

January 2007

Geoff Huston

32-bit AS Numbers – The View from the old BGP World

The Internet Assigned Numbers Authority (IANA) has now expanded the AS number registry from its original 16 bit range (AS numbers 0 through 65535) to a 32 bit range (AS numbers 0 through 4,294,967,295). As of the 1st January 2007 the Regional Internet Registries are able to process requests for 32-bit AS numbers, if so requested. It may well be that we'll be seeing some of these extended length AS numbers being used in the public network very soon after that date.

This opening up of the AS number registry is a timely action, in that we were running low with AS numbers in the 16 bit AS number range. The recent rate of as number consumption was such that by October 2010 the AS number range would've been completely exhausted. The Regional Internet Registries have each adopted policies that allow ISPs to transition their networks into using this extended AS number range over some years, without the need for last second deployment of rushed changes in BGP. For the next two years, from 1 January 2007 until 31 December 2008, ISPs may specifically request an AS number from the extended 32-bit number pool (i.e. an AS number greater than 65535), but, by default they will be assigned an AS Number from the original 16 bit number pool (i.e. a number less than 65536). From 1 January 2009 the allocation practice will be reversed, and unless a 16-bit AS number is specifically requested, AS numbers will be allocated from the extended 32-bit number pool.

What are the implications for ISPs with this AS Number allocation policy? What should an ISP have as a prerequisite to requesting a 32-bit AS Number?

AS Numbers are used in the context of inter-domain routing, and in the BGP protocol in particular. If an ISP wants to use an AS number that is greater than 65535 then it will need to deploy a "new" version of BGP. That is, it will need to deploy a version of the BGP protocol in its routers that understands 32-bit AS numbers, as most existing BGP implementations use 16 bit data structures for AS numbers. So if you plan to deploy a 32 bit AS number in your network you are going to need to deploy a version of BGP that "understands" these 32 bit AS numbers.

But what about everyone else? What about the existing "old" BGP world that uses 16-bit AS numbers? Does the first public deployment of a 32 bit AS number force everyone else to also upgrade their versions of BGP? Even though every other network already has a 16-bit AS number, will every network also need to upgrade their BGP to see these new extended-length AS numbers? Will everyone need to apply some form of upgrade to their equipment and operational support systems before the first extended length AS number is used in the public Internet?

The issue here is actually one of "transition", and the way in which transition has been integrated into the specification of BGP properties to support 32-bit AS Numbers. The approach in the 32-bit AS number transition has been carefully constructed to be backward compatible, and the changes to BGP only affect "new" 32-bit BGP implementations. The reassuring news is that if you have a 16-bit AS number and are running 16-bit AS BGP today then you need to change nothing at all in your routers. The Internet will still work and you will continue to see routes to all advertised networks, irrespective of the existence of 32-bit AS numbers in the network. You will be able to send packets to those autonomous routing domains numbered from the 32-bit AS number space, and they will be able to send packets back to you. You don't need to upgrade your version of BGP, nor make any router configuration changes in your network. The Internet will work as intended without a break in connectivity. Nothing need change.

Well, almost nothing!

Some things might change, and here's a list of some of the things to think about if you are running an "old" BGP routing environment that supports only 16-bit AS numbers.

First, some background about the use of AS numbers in the BGP protocol. AS numbers appear in two critical BGP protocol messages, the OPEN message and the UPDATE message.

In the OPEN message a BGP speaker inserts its own AS in 16 bit MYAS field of the OPEN message. In order to be backward compatible with existing 16-bit BGP implementations it is not an option to be able to simply extend this field to 32 bits in length. Instead a new capability is used, where the 32-bit BGP announces to its peer that it uses 32 bit AS numbers, and additionally announces its local 32 bit AS number as the data part of the capability message. If the local AS number of the 32-bit BGP speak is less than 65536 (i.e. if the local AS number can fit into a 16 bit field) then the number is also used in the MYAS field of the OPEN message. Otherwise the reserved AS transition number (AS 23456) is used as MYAS.

In the BGP UPDATE message AS numbers are used in the AS PATH attribute and the AGGREGATOR attribute. In BGP the AS PATH attribute is used for two essential routing roles. It's a metric of routing path length, where, by default BGP will prefer a shorter AS PATH over a longer one for the same advertised prefix. It's also a routing loop detector, where each AS is capable of detecting, and preventing, a potential routing loop by seeing its own AS already in the AS PATH of received BGP advertisements. Strictly speaking, the AS PATH does not have to be an entirely accurate description of the inter-AS forwarding path to reach the destination prefix, but it does need to preserve these capabilities of a routing path metric and routing loop detection.

