New World BGP

Geoff Huston
January 2010
APNIC
16-bit AS Number Map
16-bit AS Number Map

Unadvertised AS Numbers

Advertised AS Numbers

RIR Pool AS Numbers

IANA Pool

Current AS States

AS Blocks (x256)

Advertised
Unadvertised
RIR_Pool
IANA
IETF
Unadvertised AS Numbers

Advertised AS Numbers

RIR Pool AS Numbers

IANA Pool

What's "USED"

What's "LEFT"
Consumption of AS Numbers

![Graph showing the consumption of AS numbers over time. The graph indicates a declining trend from 1996 to 2014.]
Consumption Rate

First order differential of allocations (smoothed)

Date

AS Assignments per day

Daily Allocations (smoothed)  Linear Best Fit
Consumption Rate

3 year average of 12.3 ASNs/day, with a growth factor slightly higher than simple linear
How long have we got?
How long have we got?

We are here!

Projection

Run Out Date

We are here!
16-bit AS Number Exhaustion

- We are exhausting the 16-bit AS Number pool
  - IANA will allocate its last AS number block in March 2011
  - RIPE will exhaust its 16 bit AS Number pool in December 2011

See [http://www.potaroo.net/tools/asns](http://www.potaroo.net/tools/asns)
This is not exactly news!

Combining these views

AS Number Exhaustion – March 2003
2003 Projection

Current AS Forecast

- The available AS number pool will exhaust in the timeframe of 2009-2011 if current AS use trends continue

**2009**
- no significant reclamation in old AS number space
- No coordinated effort to increase utilization density of AS numbers

**2011**
- reclamation and increased deployment efficiency
The Agenda for AS Transition

Developed in 2004 as a 4 step process:

1. IETF to complete BGP Standards to support transition mechanisms to 32-bit AS numbers
   ~2 years
2. RIRs to start making 32-bit AS numbers available
   ~½ year
3. Vendors to provide 32-bit AS number capable BGP implementations
   ~1 year
4. BGP networks to commence deployment
   • ready for deployment by 2008!
The Agenda for AS Transition

1. IETF to complete BGP Standards to support transition mechanisms to 32-bit AS numbers
2. RIRs to start making 32-bit AS numbers available
3. Vendors to provide 32-bit AS number capable BGP implementations
4. BGP networks to commence deployment
1. IETF Standards Activity

- 4-Byte AS Specification
  - Initial draft prepared in Feb 2001
    - Change BGP Attribute Definitions to extend AS components from 16 to 32 bits
    - Change BGP OPEN message to include capability negotiation for peer 4 byte support
    - Carry 32-bit AS path across 16-bit AS domains using new opaque transitive attribute (AS4_PATH)
    - Transition mechanism via translation and tunneling that allows piecemeal introduction of 4-byte AS numbers into the Internet
  - Specification ready for publication in late 2005
- IANA 32 bit AS number registry created in November 2006
- RFC 4893 published in May 2007
Objectives of the 32-bit ASN design:

- **Change as little as possible in the BGP spec**
- **Be ‘backward compatible’ with 16-bit AS BGP implementations**
  - Attempt to negotiate 32-bit capability when opening a BGP session
  - Automatically adjust behaviour when peering with 16-bit BGP peers and assume a 16-bit “persona” with 16-bit peers
  - Use 32-bit “persona” only with 32-bit peers
- **Preserve ‘basic’ AS semantics in BGP when peering with 16-bit AS BGP peers**
  - Preserve BGP’s loop detection properties
  - Preserve AS Path length metric properties
- **No ‘flag day’ transition**
  - Allow 16-bit BGP implementations to continue to operate indefinitely in a mixed 16 / 32-bit AS bgp world with complete address prefix reachability
  - Allow for piecemeal deployment of 32-bit BGP implementations
AS Path Semantics in BGP

- It’s a **path metric** where the length of the AS Path is used as in path selection

- It’s a **loop detector** where the presence of your own AS in a PATH is an indicator of a distance-vector “I’m-going-to-loop-to-infinity-unless-you-stop-me” loop

You don’t have to have an *entirely* accurate AS Path – but at a minimum you do have to have path-metric and loop-detecting properties for BGP to function correctly
32-bit AS Transition

