IPv6 and 4-byte ASN Update

Philip Smith <pfs@cisco.com>

PacNOG 8

Pohnpei, FSM

22nd-27th November 2010

IPv6 Update

2004 → Today

- Resurgence in demand for IPv4 address space
  - 5% address space still unallocated (10/2010)
  - Exhaustion predictions have ranged from wild to conservative
  - …but early 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 NATs
- IPv6 offers vast address space
  - **The only compelling reason for IPv6**

Current Situation

- General perception is that “IPv6 has not yet taken hold”
  - IPv4 Address run-out has now made it into “headline news”
    - More discussions and run-out plans proposed
  - Private sector still demanding 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?

Status in Internet Operational Community

- Service Providers get an IPv6 prefix from their regional Internet Registries
  - Very straight forward process when compared with IPv4
  - APNIC members simply “tick a box”
- 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

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/Verio has been long time IPv6 capable
  - HE, OpenTransit/FT, TATA International, Telecom Italia, GlobalCrossing, Telefonica, C&W (EU),…
  - OCCAID
    - IPv6-only transit ISP effort (linking Asia, N-America, EU)

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

Why are we still waiting…?

- That killer application?
  - Internet Gaming or Peer to Peer applications?
  - Windows 7 (?), Apple iPad (?)
- 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?
- Content
  - Do the content providers know about 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

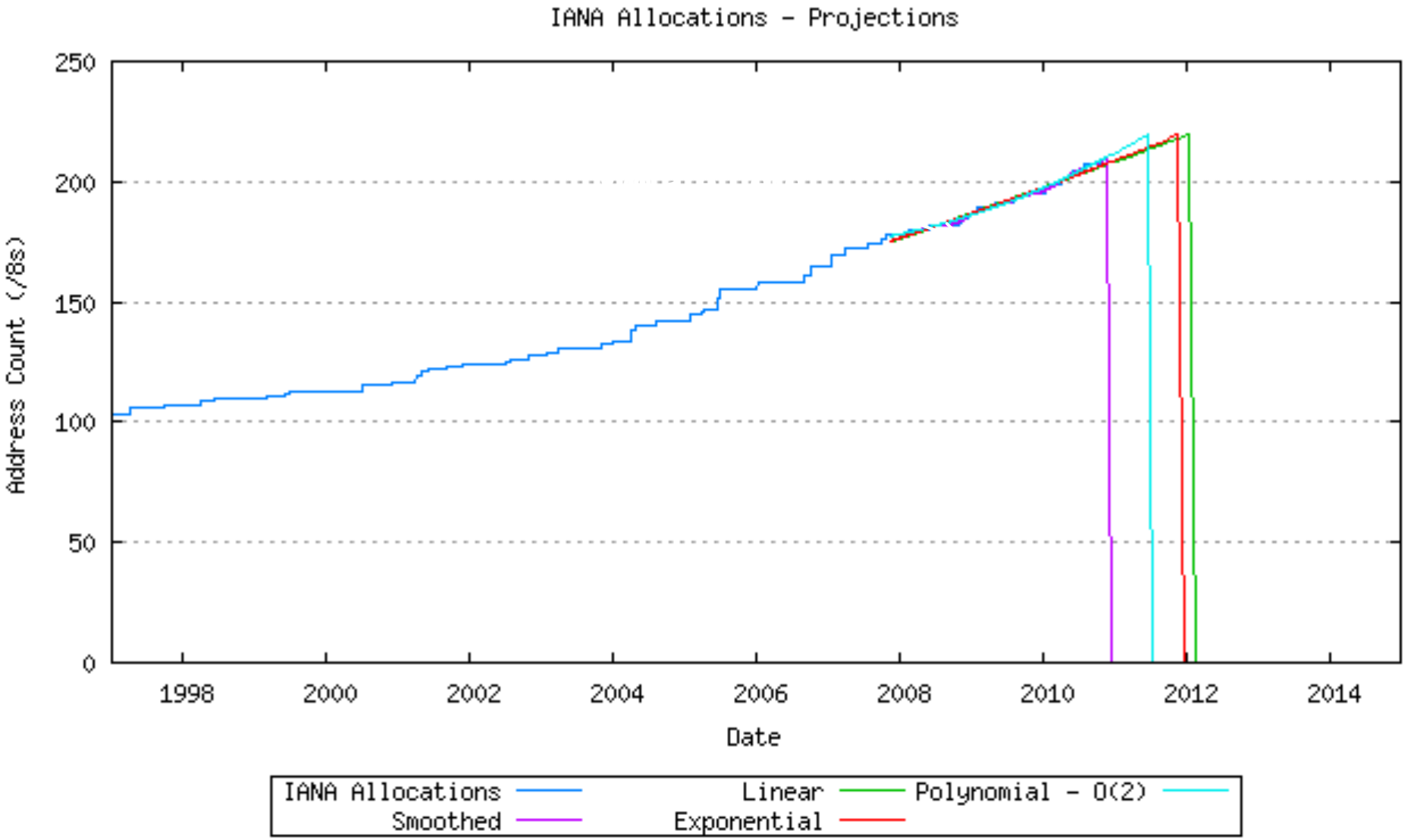
  Realistically IPv6 and IPv4 will co-exist for many years

  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?



