The Next Three Years

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

Philip Smith <pfs@cisco.com>

CTO Consulting Engineer

© 2009 Cisco Systems, Inc. All rights reserved. Cisco Public

Early Internet History

- Late 1980s
  - Exponential growth of the Internet
- Late 1990: CLNS proposed as IP replacement
- 1991-1992
  - Running out of “class-B” network addresses
  - Rapid growth of the “default-free” routing table
  - Imminent exhaustion of 32-bit address space
- Two efforts – short-term versus long-term
  - More at “The Long and Windy ROAD”
    http://rms46.vlsm.org/1/42.html

© 2009 Cisco Systems, Inc. All rights reserved.

Early Internet History

- CIDR and Supernetting proposed in 1992-3
  - Deployment started in 1994
- IETF “ipng” solicitation – RFC1550, Dec 1993
- Direction and technical criteria for ipng choice – RFC1719 and RFC1726, Dec 1994
- Proliferation of proposals:
  - TUBA – RFC1347, June 1992
  - PIP – RFC1621, RFC1622, May 1994
  - CATNIP – RFC1707, October 1994
  - SIP – RFC1710, October 1994
  - NIMROD – RFC1753, December 1994
  - ENCAPS – RFC1955, June 1996

© 2009 Cisco Systems, Inc. All rights reserved.

Early Internet History → 1996

- Other activities included:
  - Development of NAT, PPP, DHCP,…
  - Some IPv4 address reclamation
  - The RIR system was introduced
- → Brakes were put on IPv4 address consumption
- IPv4 32 bit address = 4 billion hosts
  - HD Ratio (RFC3194) realistically limits IPv4 to 250 million hosts

© 2009 Cisco Systems, Inc. All rights reserved.

Recent Internet History
The “boom” years → 2001

- IPv6 Development in full swing
  - Rapid IPv4 consumption
  - IPv6 specifications sorted out
  - (Many) Transition mechanisms developed
- 6bone
  - Experimental IPv6 backbone sitting on top of Internet
  - Participants from over 100 countries
- Early adopters
  - Japan, Germany, France, UK,…

© 2009 Cisco Systems, Inc. All rights reserved.

Recent Internet History
The “bust” years: 2001 → 2003

- The DotCom “crash”
  - i.e. Internet became mainstream
- IPv4:
  - Consumption slowed
  - Address space pressure “reduced”
- IPv6 Indifference
  - Early adopters surging onwards
  - Sceptics more sceptical
  - Yet more transition mechanisms developed

© 2009 Cisco Systems, Inc. All rights reserved.

2004 → Today

- Resurgence in demand for IPv4 address space
  - 13.6% address space still unallocated (07/2009)
  - Exhaustion predictions range from wild to conservative
  - …but mid 2011 seems realistic at current rates
  - …but what about the market for address space?
- Market for IPv4 addresses:
  - Creates barrier to entry
  - Condemns the less affluent to multiple NATs
- IPv6 offers vast address space
  - **The only compelling reason for IPv6**

© 2009 Cisco Systems, Inc. All rights reserved.

Current Situation

- General perception is that “IPv6 has not yet taken hold”
  - IPv4 Address run-out is not “headline news” yet
    - More discussions plus run-out plans being proposed
  - Private sector requires a business case to “migrate”
    - No easy Return on Investment (RoI) computation
- 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?

© 2009 Cisco Systems, Inc. All rights reserved.

Is there a need for a larger address space?

- Internet population
  - ~600 million users in Q4 CY2002
  - ~945M by end CY 2004 – only 10-15%
  - Future Worldwide population? (~9B in 2050)
- US uses 88 /8s – this is 4.8 IPv4 addresses per person
  - Repeat this the world over…
  - 6 billion population could require 29 billion IPv4 addresses
  - (7 times larger than the IPv4 address pool)
- Emerging Internet economies need address space:
  - China uses more than 210 million IPv4 addresses today (12.5 /8s)

© 2009 Cisco Systems, Inc. All rights reserved.

Is there a need for a larger address space?

- RFC 1918 is not sufficient for large environments
  - Cable Operators (e.g. Comcast – NANOG37 presentation)
  - Mobile providers (fixed/mobile convergence)
  - Large enterprises
- The Policy Development process of the RIRs turned down a request to increase private address space
  - RIR membership guideline is to use global addresses instead
  - This leads to an accelerated depletion of the global address space
- Some want 240/4 as new private address space
  - But how to back fit onto all TCP/IP stacks released since 1995?

© 2009 Cisco Systems, Inc. All rights reserved.

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

© 2009 Cisco Systems, Inc. All rights reserved.

Service Provider Status

- Many transit ISPs have “quietly” made their backbones IPv6 capable as part of infrastructure upgrades
  - Native is common (dual stack)
  - Providers using MPLS use 6PE
  - Tunnels still used (unfortunately)
- Examples:
  - NTT has been long time IPv6 capable
  - OpenTransit/FT, TATA International, Telecom Italia, GlobalCrossing, Telefonica, C&W (EU),…
- OCCAID
  - IPv6-only transit ISP effort (linking Asia, N-America, EU)

© 2009 Cisco Systems, Inc. All rights reserved.

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

© 2009 Cisco Systems, Inc. All rights reserved.

Why are we still waiting…?

- That killer application?
  - Internet Gaming or Peer to Peer applications?
  - Windows Vista or 7 (?)
- Our competitors?
  - Any network deployed in last 3 years will be IPv6 capable
  - Even if not enabled!
- The end-user should not have to choose protocols
  - Remember “Turbo” button on early IBM PC clones?
- The “Chattering Classes”
  - People looking for problems, not solutions

© 2009 Cisco Systems, Inc. All rights reserved.

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?

© 2009 Cisco Systems, Inc. All rights reserved.

