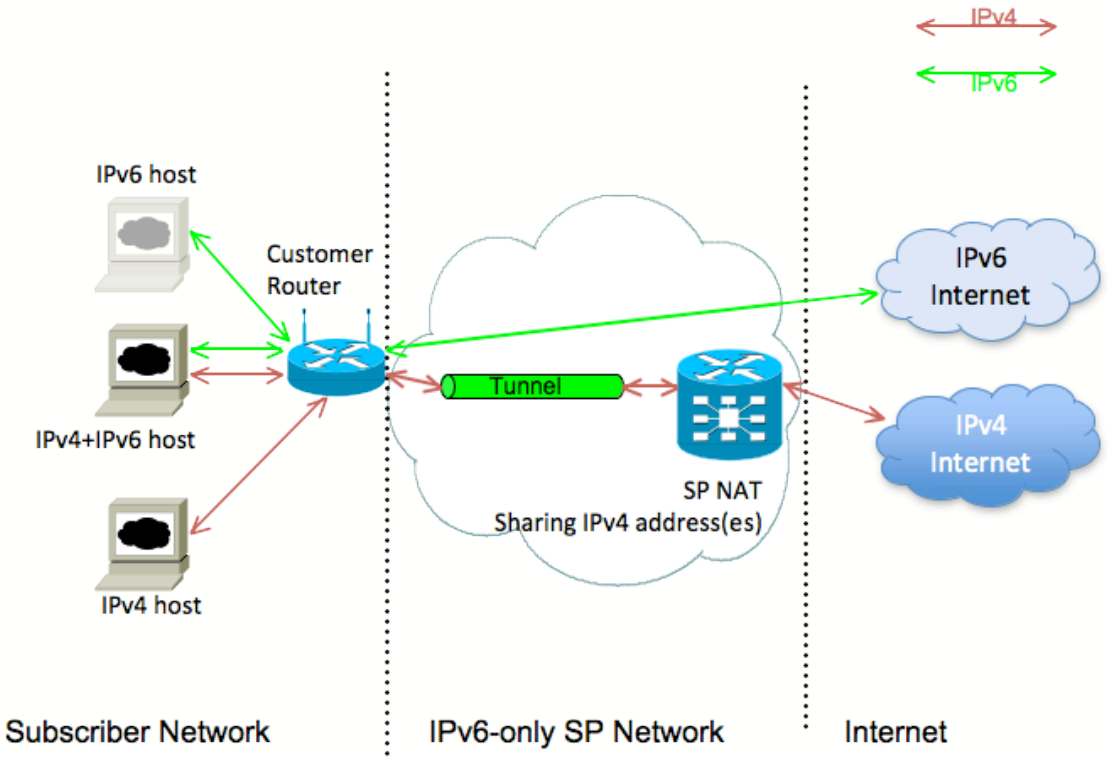
- The original transition scenario, but dependent on:
  - IPv6 being available all the way to the consumer
  - Sufficient IPv4 address space for the consumer