- Think about this space as a set of NEW / OLD boundaries
- Define the NEW / OLD and the OLD / NEW transitions
- Preserve all BGP information at the transition interfaces
  - **Translate** 32-bit AS Path information into a 16-bit representation
  - **Tunnel** 32-bit AS Path information through 16-bit AS domain as an update attribute
32-bit AS Transition

- Think about this space as a set of NEW / OLD boundaries
- Define the NEW / OLD and the OLD / NEW transitions
- Preserve all BGP information at the transition interfaces
  - Translate 32-bit AS Path information into a 16-bit representation
  - Tunnel 32-bit AS Path information through 16-bit AS domain as an update attribute

**Translate all 32-bit-only AS numbers to AS23456**
32-bit AS Transition

- Think about this space as a set of NEW / OLD boundaries
- Define the NEW / OLD and the OLD / NEW transitions
- Preserve all BGP information at the transition interfaces
  - **Translate** 32-bit AS Path information into a 16-bit representation
  - **Tunnel** 32-bit AS Path information through 16-bit AS domain as an update attribute

![Diagram of 32-bit AS Transition]

**TUNNEL 32-bit AS_PATH as NEW_AS_PATH**
32-bit AS Transition

- Think about this space as a set of NEW / OLD boundaries
- Define the NEW / OLD and the OLD / NEW transitions
- Preserve all BGP information at the transition interfaces
  - Translate 32-bit AS Path information into a 16-bit representation
  - Tunnel 32-bit AS Path information through 16-bit AS domain as an update attribute
32-bit / 16-bit BGP Example

AS Path in the RIB

NEW 131072 — 131074 — 1221 — 4637 — 131075

NEW  OLD  OLD  NEW
32-bit / 16-bit BGP Example

AS Path in the RIB

AS Path Attribute in the UPDATE Message

131072 131074 1221 4637 131075
32-bit / 16-bit BGP Example

AS Path in the RIB

AS Path Attribute in the UPDATE Message
**32-bit / 16-bit BGP Example**

**AS Path in the RIB**
- NEW: 131072
- NEW: 131074
- OLD: 1221
- OLD: 4637
- NEW: 131075

**AS Path Attribute** in the UPDATE Message:
- 131072

**NEW_AS_PATH Attribute** in the UPDATE Message:
- 131072
- 23456 23456
- 131074 131072
32-bit / 16-bit BGP Example

AS Path in the RIB
- 131072
- 131074
- 1221
- 4637
- 131075

AS Path Attribute in the UPDATE Message
- 131072
- 23456 23456

NEW AS_PATH Attribute in the UPDATE Message
- 131074 131072
32-bit / 16-bit BGP Example

AS Path in the RIB:
131072 131074 1221 4637 131075

AS Path Attribute in the UPDATE Message:
131072 23456 23456

NEW_AS_PATH Attribute in the UPDATE Message:
131074 131072 131074 131072
32-bit / 16-bit BGP Example

AS Path in the RIB
- NEW 131072
- NEW 131074
- OLD 1221
- OLD 4637
- NEW 131075

AS Path Attribute in the UPDATE Message
- 131072
- 23456 23456
- 1221 23456 23456

NEW_AS_PATH Attribute in the UPDATE Message
- 131074 131072
- 131074 131072
32-bit / 16-bit BGP Example

**AS Path in the RIB**

131072 → 131074 → 1221 → 4637 → 131075

**AS Path Attribute in the UPDATE Message**

131072 → 23456 23456 → 1221 23456 23456 → 4637 1221 23456 23456

**NEW_AS_PATH Attribute in the UPDATE Message**

131074 131072 → 131074 131072 → 131074 131072

NEW  NEW  OLD  OLD  NEW
32-bit / 16-bit BGP Example

AS Path in the RIB

NEW 131072 — 131074 — 1221 — 4637 — 131075

AS Path Attribute in the UPDATE Message

NEW 131072 — 23456 23456 — 1221 23456 23456 — 4637 1221 131074 131072

NEW_AS_PATH Attribute in the UPDATE Message

NEW 131074 131072 — 131074 131072 — 131074 131072

OLD 1221 — 4637 — 131074

OLD 131072 — 23456 23456 — 1221 23456 23456

NEW 131072 — 131074 — 1221 — 4637 — 131075
32-bit / 16-bit BGP Example

AS Path in the RIB

32-bit AS Paths  Equivalent 16-bit AS Paths  32-bit AS Path

NEW 131072 — 131074 — 1221 — 4637 — 131075

NEW 131072

131074

1221

4637

131075

23456 23456

1221 23456 23456

4637 1221 131074 131072
Can old-BGP get Confused?