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

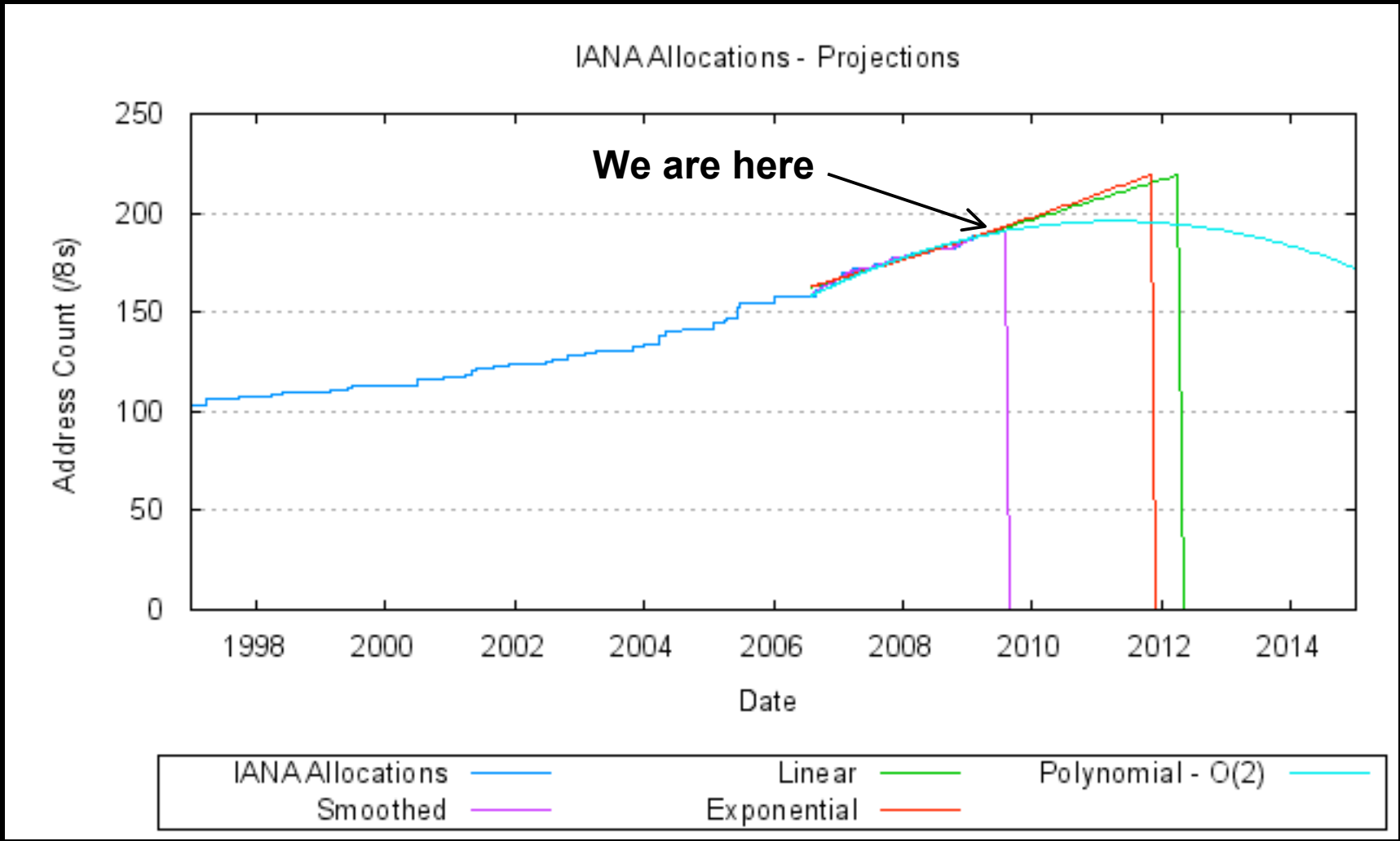
© 2009 Cisco Systems, Inc. All rights reserved.

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
  - 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?

© 2009 Cisco Systems, Inc. All rights reserved. Cisco Public

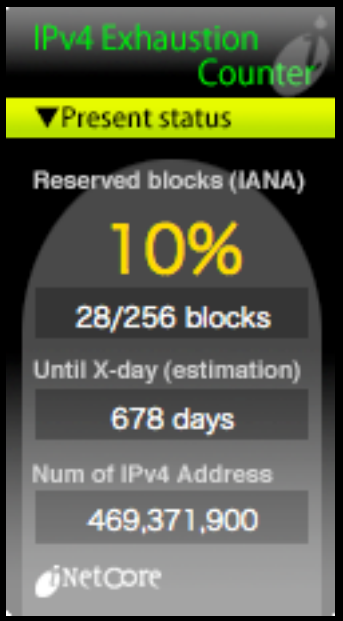
Is IPv4 really running out?



© 2009 Cisco Systems, Inc. All rights reserved. Cisco Public

Is IPv4 really running out?

- Yes
  - IANA IPv4 free pool runs out in June 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/



© 2009 Cisco Systems, Inc. All rights reserved.

IPv4 run-out

- RIR Policy Development process in each RIR region is now handling many 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)

© 2009 Cisco Systems, Inc. All rights reserved.

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

© 2009 Cisco Systems, Inc. All rights reserved.

Options 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

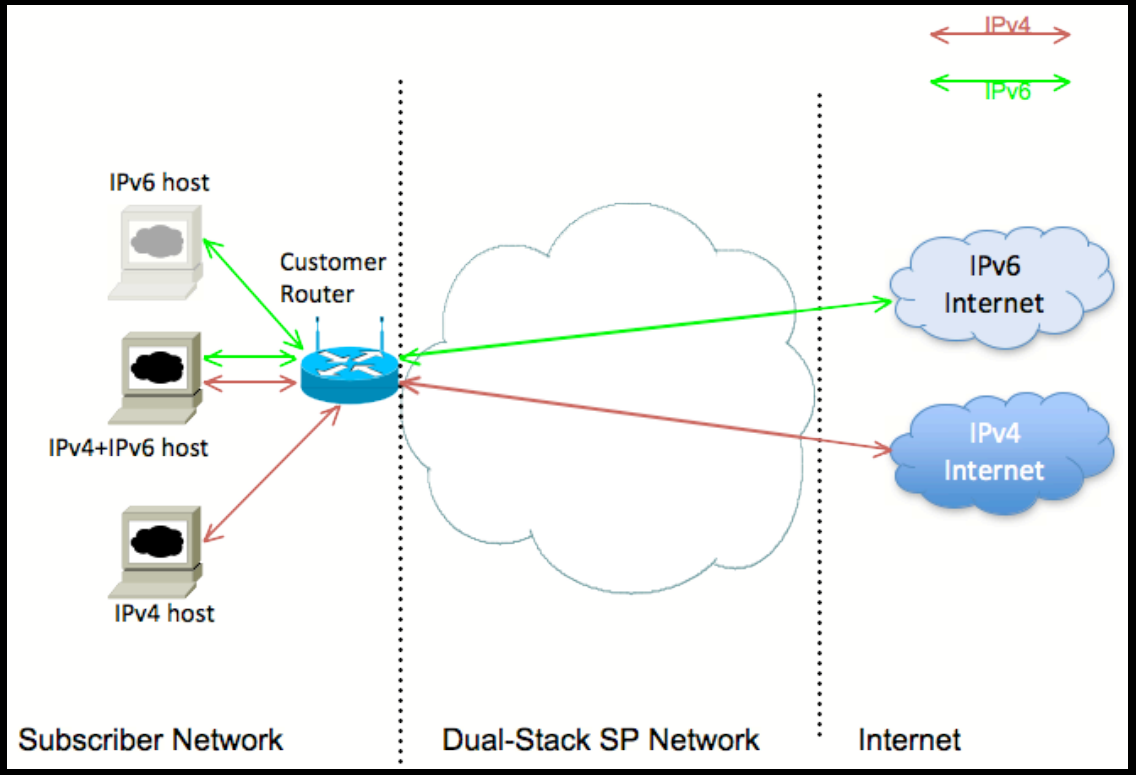
© 2009 Cisco Systems, Inc. All rights reserved.