NAT444/SP NAT



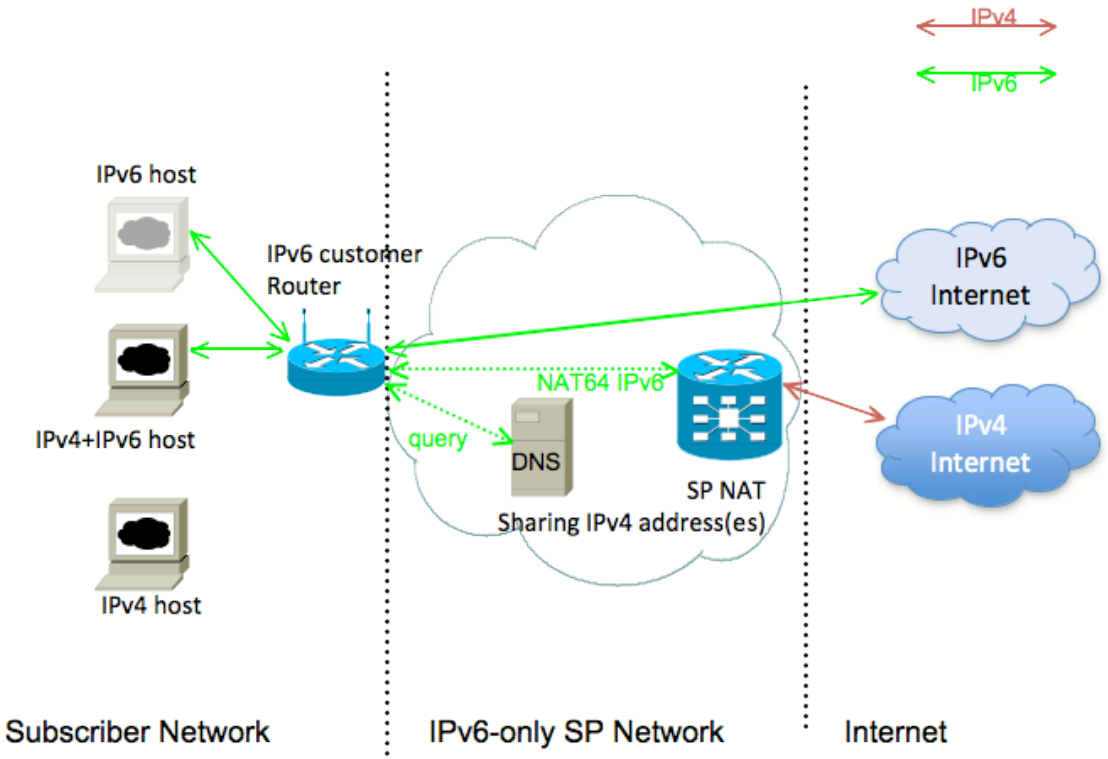
- Consumer uses private IPv4 and native IPv6
- SP uses private IPv4 and native IPv6 for backbone

DualStack-Lite



- SP has IPv6 only infrastructure
- For consumer, IPv4 tunnel to SP NAT, IPv6 native

NAT64



- Consumer uses only IPv6 plus Protocol Translation to reach IPv4
- Service provider uses only IPv6

IPv4 Address Markets

- Address Market:

  When organisations don't return unused address space to their RIR (as they are supposed to do)

  But give it to other organisations (in exchange for some form of compensation)

- If markets happen:

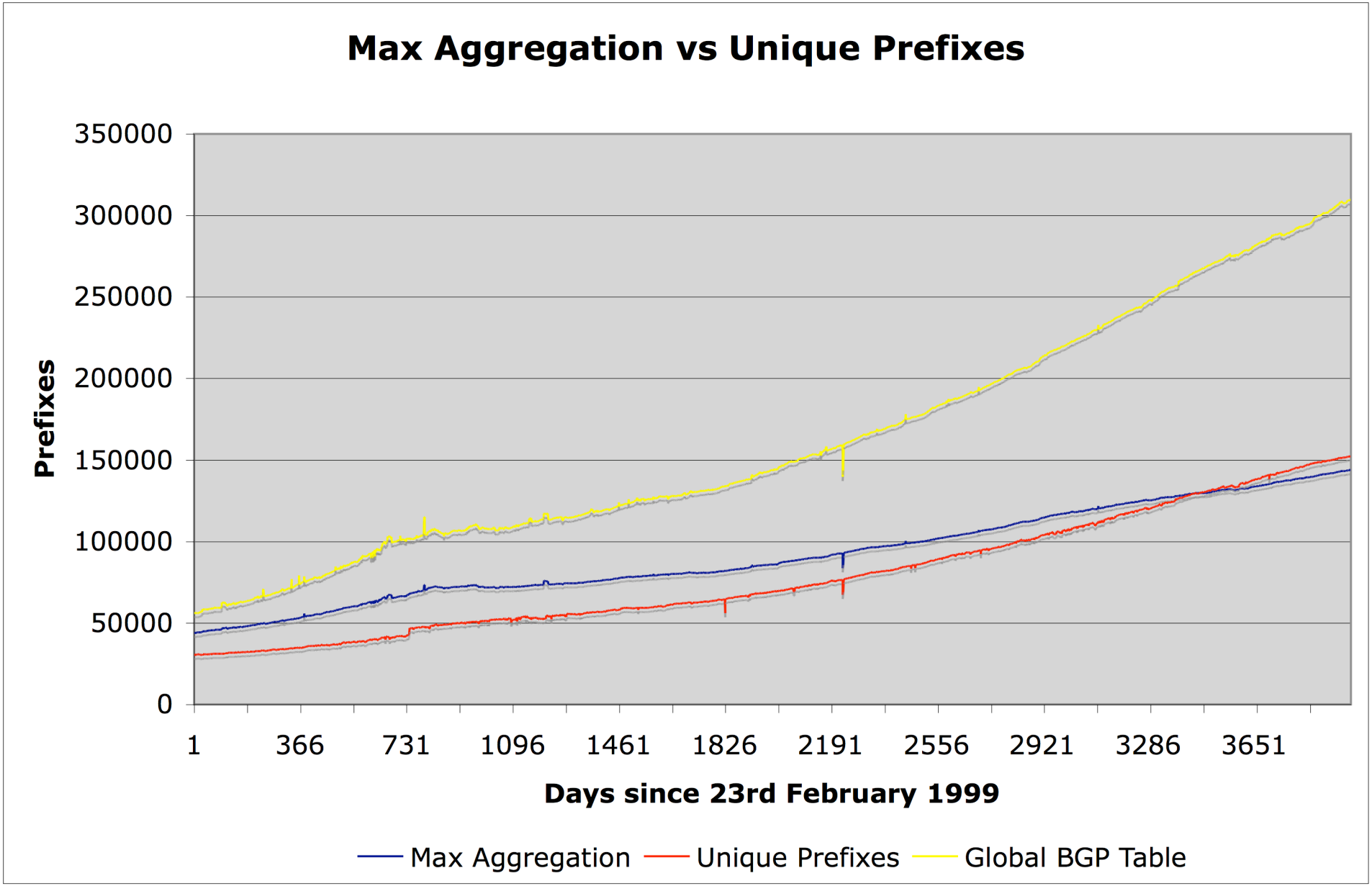
  Organisations will "sell" unused portions of IPv4 address space to other organisations