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0.1.0/24</td>
<td>Path 23456 23456</td>
</tr>
<tr>
<td>10.0.2.0/24</td>
<td>Path 23456</td>
</tr>
</tbody>
</table>

RIB Contents:

131072 — 131074 — 131076 — A — 1221 — B — 4637 — 131075
NO! BGP Nextthop is the key!

Traffic from AS 1221 to 10.0.1.0/24 will be forwarded on interface A
Traffic from AS 1221 to 10.0.2.0/24 will be forwarded on interface B

This is standard BGP behaviour – nothing changes here for BGP as it is used today
The Agenda for AS Transition

1. IETF to complete BGP Standards to support transition mechanisms to 32-bit AS numbers
2. RIRs to start making 32-bit AS numbers available
3. Vendors to provide 32-bit AS number capable BGP implementations
4. BGP networks to commence deployment
2. RIR ASN Allocation Policy

- Globally coordinated policy proposal 2005 / 2006
- Intended to avoid surprises and disappointment during the run-out of the 16-bit AS number space
- State clear milestones for vendors, ISPs and network admins for 32-bit ASN uptake
- Phased transition to the 32-bit AS number pool:
  - 2007 – 32 bit ASNs available upon request
  - 2009 – 32 bit ASNs available by default
  - 2010 – transition projected to be complete
The Agenda for AS Transition

1. IETF to complete BGP Standards to support transition mechanisms to 32-bit AS numbers
2. RIRs to start making 32-bit AS numbers available
3. Vendors to provide 32-bit AS number capable BGP implementations
4. BGP networks to commence deployment
## 3. Vendor Support in BGP

<table>
<thead>
<tr>
<th>Name</th>
<th>Version</th>
<th>Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcatel-Lucent SR OS</td>
<td>7.0</td>
<td>asplain</td>
</tr>
<tr>
<td>Arbor Peakflow SP</td>
<td>5.5</td>
<td>asplain</td>
</tr>
<tr>
<td>BIRD</td>
<td>1.0.12</td>
<td>asplain</td>
</tr>
<tr>
<td>Brocade (Foundry) IronWare</td>
<td>4.0.00 for the Netiron MLX and XMR, 2.8.00 for the Bigiron RX</td>
<td>asdot, asdot+, asplain</td>
</tr>
<tr>
<td>Cisco IOS</td>
<td>12.0(32)S12, 12.0(32)SY8, 12.2(33)SX11, 12.4(24)T</td>
<td>asdot (asplain planned for future)</td>
</tr>
<tr>
<td>Cisco IOS XE</td>
<td>2.3</td>
<td>asplain (asdot optional)</td>
</tr>
<tr>
<td>Cisco IOS XR</td>
<td>3.4(1)</td>
<td>asdot (asplain planned for 3.9)</td>
</tr>
<tr>
<td>Cisco NX-OS</td>
<td>4.0(1)</td>
<td>asdot (asplain planned for 4.1(3))</td>
</tr>
<tr>
<td>ExtremeXOS</td>
<td>Need Information</td>
<td>Need Information</td>
</tr>
<tr>
<td>Juniper JUNOS</td>
<td>9.1R1</td>
<td>asplain (asdot optional)</td>
</tr>
<tr>
<td>Juniper JUNOSe</td>
<td>4.1.0</td>
<td>asplain</td>
</tr>
<tr>
<td>Force10 FTOS</td>
<td>7.7.1.0</td>
<td>asplain (asdot, asdot+ optional)</td>
</tr>
<tr>
<td>OpenBGPD</td>
<td>4.2, patches for 3.9 and 4.0</td>
<td>asdot</td>
</tr>
<tr>
<td>Quagga</td>
<td>0.99.10, patches for 0.99.6 and other versions</td>
<td>asplain</td>
</tr>
<tr>
<td>Redback SEOS</td>
<td>2.0</td>
<td>ascoion (asplain planned for end of 2009)</td>
</tr>
</tbody>
</table>
The Agenda for AS Transition