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

© 2009 Cisco Systems, Inc. All rights reserved.

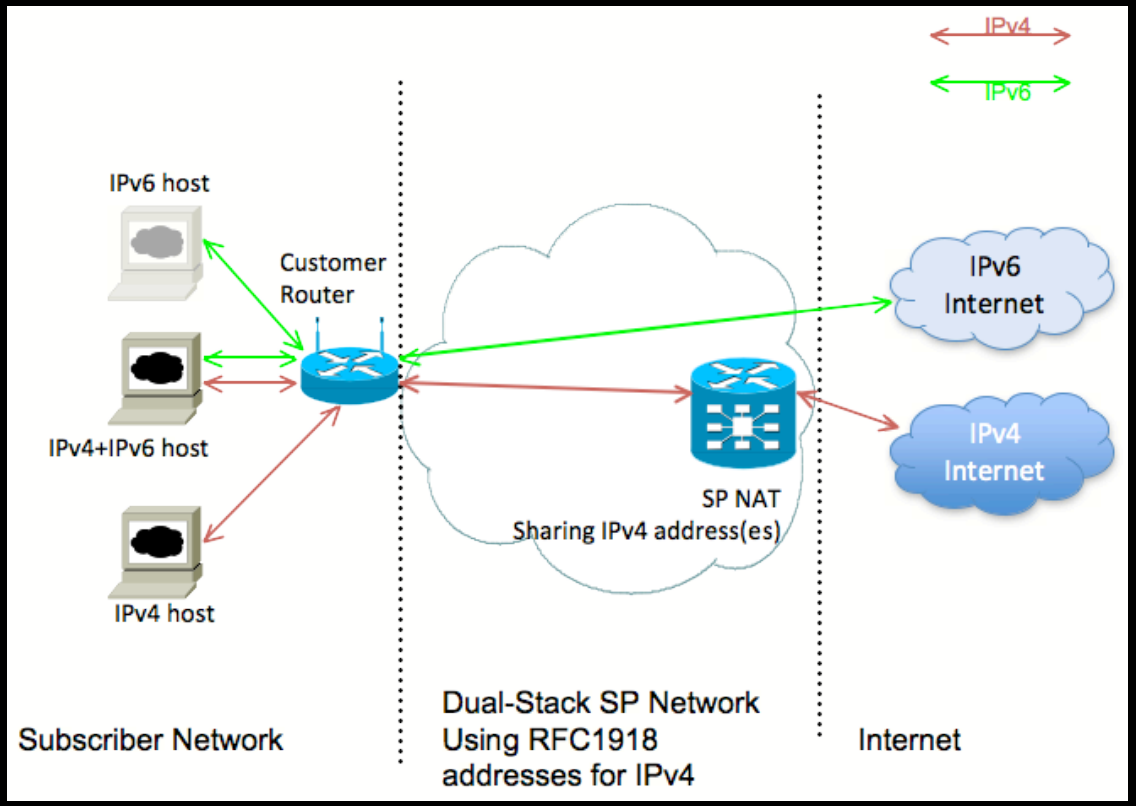
Dual Stack Network



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

© 2009 Cisco Systems, Inc. All rights reserved.