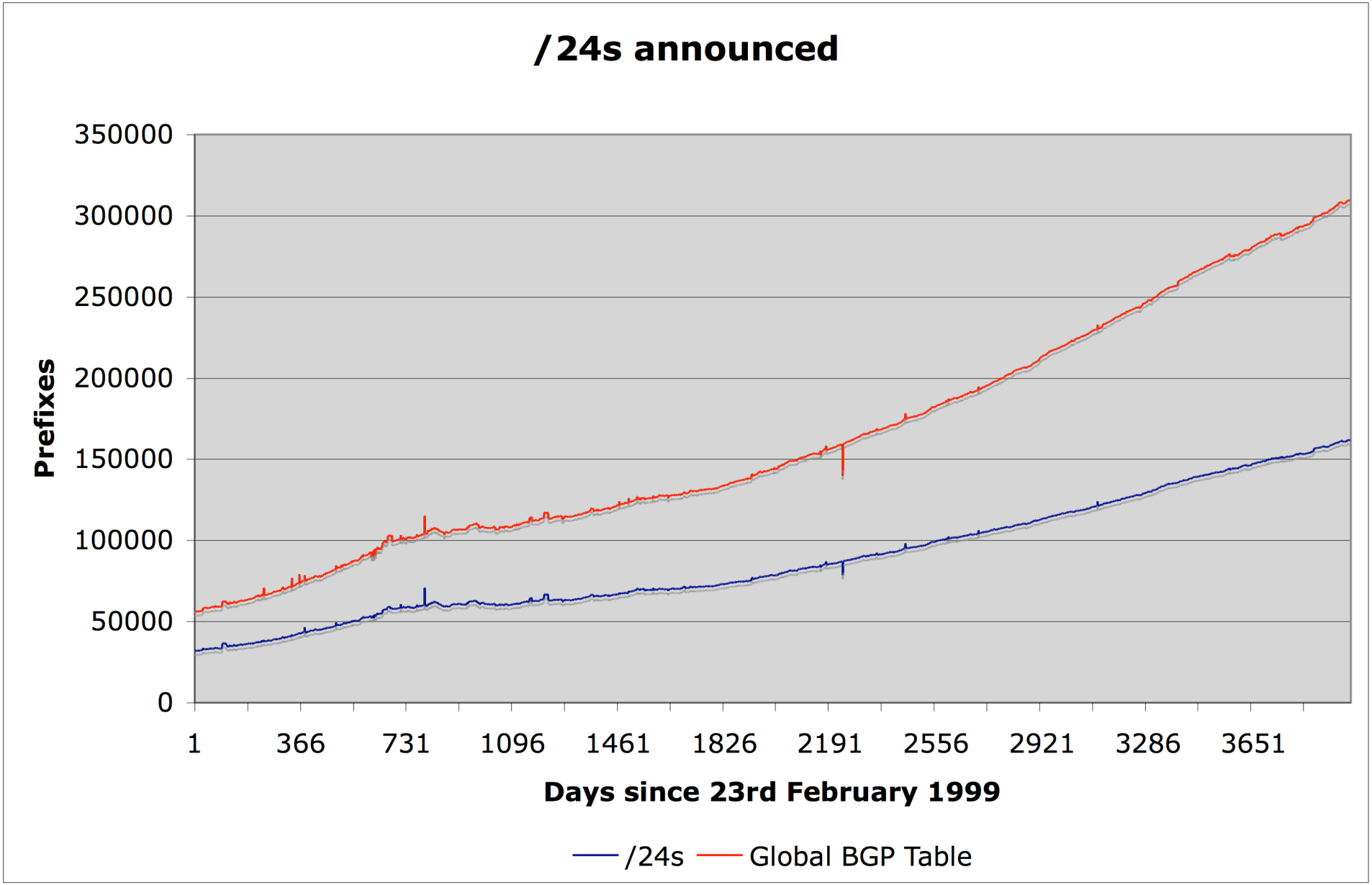
  e.g. have a /16, but two /24s are unused

  Bypasses their RIR (but RIR will still have to register address space so that it can be routed by ISPs)

Routing Table Implications

- Assuming markets happen
  - e.g. organisation with /16 disposes of two /24s
  - Can no longer announce just the /16
  - Have to announce component parts, excluding two /24s
  - One routing announcement replaced by many
- What will happen to the IPv4 Routing Table?
  - Table today is 310k prefixes, of which 162k are /24s
  - Growth is faster than it has been since introduction of CIDR
  - Deaggregation is growing too – Routing Table could theoretically be reduced to 143k prefixes today
  - Source: http://thyme.apnic.net/current/





Deaggregation Effects & Solutions

- If entire Internet deaggregated to /24s
  - 2179238688 host addresses being announced today
  - Equivalent to 8.5 million /24s
- Issues:
  - Router memory (RIB and FIB)
  - Routing System convergence
- Industry aggregation efforts:
  - BGP Features
  - CIDR Report – http://www.cidr-report.org
  - Routing Table Report – http://thyme.apnic.net/current
  - RIPE-399 – http://www.ripe.net/ripe/docs/ripe-399.html

Deaggregation Impacts

- Router memory (RIB & FIB)
  - Shortens router life time & depreciation cycle
  - Increased costs for ISP and customers
- Router processing power
  - Processors are underpowered, depreciation cycle shortened
  - Increased costs for ISP and customers
- Routing System convergence
  - Larger routing table → slower convergence → greater instability
  - Can be improved by faster control plane processors
- Network Performance & Stability
  - Slower convergence → slower recovery from failure → longer downtime
  - Longer downtime → unhappier customers