1. IETF to complete BGP Standards to support transition mechanisms to 32-bit AS numbers
2. RIRs to start making 32-bit AS numbers available
3. Vendors to provide 32-bit AS number capable BGP implementations
4. BGP networks to commence deployment
RIR Allocation Data of 32-bit AS’s
32-bit ASNs in BGP
32-bit ASN Deployment

- Allocation status as of January 2010:
  - Advertised: 183
  - Unadvertised: 262

- In 2009 the RIRs allocated 4,761 ASNs
  - 4,445 were 16-bit ASNs
  - 316 were 32 bit ASNs
The Agenda for Transition

1. IETF to complete BGP Standards to support transition mechanisms to 32-bit AS numbers
2. RIRs to start making 32-bit AS numbers available
3. Vendors to provide 32-bit AS number capable BGP implementations
4. BGP networks to commence deployment

RFC 4893 – May 2007
32 bit AS Policy – 2006
Currently gathering pace
2009 – 2010
Lagging – badly!
How can we assist with 32-bit AS deployment?

- Information and education
  - Keep the community informed
  - Address some common misunderstandings about 4 byte AS numbers
- Supply chain pressure
  - Add 4 byte AS support to your “mandatory to support” in your next BGP purchase
NEW_AS_PATH Attribute

- BGP speakers in 16-bit AS domains should support **NEW_AS_PATH** as a transitive optional attribute in UPDATE messages
  - because that’s where the 32-bit path is hiding
  - That’s a “SHOULD” not a “MUST”, by the way
  - Its better if you do, but nothing fatally breaks if you don’t
    - Mixed 2 / 4 Byte loops will get detected in the 16-bit world as a fallback

*Default BGP configurations will do the right thing here*
NEW_AGGREGATOR Attribute

- BGP speakers in 16-bit AS domains should support **NEW_AGGREGATOR** as a transitive optional attribute in UPDATE messages
  - because that’s where the 32-bit Aggregator AS is hiding
  - That’s a “SHOULD” not a “MUST”, by the way
  - Its better if you do, but nothing fatally breaks if you don’t

*Default BGP configurations should do the right thing here*
AS 23456

- AS 23456 is going to appear in many 16-bit AS paths – both origin and transit

This is not an error – it’s a 16-bit token holder for a 32-bit AS number
Netflow and Sflow

- Netflow analyzers may need to be reviewed
  - Netflow version 9 supports 32-bit AS numbers
    - But may not report the 32-bit ASN unless the netflow collector is a 32-bit BGP
    - Does your analyzer support 32-bit AS numbers?
  - Netflow version 8 and earlier are 16-bit AS constrained
    - Which implies that you’ll be seeing AS 23456 more than you may want!

- Sflow
  - Appears to define a source and dest AS using a 32 bit field
  - So it *should* be ok!
BGP Communities

- If you want to explicitly signal to a 32-bit AS using communities in BGP then you will need to explicitly signal the 32-bit AS using **BGP Extended Communities**
  - Attempting to use AS 23456 in this context will have unintended consequences!

See:
- RFC 4630
- RFC 5568
BGP Memory requirements

- BGP memory requirements will increase
  - 32-bit BGP speakers will need twice the memory used to hold AS paths\(^1\)
  - 16-bit BGP speakers will need up to three times the memory used to hold AS paths plus NEW_AS_PATH extended community attribute\(^2\)

- 30,000 unique AS paths with an average length of 4 implies an additional memory requirement of 240Kb for 32-bit BGP and up to a further 480Kb for 16-bit BGP

\(^1\) - Not “twice the memory” but “twice the memory used for AS Path storage”
\(^2\) – Not “three times the memory”, but “three times the memory used for AS Path Storage”
BGP Bandwidth requirements