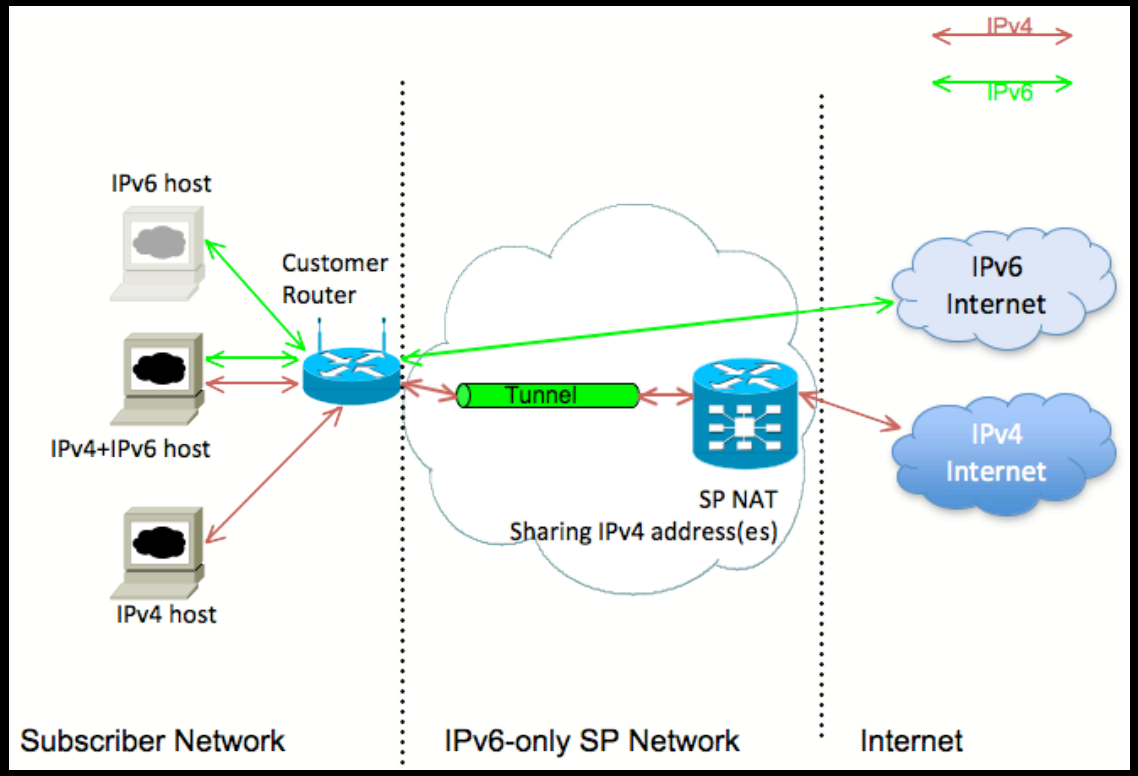
NAT444/SP NAT



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

© 2009 Cisco Systems, Inc. All rights reserved.

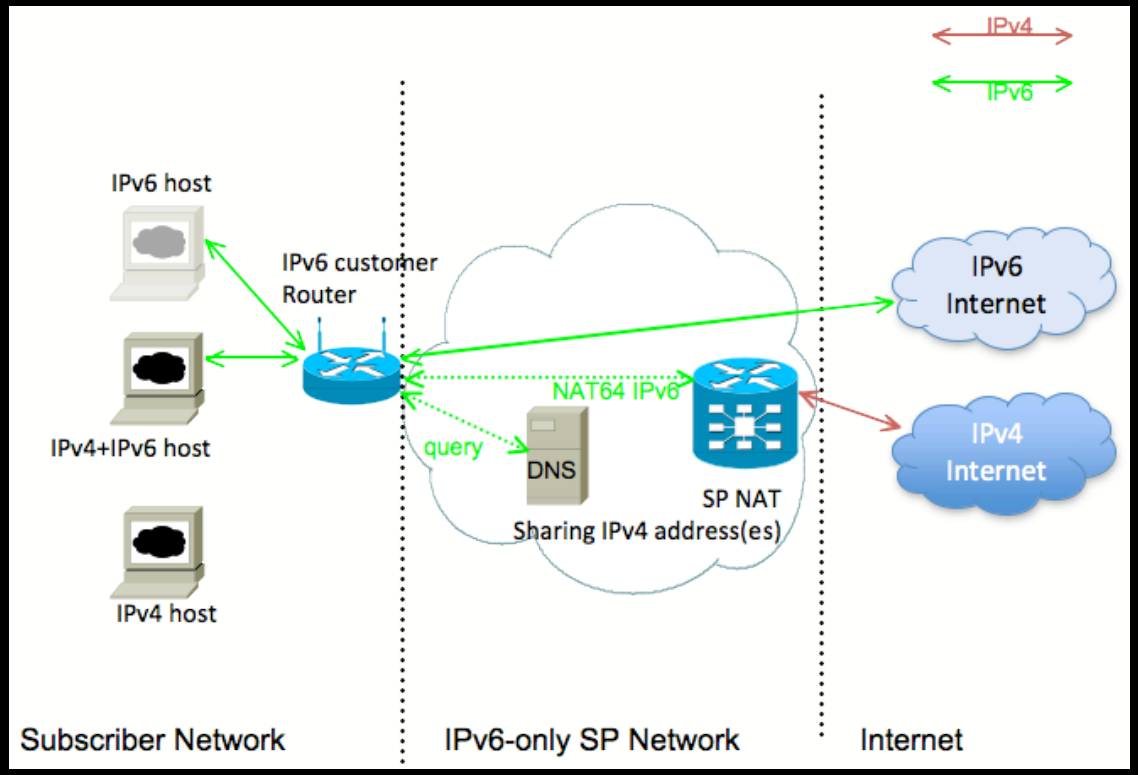
DualStack-Lite



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

© 2009 Cisco Systems, Inc. All rights reserved.

NAT64



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

© 2009 Cisco Systems, Inc. All rights reserved.