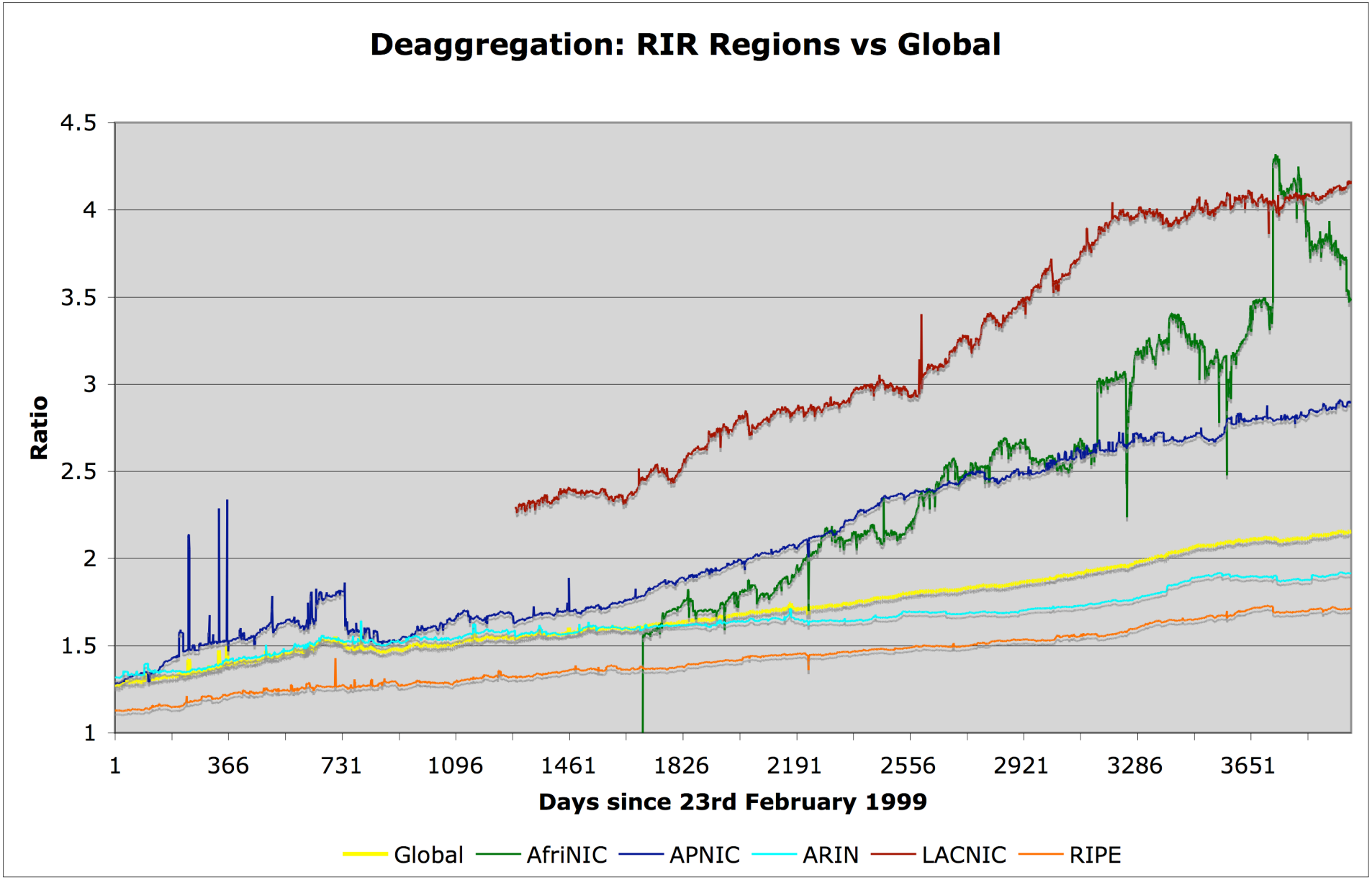
Deaggregation by region: January 2010

Total Prefixes

- Global BGP Table
  310k prefixes
- Europe & Middle East
  71k prefixes
- North America
  129k prefixes
- Asia & Pacific
  75k prefixes
- Africa
  6k prefixes
- Latin America & Caribbean
  27k prefixes