Is IPv4 really running out?

- Yes

  IANA IPv4 free pool runs out in June 2011

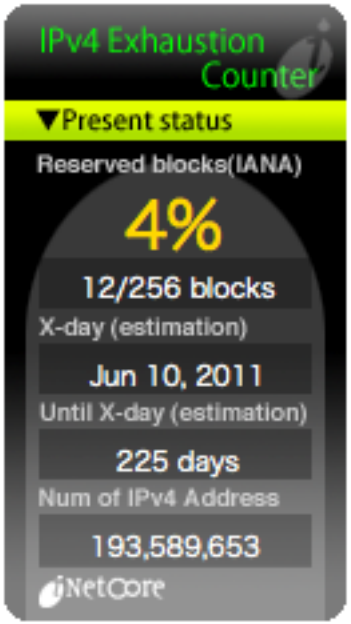
  RIR IPv4 free pool runs out within 2-3 months after

  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 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)

4-byte ASN Update

Autonomous System Number (ASN)

- Two ranges

0-65535	(original 16-bit range)
65536-4294967295	(32-bit range - RFC4893)

- Usage:

0 and 65535	(reserved)
1-64495	(public Internet)
64496-64511	(documentation - RFC5398)
64512-65534	(private use only)
23456	(represent 32-bit range in 16-bit world)
65536-65551	(documentation - RFC5398)
65552-4294967295	(public Internet)

- 32-bit range representation specified in RFC5396
  - Defines “asplain” (traditional format) as standard notation

Autonomous System Number (ASN)

- ASNs are distributed by the Regional Internet Registries
  - They are also available from upstream ISPs who are members of one of the RIRs
- Current 16-bit ASN allocations up to 56319 have been made to the RIRs
  - Around 35000 are visible on the Internet
- The RIRs also have received 1024 32-bit ASNs each
  - Out of 825 assignments, around 500 are visible on the Internet
- See **www.iana.org/assignments/as-numbers**

32-bit ASNs

- Standards documents
  - Description of 32-bit ASNs www.rfc-editor.org/rfc/rfc4893.txt
  - Textual representation www.rfc-editor.org/rfc/rfc5396.txt
  - New extended community www.rfc-editor.org/rfc/rfc5668.txt
- 16-bit ASNs
  - Refers to the range 0 to 65535
- 32-bit ASNs
  - Refers to the range 65536 to 4294967295
  - (or the extended range)
- 32-bit ASN pool
  - Refers to the range 0 to 4294967295

Getting a 32-bit ASN

- Sample RIR policy

  www.apnic.net/docs/policy/asn-policy.html

- From 1st January 2007

  32-bit ASNs were available on request

- From 1st January 2009

  32-bit ASNs were assigned by default

  16-bit ASNs were only available on request

- From 1st January 2010

  No distinction – ASNs assigned from the 32-bit pool

Representation

- Representation of 0-4294967295 ASN 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+ !!!**
- For example:
  - ^[0-9]+$ matches any ASN (16-bit and asplain)
  - This and equivalents extensively used in BGP multihoming configurations for traffic engineering
- Equivalent regexp for asdot is: ^([0-9]+)|([0-9]+\.[0-9]+)$
- Equivalent regexp for asdot+ is: ^[0-9]+\.[0-9]+$

Changes

- 32-bit ASNs are backward 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
- You do need to be aware that:
  - Your customers will come 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 locally
- If you have a proper BGP implementation, 32-bit ASNs will be transported silently across your network

How does it work?