IPv4 Address Markets

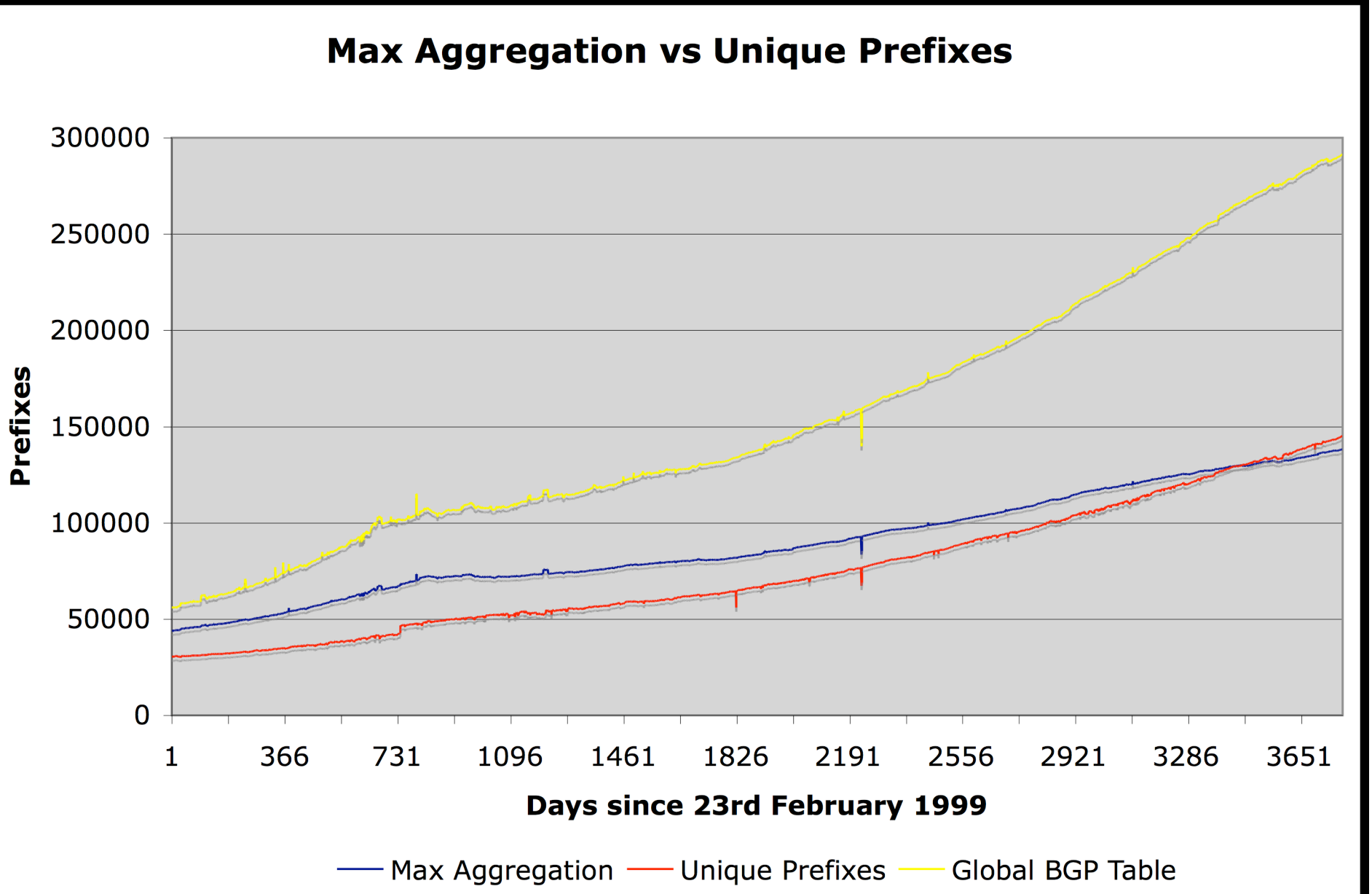
- Address Market:
  - When organisations don't return unused address space to their RIR (as they are supposed to do)
  - But give it to other organisations (in exchange for some form of compensation)
- If markets happen:
  - Organisations will "sell" unused portions of IPv4 address space to other organisations
    - 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)

© 2009 Cisco Systems, Inc. All rights reserved.

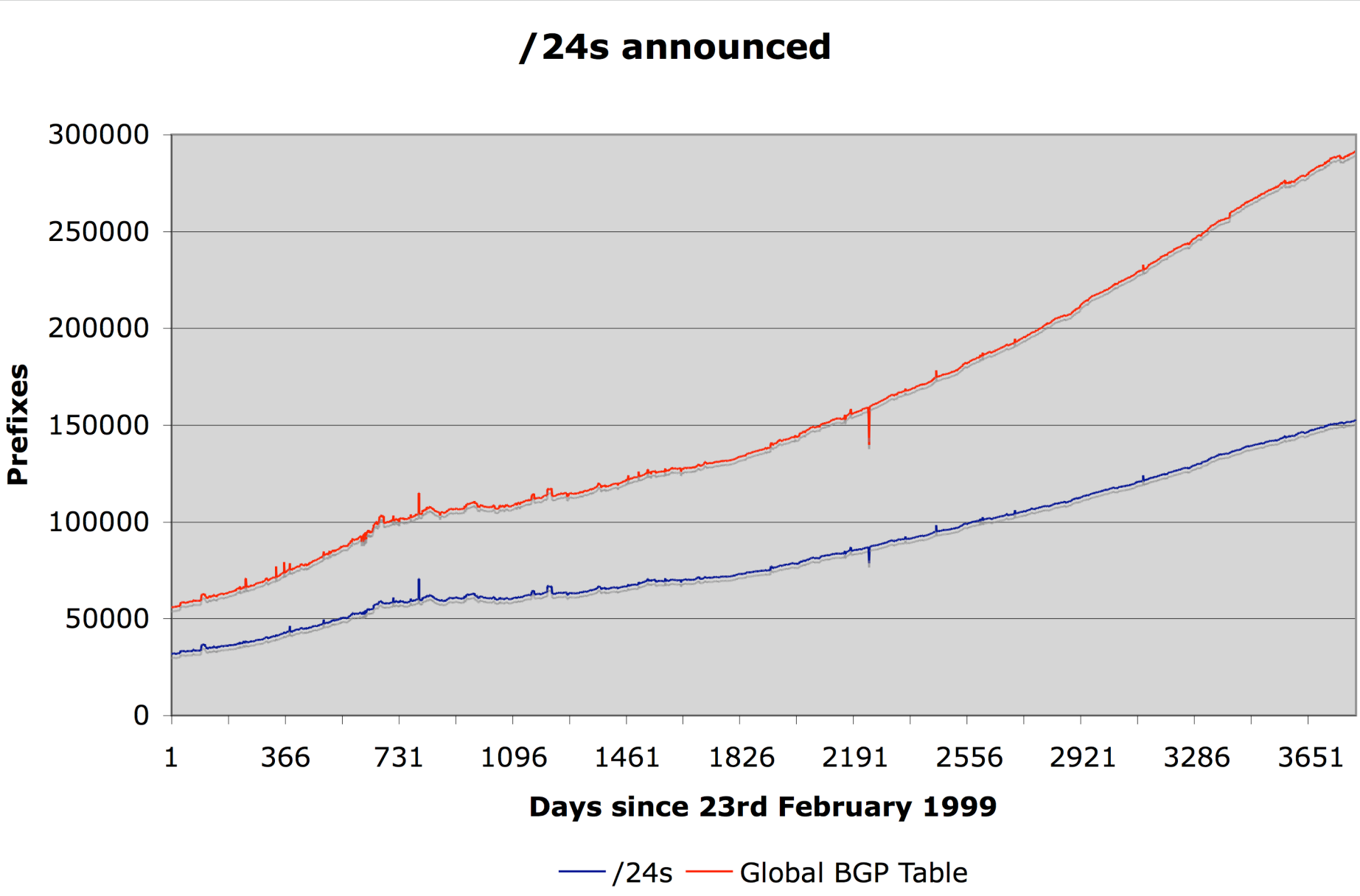
Routing Table Implications

- 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 293k prefixes, of which 154k are /24s
  - Growth is faster than it has been since introduction of CIDR
  - Deaggregation is growing too – Routing Table could theoretically be reduced to 139k prefixes today
  - Source: http://thyme.apnic.net/current/

© 2009 Cisco Systems, Inc. All rights reserved. Cisco Public



© 2009 Cisco Systems, Inc. All rights reserved.