Deaggregation Factor

- Global Average
  2.15
- Europe & Middle East
  1.72
- North America
  1.91
- Asia & Pacific
  2.90
- Africa
  3.48
- Latin America & Caribbean
  4.15



Asia Pacific Aggregation Savings Summary

ASN	No of Nets	Savings	Description
4766	1860	1388	Korea Telecom (KIX)
4755	1311	1175	TATA Communications formerly
17488	1278	1138	Hathway IP Over Cable Interne
18101	1044	1008	Reliance Infocom Ltd Internet
17974	881	830	PT TELEKOMUNIKASI INDONESIA
7545	920	822	TPG Internet Pty Ltd
9829	840	819	BSNL National Internet Backbo
17908	764	709	Tata Communications
24560	839	667	Bharti Airtel Ltd., Telemedia
9299	663	642	Philippine Long Distance Tele
4808	836	623	CNCGROUP IP network: China169
4134	1019	621	CHINANET-BACKBONE
9498	663	617	BHARTI Airtel Ltd.
4780	603	531	Digital United Inc.
17676	563	501	Softbank BB Corp.
9583	986	495	Sify Limited
9808	442	432	Guangdong Mobile Communicatio
9443	510	431	Primus Telecommunications
4804	455	387	Microplex PTY LTD
4802	523	360	iiNet Limited

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

Observations

- Service Providers already need to be more vigilant about routing announcements to Internet
  - Applies to every organisation using BGP
- BGP Instability Report
  - http://bgpupdates.potaroo.net/instability/bgpupd.html
  - Some ISPs have been generating >5 updates per minute!!
- IPv6 transition will create more stress on IPv4
  - Both at consumer level and at infrastructure level
  - Transfer markets might result in many more /24s appearing and many more unstable announcements