- BGP bandwidth requirements will increase (ever so slightly!)
  - 32-bit BGP speakers will need twice the size used to carry AS paths
  - 16-bit BGP speakers will need up to three times the size used to carry AS paths (factoring in the NEW_AS_PATH attribute)

- The update will grow by an average of 20 bytes, assuming an average AS path length of 4
Performance

- 32-bit to 16-bit BGP session startup may be considerably slower
  - The 32-bit speaker will need to compress all the AS Paths into their 16-bit equivalent prior to generating updates
    (assuming that the 16-bit Paths for Update messages are generated on demand)
  - This may take some time to compute for some 30,000 distinct AS Paths (depending on the internal structure of the BGP implementation)
Performance

- BGP convergence times may increase in some cases
  - Any instance of 16-bit BGP world destruction of the tunnelled NEW_AS_PATH attribute implies extended times on loop detection in order to fully complete prefix withdrawal
  - It's not that the withdrawal will loop forever, it's that the loop will take additional AS hops before it is detected in the 16-bit realm
  - The time to complete the withdrawal of a route may be extended
Proxy Aggregation

- If you proxy aggregate in the 16-bit world then make sure that the aggregate is strictly larger than the components
  - Or loop detection may be harder
    - As the AS Set object generated in the 16-bit word as a result of this proxy aggregation is not cleanly translatable into the 32-bit world, so 32-bit information is lost

- But proxy aggregation is not a common occurrence in today’s BGP environment
Mixed environments

- No dynamic capability for 16/32-bit ASN mode shift
  - You cannot flick from “16-bit OLD” to “32-bit NEW” mode within an active BGP session
  - You need to clear the session and then perform a clean start to trigger the initial capability exchange
Transition within an AS

- In a complex iBGP AS that wants to transition to using a 32-bit “home” AS then you are going to have to think about the transition VERY carefully
  - You can undertake this transition one router at a time, but care and attention are required
What happens when you have a customer / transit / peer with a 32-bit AS Number?

- What’s in the route registries and what your customers tell you about their AS and what’s in your OSS and your routing system will differ:
  - E.g.: AS 65538 needs to be auto-translated into AS 23456 in a number of places, including in your OSS
  - Your BGP routers may need to peer with AS 23456, transit across AS 23456, and have multiple customers on AS 23456 at the same time, while also understanding that these refer to different external parties
Related Systems

- Anything that wants to manipulate AS numbers, including your local support systems, scripts and databases
1. Someone out there is using 4 byte AS numbers. Do I have to upgrade my BGP to support 4-byte AS numbers in order to reach the prefixes that they are announcing?
Common Questions

1. Someone out there is using 4 byte AS numbers. Do I have to upgrade my BGP to support 4-byte AS numbers in order to reach the prefixes that they are announcing?

NO!
Common Questions

1. Someone out there is using 4 byte AS numbers. Do I have to upgrade my BGP to support 4-byte AS numbers in order to reach the prefixes that they are announcing?

- BGP uses a translation approach to mapping 4-byte AS numbers into a 2-byte AS number
- The 4 byte BGP speaker does all the translation work, so the existing BGP world will not need to upgrade to “see” these additional networks that lie within 4-byte ASNs in the routing space
- All that you will see is:
  - AS 23456 appearing in many AS paths
  - A very minor increase in memory use by BGP - associated with the storage of the additional AS4_PATH attribute
    - which contains the 4-byte AS path
    - but its an opaque transitive attribute to you, so you don’t care about its contents
2. My customers / peers/ upstreams are using 4-byte AS numbers. Do I have to upgrade my BGP to support 4-byte AS numbers?
2. My customers / peers/ upstreams are using 4-byte AS numbers. Do I have to upgrade my BGP to support 4-byte AS numbers?

NO!
2. My customers / peers/ upstreams are using 4-byte AS numbers. Do I have to upgrade my BGP to support 4-byte AS numbers?