© 2009 Cisco Systems, Inc. All rights reserved.

Deaggregation Effects & Solutions

- If entire Internet deaggregated to /24s
  - 2102217280 host addresses being announced today
  - Equivalent to 8.2 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

© 2009 Cisco Systems, Inc. All rights reserved.

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

© 2009 Cisco Systems, Inc. All rights reserved.

Deaggregation by Region: July 2009

Total Prefixes

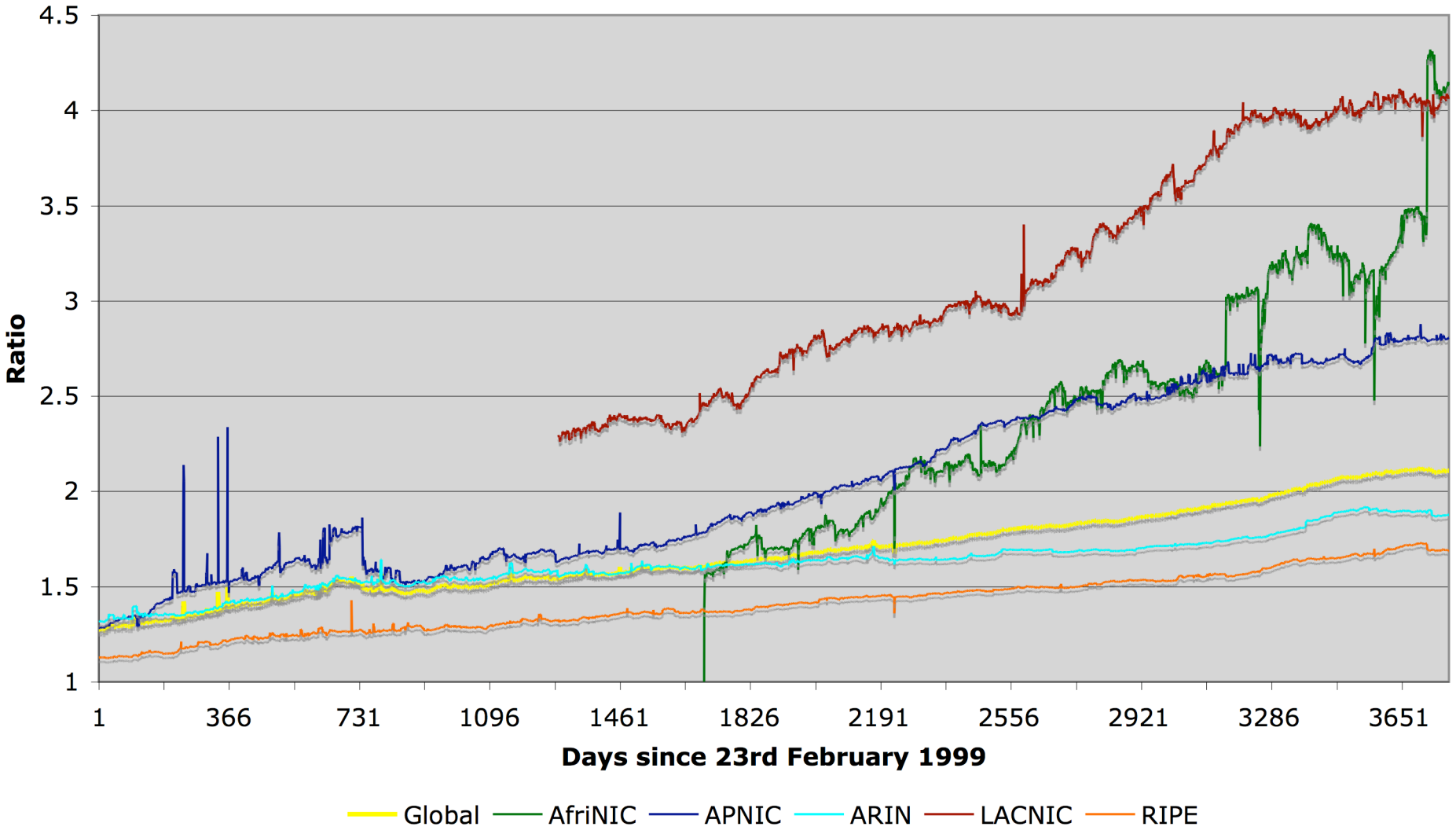
- Global BGP Table
  291k prefixes
- Europe & Middle East
  67k prefixes
- North America
  124k prefixes
- Asia & Pacific
  70k prefixes
- Africa
  6k prefixes
- Latin America & Caribbean
  25k prefixes

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

© 2009 Cisco Systems, Inc. All rights reserved. Cisco Public

Deaggregation: RIR Regions vs Global



© 2009 Cisco Systems, Inc. All rights reserved.

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

© 2009 Cisco Systems, Inc. All rights reserved. Cisco Public

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 are 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

© 2009 Cisco Systems, Inc. All rights reserved.