50 Most active ASes for the past 7 days

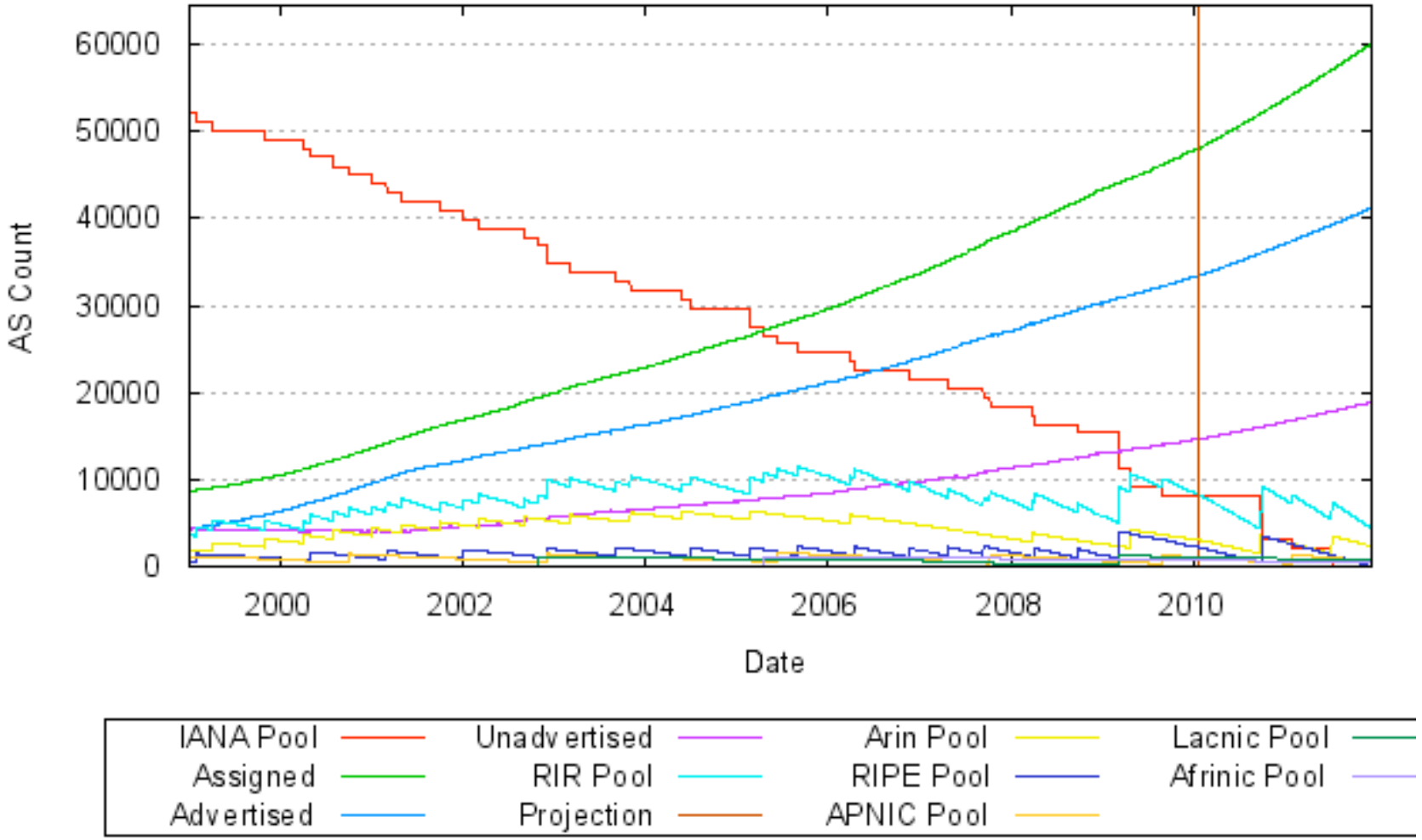
RANK	ASN	UPDs	%	Prefixes	UPDs/Prefix	AS NAME
1	2686	27424	2.75%	230	119.23	AT&T Global Network Services - EMEA
2	7643	20418	2.05%	660	30.94	VNPT-AS-VN Vietnam Posts and Telecommunications (VNPT)
3	5800	16112	1.61%	197	81.79	DNIC-ASBLK-05800-06055 - DoD Network Information Center
4	37986	11414	1.14%	87	131.20	TULIP Tulip Telecom Ltd.
5	45408	10918	1.09%	2	5459.00	
6	9829	9578	0.96%	840	11.40	BSNL-NIB National Internet Backbone
7	4270	8895	0.89%	5	1779.00	Red de Interconexion Universitaria
8	8452	7871	0.79%	1022	7.70	TEDATA TEDATA
9	4134	7312	0.73%	1024	7.14	CHINANET-BACKBONE No.31,Jin-rong Street
10	5668	7172	0.72%	796	9.01	AS-5668 - CenturyTel Internet Holdings, Inc.
11	5803	6494	0.65%	93	69.83	DNIC-ASBLK-05800-06055 - DoD Network Information Center
12	14420	6311	0.63%	372	16.97	CORPORACION NACIONAL DE TELECOMUNICACIONES CNT S.A.
13	151	6272	0.63%	16	392.00	IND-NTC-AS - Hewlett-Packard Company
14	14522	6230	0.62%	347	17.95	Satnet
15	35805	5855	0.59%	557	10.51	UTG-AS United Telecom AS
16	11139	5802	0.58%	466	12.45	CWRIN CW BARBADOS
17	1237	5349	0.54%	143	37.41	KREONET-AS-KR Korea Institute of Science and Technology Information
18	17964	5061	0.51%	192	26.36	DXTNET Beijing Dian-Xin-Tong Network Technologies Co., Ltd.
19	8151	4901	0.49%	1589	3.08	Uninet S.A. de C.V.
20	4249	4822	0.48%	185	26.06	LILLY-AS - Eli Lilly and Company
21	18170	4709	0.47%	22	214.05	CHANGWON-AS-KR Changwon National University
22	17974	4662	0.47%	927	5.03	TELKOMNET-AS2-AP PT Telekomunikasi Indonesia
23	7738	4348	0.44%	433	10.04	Telecomunicacoes da Bahia S.A.
24	8668	4155	0.42%	7	593.57	TELONE-AS TelOne Zimbabwe P/L
25	747	3894	0.39%	108	36.06	TAEGU-AS - Headquarters, USAISC
26	19647	3782	0.38%	33	114.61	HPOD20001 - Hewlett-Packard Operation Division