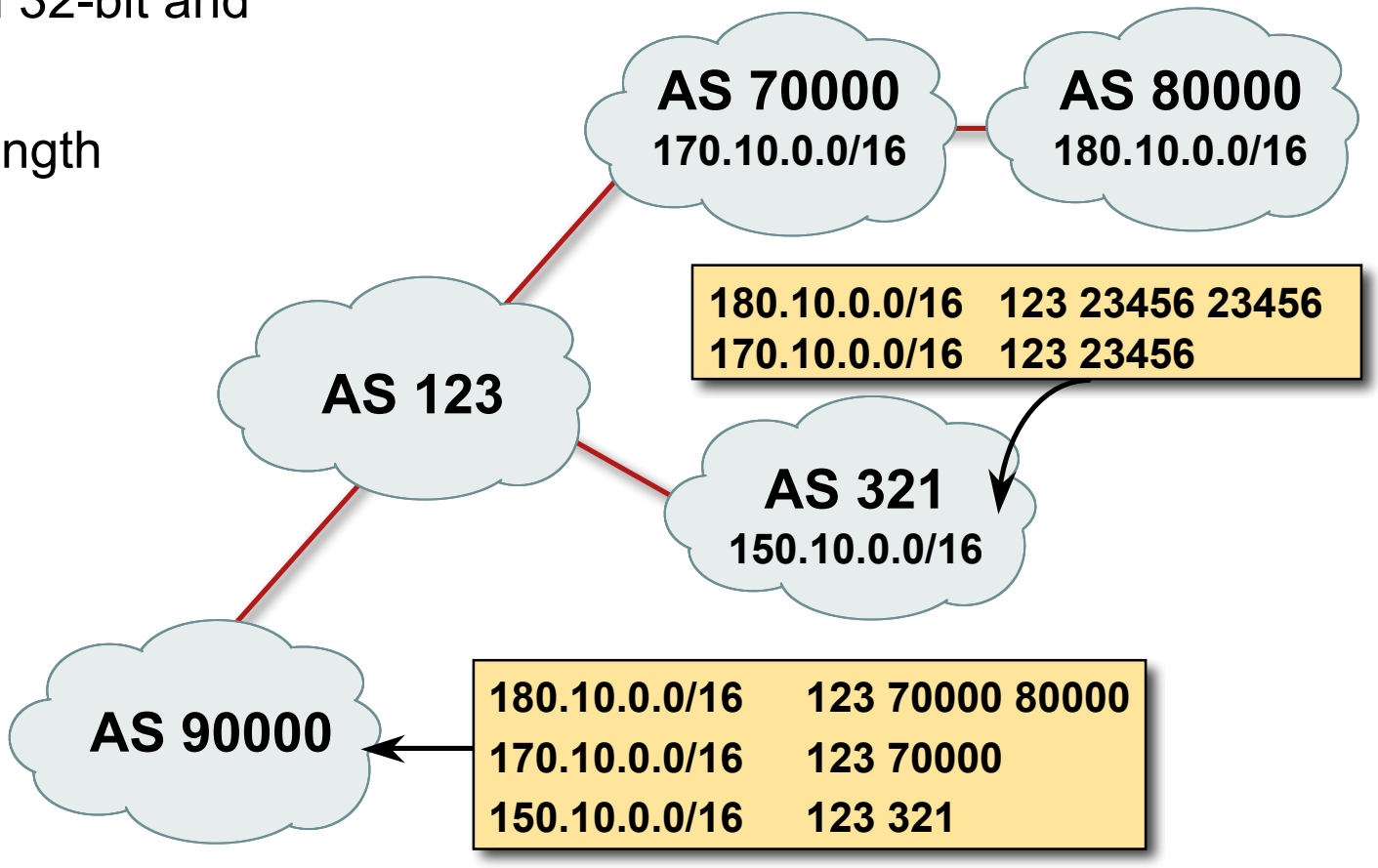
- If local router and remote router supports configuration of 32-bit ASNs
  - BGP peering is configured as normal using the 32-bit ASN
- If local router and remote router does not support configuration of 32-bit ASNs
  - BGP peering can only use a 16-bit ASN
- If local router only supports 16-bit ASN and remote router/network has a 32-bit ASN
  - Compatibility mode is initiated…

Compatibility Mode:

- Local router only supports 16-bit ASN and 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
- 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

Example:

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



What do they look like?

- IPv4 prefix originated by AS196613

asplain format

```
as4-7200#sh ip bgp 145.125.0.0/20
BGP routing table entry for 145.125.0.0/20, version 58734
Paths: (1 available, best #1, table default)
  131072 12654 196613
    204.69.200.25 from 204.69.200.25 (204.69.200.25)
      Origin IGP, localpref 100, valid, internal, best
```

- IPv4 prefix originated by AS3.5

asdot format

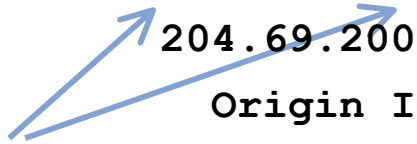
```
as4-7200#sh ip bgp 145.125.0.0/20
BGP routing table entry for 145.125.0.0/20, version 58734
Paths: (1 available, best #1, table default)
  2.0 12654 3.5
    204.69.200.25 from 204.69.200.25 (204.69.200.25)
      Origin IGP, localpref 100, valid, internal, best
```

What do they look like?

- IPv4 prefix originated by AS196613

  But 16-bit AS world view:

```
BGP-view1>sh ip bgp 145.125.0.0/20
BGP routing table entry for 145.125.0.0/20, version 113382
Paths: (1 available, best #1, table Default-IP-Routing-
Table)
23456 12654 23456
    204.69.200.25 from 204.69.200.25 (204.69.200.25)
      Origin IGP, localpref 100, valid, external, best
```



Transition AS

If 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!)
- 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 onwards
- Cisco IOS-XE 2.3 onwards
- Cisco IOS 12.0(32)S12, 12.4(24)T, 12.2SRE, 12.2(33)SXI1 onwards
- Cisco NX-OS 4.0(1) onwards
- Quagga 0.99.10 (patches for 0.99.6)
- OpenBGPd 4.2 (patches for 3.9 & 4.0)
- Juniper JunOSe 4.1.0 & JunOS 9.1 onwards
- Redback SEOS
- Force10 FTOS7.7.1 onwards

http://as4.cluepon.net/index.php/Software_Support for a complete list

IPv6 and 4-byte ASN Update

Philip Smith <pfs@cisco.com>

PacNOG 8

Pohnpei, FSM

22nd-27th November 2010