50 Most active ASes for the past 7 days

RANK	ASN	UPDs	%	Prefixes	UPDs/Prefix	AS NAME
1	9198	97867	6.44%	374	261.68	KAZTELECOM-AS Kazakhtelecom Corporate Sales Administration
2	8151	36334	2.39%	4865	7.47	Uninet S.A. de C.V.
3	11	30681	2.02%	15	2045.40	HARVARD - Harvard University
4	33783	22935	1.51%	157	146.08	EEPAD
5	47408	15350	1.01%	22	697.73	MANDARIN-AS Mandarin WIMAX Sicilia SpA
6	8452	14603	0.96%	1027	14.22	TEDATA TEDATA
7	6389	9496	0.62%	4241	2.24	BELLSOUTH-NET-BLK - BellSouth.net Inc.
8	7011	9433	0.62%	988	9.55	FRONTIER-AND-CITIZENS - Frontier Communications of America, Inc.
9	5668	8455	0.56%	1056	8.01	AS-5668 - CenturyTel Internet Holdings, Inc.
10	7018	7826	0.52%	1541	5.08	ATT-INTERNET4 - AT&T WorldNet Services
11	24863	7755	0.51%	910	8.52	LINKdotNET-AS
12	4249	7419	0.49%	182	40.76	LILLY-AS - Eli Lilly and Company
13	9829	7368	0.48%	807	9.13	BSNL-NIB National Internet Backbone
14	12479	6988	0.46%	475	14.71	UNI2-AS Uni2 Autonomous System
15	4323	6724	0.44%	4333	1.55	TWTC - tw telecom holdings, inc.
16	20115	6451	0.42%	1429	4.51	CHARTER-NET-HKY-NC - Charter Communications
17	12066	6423	0.42%	146	43.99	TRICOM
18	7738	6347	0.42%	410	15.48	Telecomunicacoes da Bahia S.A.
19	10620	6023	0.40%	999	6.03	TV Cable S.A.
20	35805	6010	0.40%	423	14.21	UTG-AS United Telecom AS
21	5050	5889	0.39%	16	368.06	PSC-EXT - Pittsburgh Supercomputing Center
22	17974	5778	0.38%	710	8.14	TELKOMNET-AS2-AP PT Telekomunikasi Indonesia
23	3356	5687	0.37%	1210	4.70	LEVEL3 Level 3 Communications
24	1785	5685	0.37%	1728	3.29	AS-PAETEC-NET - PaeTec Communications, Inc.
25	2386	5570	0.37%	1287	4.33	INS-AS - AT&T Data Communications Services
26	22773	5475	0.36%	1082	5.06	ASN-CXA-ALL-CCI-22773-RDC - Cox Communications Inc.

Key: 10080 updates in 7 days = 1 per minute

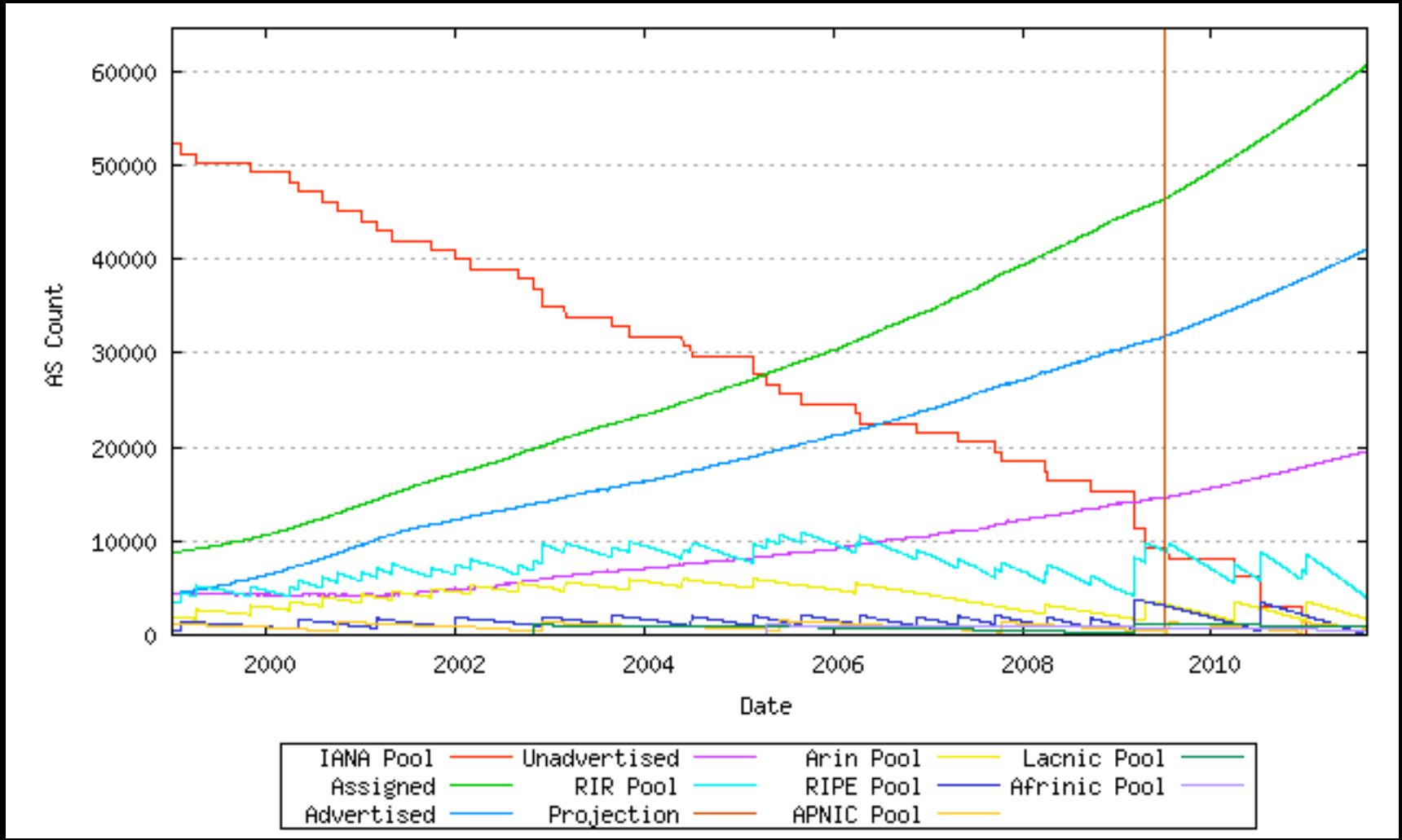
© 2009 Cisco Systems, Inc. All rights reserved. Cisco Public

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 August 2011
  - Current allocations up to 55294 have been made to the RIRs
- Work started in 2001 to extend the ASN pool to 32-bits

© 2009 Cisco Systems, Inc. All rights reserved.

Running out of 16-bit ASNs



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

© 2009 Cisco Systems, Inc. All rights reserved. Cisco Public

32-bit ASNs

- Standards 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-00.txt
- AS 23456 is reserved as interface between 16-bit and 32-bit ASN world

© 2009 Cisco Systems, Inc. All rights reserved.

Representation

- 32-bit ASNs extend the pool:
  - 0-65535 extended to 0-4294967295
- Representation of 65536-4294967295 range
  - Most operators favour traditional format (asplain)
  - A few prefer dot notation (X.Y):
    - asdot for 65536-4294967295, e.g 2.4
    - asdot+ for 0-4294967295, e.g 0.64513
  - But regular expressions will have to be completely rewritten for asdot and asdot+ !!!