Key: 10080 updates in 7 days = 1 per minute

The Forgotten Run-Out: ASNs

- AS Numbers as used for BGP are also running out
  - Analysis at http://www.potaroo.net/tools/asns/
  - Current estimates are that the 16-bit ASN pool will be exhausted by December 2011
  - Current allocations up to 56318 have been made to the RIRs
- Work started in 2001 to extend the ASN pool to 32-bits

Running out of 16-bit ASNs



Source: http://www.potaroo.net/tools/asns/fig28.png

32-bit ASNs

- 32-bit ASNs extend the pool:
  - 0-65535 extended to 0-4294967295
- Specification documents
  - Description of 32-bit ASNs
    - http://www.rfc-editor.org/rfc/rfc4893.txt
  - Textual representation
    - http://www.rfc-editor.org/rfc/rfc5396.txt
  - New extended community
    - http://www.ietf.org/internet-drafts/draft-ietf-idr-as4octet-extcomm-generic-subtype-01.txt
- AS 23456 is reserved as interface between 16-bit and 32-bit ASN world

Changes

- 32-bit ASNs are backwardly compatible with 16-bit ASNs
  - No flag day, no need to throw out old routers, or replaced existing 16-bit ASN with a 32-bit ASN
- You need to be aware that:
  - You may connect to organisations with 32-bit ASNs
  - ASN 23456 is not a bogon!
  - You will need a router supporting 32-bit ASNs to use a 32-bit ASN
- Proper BGP implementations will silently transport 32-bit ASNs across the network

If 32-bit ASN not supported:

- Inability to distinguish between peer ASes using 32-bit ASNs
  - They will all be represented by AS23456 - problematic for policy!
- Inability to distinguish prefix's origin AS
  - How to tell whether origin is real or fake?
  - The real and fake both represented by AS23456
  - (There should be a better solution here!)
- Incorrect NetFlow summaries:
  - Prefixes from 32-bit ASNs will all be summarised under AS23456
  - Traffic statistics need to be measured per prefix and aggregated
  - Makes it hard to determine peerability of a neighbouring network

Implementations (Jan 2010)

- Cisco IOS-XR 3.4
- Cisco IOS-XE 2.3
- Cisco IOS 12.0(32)S12 & 12.4(24)T
- Cisco NX-OS 4.0(1)
- Quagga (patches for 0.99.6)
- OpenBGPd (patches for 3.9 & 4.0)
- Juniper JunOSe 4.1.0 & JunOS 9.1
- Redback SEOS
- Force10 FTOS7.7.1 onwards

- http://as4.cluepon.net/index.php/Software_Support

Closing Thoughts

- IPv6 is part of our lives now
  - Not totally clear exactly how pervasive it will become
  - IPv4 is not going away any time soon either
- Pressure on Internet Routing System is growing
  - Deaggregation due to increasing carelessness
  - Potential impact of IPv4 runout plans and address transfer markets
- ASN range is increased to 32-bits
  - No flag day – but how many ISPs are prepared for customers with 32-bit ASNs?