- You need to do **nothing**!
- The new 4-byte BGP speaker figures out its talking to your old 2-byte BGP speaker and the 4-byte BGP speaker does **all** the work
  - it translates all instances of 4 byte AS numbers in the AS Path and Aggregator attributes to 23456 and stores the original 4-byte AS Path and Aggregator in new opaque transitive attributes (tunneling) before sending you the update
  - and restores the 4-byte information in any updates it received from you from the tunneled attribute information
2. My customers / peers/ upstreams are using 4-byte AS numbers. Do I have to upgrade my BGP to support 4-byte AS numbers?

- **But** you should’ve checked out your operational support system by now to make sure it can cope:
  - because you will need to support multiple peers / customers / upstreams who will have 4-byte AS numbers
  - and you will want to differentiate between them
  - but your routers’ BGP configs will be peering with AS 23456 for each instance
  - so your support system better be able to work this all out and not get confused!
3. Can I use communities for 4-byte ASNs?
Common Questions

3. Can I use communities for 4-byte ASNs?

YES and NO
3. Can I use communities for 4-byte ASNs?

- **NO** - if your BGP does not support RFC5668
  - because there is only a 2 byte field for the ASN in the conventional BGP community
  - You need to use a BGP extended community to define a set of communities for 4-byte origin and target AS values
  - This is specified in RFC5668
  - Ask your vendor when they will be supporting BGP extended communities with 4-byte ASNs – RFC5668

- **YES** – if your BGP supports RFC5668
Common Questions

4. If I upgrade BGP, will BGP crash?
4. If I upgrade BGP, will BGP crash?

MAYBE!
4. If I upgrade BGP, will BGP crash?

- Some Cisco implementations of BGP with 4-byte ASN support get unhappy when the number of elements in the AS path gets to over 1,000
- The maxas-limit setting is your friend
Common Questions

4. If I upgrade BGP, will BGP crash?
4. If I upgrade BGP, will BGP crash?

Also, there is the issue of the “standard” method for handling invalid components in the AS4_PATH attribute:

- AS Confederation path segments are declared invalid in the AS4_PATH attribute (RFC4893)
- If an optional attribute in an UPDATE is recognised then it must be checked, and if it is detected as invalid then a NOTIFICATION message must be sent and the BGP session is closed (RFC4271)
- A literal implementation of 4-byte AS BGP will be triggered to repeatedly tear down the local BGP session if AS Confederation elements are added into the AS PATH by a 4-byte AS BGP speaker, and then immediately propagated to a 2-byte AS BGP peer
4. If I upgrade BGP, will BGP crash?

The “safest” option is for the 4-byte BGP speaker to remove the offending element and reconstruct the AS Path as best it can, and log the error:

- Which appears to be what many BGP implementations now do
- And this consideration of “soft handling” of update errors applies to any BGP update, not only those with the AS4_PATH attribute, such as the use of AS0 in an AS Path
- The IETF is working on refining the BGP specification to treat such BGP update attribute errors with some circumspection, rather than a rather brutal “just drop the session” response!
Common Questions

5. I see AS 23456 in a 4-byte AS path – Is the Internet about the crash and die?
5. I see AS 23456 in a 4-byte AS path – Is the Internet about the crash and die?

Calm down!
Common Questions

5. I see AS 23456 in a 4-byte AS path – Is the Internet about the crash and die?

It may be abnormal, but it's not fatal
Common Questions

5. I see AS 23456 in a 4-byte AS path – Is the Internet about the crash and die?
5. I see AS 23456 in a 4-byte AS path – Is the Internet about the crash and die?

It may be abnormal, but it’s not fatal

- The AS Path is used for loop detection and path metric
- Even when AS23456 appears in the AS path, routing loops cannot form in BGP
  - but such “hybrid” loops may take a few more AS hops to detect and kill
32-bit ASN Resources

- **IETF Specifications**
  - RFC4893 – the 4-byte AS specification
  - draft-ietf-idr-rfc4893bis – working document that adds some further clarity and error handling to the specification

- **Documentation**
  - Exploring AS Numbers – Internet Protocol Journal, Vol 9, No 1
    (http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_9-1/autonomous_system_numbers.html)

- **Reports and Resources**
  - The AS Reports
    - http://www.potaroo.net/tools/asn16/
    - http://www.potaroo.net/tools/asn32/
  - ISP Resource Wiki for ASNs
    - http://as4.cluepon.net
Thank You