© 2009 Cisco Systems, Inc. All rights reserved.

Changes (1)

- 32-bit ASNs are backwardly compatible with 16-bit ASNs
- There is no flag day
- You do NOT need to:
  - Throw out your old routers
  - Replace your 16-bit ASN with a 32-bit ASN

© 2009 Cisco Systems, Inc. All rights reserved.

Changes (2)

- You do 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
- If you have a proper BGP implementation, 32-bit ASNs will be transported silently across your network

© 2009 Cisco Systems, Inc. All rights reserved.

How does it work (1)?

- Local router only supports 16-bit ASN
- Remote router uses 32-bit ASN
- BGP peering initiated:
  - Remote asks local if 32-bit supported (BGP capability negotiation)
  - When local says “no”, remote then presents AS23456
  - Local needs to be configured to peer with remote using AS23456

© 2009 Cisco Systems, Inc. All rights reserved.

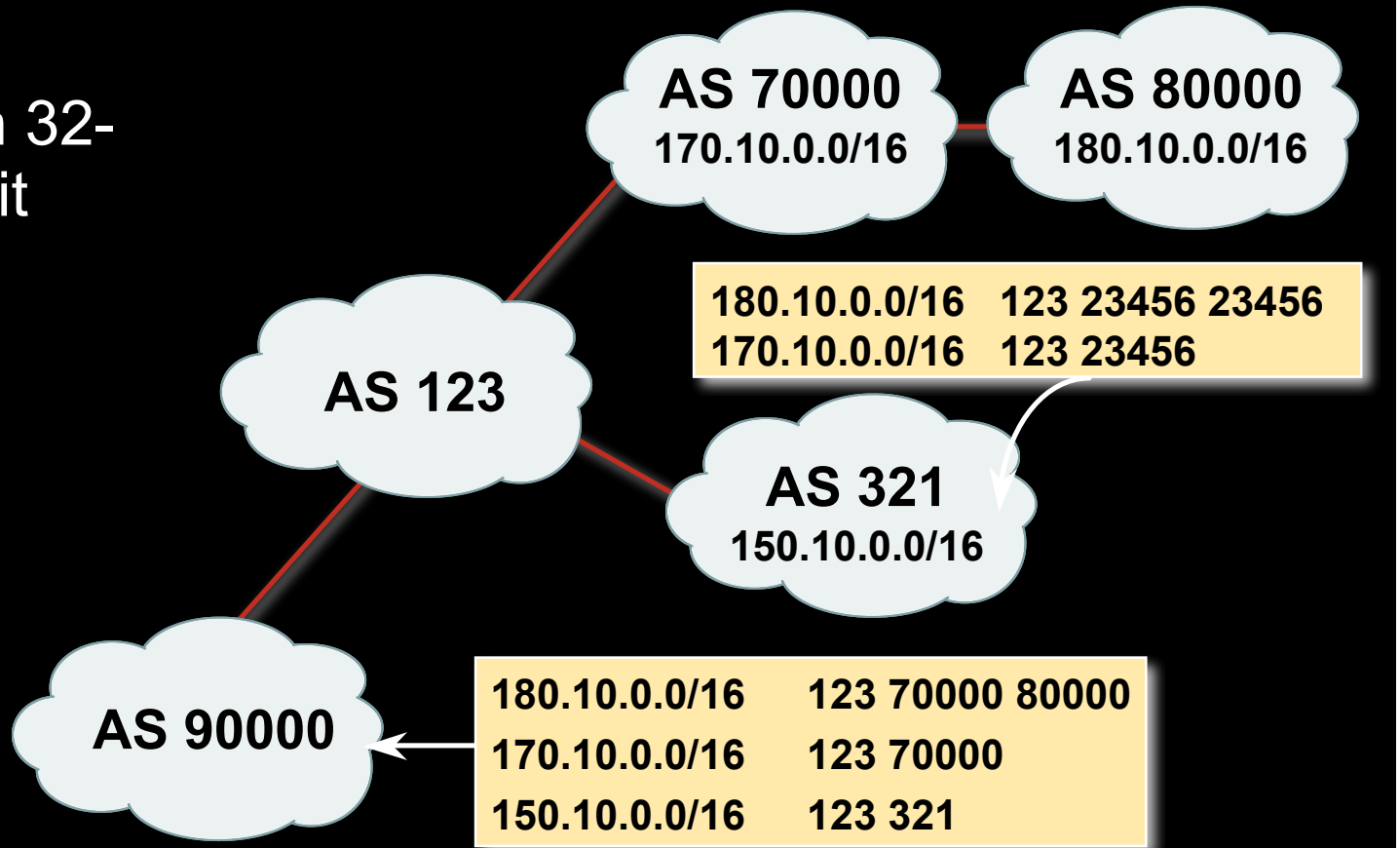
How does it work (2)?

- BGP peering initiated (cont):
  - BGP session established using AS23456
  - 32-bit ASN included in a new BGP attribute called AS4_PATH
    - (as opposed to AS_PATH for 16-bit ASNs)
- Result:
  - 16-bit ASN world sees 16-bit ASNs and 23456 standing in for 32 bit ASNs
  - 32-bit ASN world sees 16 and 32-bit ASNs

© 2009 Cisco Systems, Inc. All rights reserved.

Example:

- Internet with 32-bit and 16-bit ASNs
- AS-PATH length maintained



© 2009 Cisco Systems, Inc. All rights reserved.

32-bit ASN not supported:

- Inability to distinguish between peer ASes using 32-bit ASNs
  - They will all be represented by AS23456
  - Could be problematic for transit provider's 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!)

© 2009 Cisco Systems, Inc. All rights reserved.

32-bit ASN not supported:

- 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

© 2009 Cisco Systems, Inc. All rights reserved.

Implementations (Apr 09)

- Cisco IOS-XR 3.4 onwards
- Cisco IOS-XE 2.3 onwards
- 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

© 2009 Cisco Systems, Inc. All rights reserved.

Closing Thoughts

- IPv6 is part of our lives now
  - Not totally clear exactly how pervasive it will become
  - But IPv4 is not going away any time soon
- Pressure on Internet Routing System is growing
  - Deaggregation due to increasing carelessness
  - Plus potential impact of 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?

© 2009 Cisco Systems, Inc. All rights reserved.