The transition mechanism for 32 bit BGP UPDATE messages is a combination of translation and tunnelling.

When passing a routing update into the 16-bit "old" BGP world the 32-bit "new" BGP speaker converts all AS numbers in the AS PATH to 16-bit values. If the AS number was between 0 and 65535 then all it does is strip off the leading 16 zero bits of the AS number value to perform this conversion. If the AS number is greater than 65535 then it translates the AS number of the special 16-bit value of AS transition value of 23456. If any AS number is translated in this way the "new" BGP speaker also saves a copy of the entire 32-bit AS PATH in a new transitive opaque community attribute called "NEW_AS_PATH". So when a "new" BGP speaker talks to an "old" BGP, then it maps the AS PATH to a 16 bit representation of the AS PATH, preserving its length and also preserving the values of all 16-bit AS numbers in the path. A similar transformation is applied to the AGGREGATOR attribute, mapping the AGGREGATOR attribute to the transition value of 23456 and storing the 32 bit value in a new transition opaque community attribute called "NEW_AGGREGATOR".

When passing a BGP UPDATE message from a 16-bit "old" BGP speaker to a 32-bit "new" BGP speaker then the opposite transformation is applied. All the AS numbers in the AS Path attribute are expanded to the equivalent 32-bit values by adding the leading 16 zero bits to the AS number value. If there is a NEW_AS_PATH community attribute in the UPDATE, then this AS string is substituted back into the AS PATH. If all goes well, the 32-bit BGP world sees an accurately re-constructed 32-bit AS PATH, preserving both AS path length metrics and the routing loop detection capability. Similarly the NEW_AGGREGATOR attribute value is mapped back into the AGGREGATOR field, as required.

This assumes that nothing "unusual" has happened as the BGP UPDATE message traversed the 16 bit "old" BGP world. It may be the case that the NEW_AS_PATH attribute has been dropped, or some form of proxy aggregation in the 16 bit world has garbled the AS PATH so that the reverse substitution cannot be performed. But even if the NEW_AS_PATH attribute is not present, or cannot be substituted back into the received AS PATH, its not a fatal condition for BGP itself. Even without this substitution the AS PATH length metric is preserved, and routing loop detection still can be performed, although in a degraded fashion. Potential routing loops that exist entirely within the 32-bit "new" BGP world are detected as normal, as are potential routing loops that sit entirely within the 16-bit "old" BGP world. And in the case of a mixed 16-bit and 32-bit potential routing loops the detection will still happen when the loop formation reaches the 16-bit "old" BGP world, as all the 16 bit AS number are preserved in the AS PATH in any case. So if the NEW_AS_PATH attribute is lost in the 16-bit "old" BGP world then the only casualty is speed of routing convergence, where it may take a number of additional AS hops for a potential routing loop to be detected and removed.

So, even if you do absolutely nothing in your 16-bit "old" BGP routing domain, Internet-wide routing will still work, and reachability information will still be propagated in useful forms. Nothing will 'break'. But there are some things to check, and maybe alter, in the larger environment of your operational support framework.

The implications for "old" world BGP routing domains include the following:

-
- The first implication for the old BGP world is that it is strongly preferred if the NEW_AS_PATH is carried as an optional transitive opaque community attribute when present. The 4-byte injection will mark this attribute with the optional and transitive attribute flags. 2-byte BGP speakers should avoid using local policy configurations that alter this attribute setting or remove this attribute from the prefix. If we are using standards-compliant language, then that's a "SHOULD", not a "MUST", by the way. Nothing will break if you don't but we'll see faster routing convergence if this attribute is left intact.
 - Similarly, its strongly preferred of the NEW_AGGREGATOR attribute is also carried as an optional transitive opaque community attribute when present. Again, nothing will break if you don't but we'll see faster routing convergence if this attribute is left intact.
 - The next implication is that the "old" 16-bit BGP world will see more and more instances of AS 23456 as both an originator of prefixes and as a transit provider. This is not a mistake, its just the only way that the 16-bit world can carry a place-holder for a 32-bit AS value.
 - An "old" BGP ISP may see routing peers, both as customers, peers and possibly upstream transit providers, using 32-bit AS numbers. But as your local BGP is a "old" world BGP your routers will not be aware of these 32-bit AS values. From your routers' perspective AS 23456 is going to start popping up both as a diverse prefix originator and a possibly as a ubiquitous transit provider. The ISP's operational support systems (OSS) should be able to store the corresponding AS numbers of these BGP routing peers as 32-bit number values, simply to avoid unnecessary confusion and potential ambiguity. So you probably should ensure that your OSS is 32-bit compliant for AS numbers, and is capable of storing and displaying configuration state for AS numbers in 32 bit format, even if you are running a version of BGP that supports only 16-bit AS numbers. Depending on the way in which the OSS has been designed, implementing this requirement may vary from the trivial to the extensive.
 - If you use this OSS to generate router configuration fragments, AS path filters and similar, then you may need to ensure that your OSS is capable of generating both 16-bit BGP configuration fragments and 32-bit configuration fragments. In the case of the 16-bit version the OSS will need to transform the locally held 32-bit AS number values into the 16-bit equivalent value of AS 23456.
 - Routing Registries will also need to be updated, allowing the registry's clients the ability to deposit registry entries that refer to 32-bit AS numbers. When using a Routing Registry to generate local configuration fragments, then the generated configuration entry will need to differ, depending on whether a 16-bit or 32-bit configuration fragment is required. For example, the route registry may have an entry relating to a routing domain of AS 1.2, but your "old" BGP router will need to be provided with a generated configuration fragment that refers to AS 23456.
 - If you filter based on AS numbers than any filter generator code that you might use will need to translate the 32-bit AS numbers stored in your local routing policy database into instances of AS 23456.
 - Many ISPs used directed community attributes to signal to a remote AS. A prefix that has explicit signalling to AS65505 uses a community attribute of "65505:123", for example. But this will not work if you wish to generate a signal to a 32-bit target AS. At the very least your BGP version should support expended community attributes (RFC4630) and also support the means of entering 32-bit AS numbers into these attributes (currently described in draft-rekhter-as4octet-ext-community-01.txt).
 - Your should also expect a modest increase on memory and bandwidth requirements for BGP. While nothing much is changing in your view of the routing world you will be carrying these NEW_AS_PATH transitive community attributes along with the prefixes, and the memory and bandwidth required to hold AS Paths will triple for "old" world BGP routers. That's not saying that total BGP memory demands will triple, just that requirement relating to AS path storage.
 - We might anticipate slightly poorer performance in routing. The specific cases where convergence times will extend are in those circumstances where the NEW_AS_PATH attribute is lost on transit through the "old" BGP

16-bit world. In such cases loop detection will take slightly longer, and this will have some level of impact on convergence times.

- There is no dynamic capability to support a change from 16-bit “old” BGP to 32-bit “new” BGP. When a routing domain wants to transition from a 16-bit to a 32-bit AS number then the BGP session will need to be reset via a complete shutdown and restart. The transition from “old” BGP to “new” BGP within a domain includes a number of considerations with respect to iBGP as well as eBGP sessions, and the transition will need to be planned very carefully.
- AS numbers are used in other contexts that are not directly related to BGP. Some implementations of MPLS VPNs use an AS number as part of a Route Distinguisher value [as described in RFC2547]. Identifying such uses of AS numbers, and ensuring that they either already support 32-bit fields for the AS number, or can be upgraded appropriately, is left as an exercise for the reader!

I've generated some code releases of BGP implementations with 4-Byte AS number supported hacked in, for those who are interested in experimenting with interactions with BGP with 4-Byte AS number support. This includes:

- OpenBGPD 3.9

<http://www.potaroo.net/tools/bgpd/openbgpd-3.9-4bas.tar.gz>

- OpenBGPD 3.9, FreeBSD-patched

<http://www.potaroo.net/tools/bgpd/openbgpd-3.9-4bas-freebsd.tar.gz>

- OpenBGPD 4.0

<http://www.potaroo.net/tools/bgpd/openbgpd-4.0-4bas.tar.gz>

Further Reading

- Exploring AS Numbers, Geoff Huston, ISP Column, August 2005. <<http://www.potaroo.net/ispcol/2005-08/as.html>>
- BGP Support for Four-octet AS Number Space, Quazar Voihara and Enke Chen, Internet Draft, Work in progress. <<http://draft-ietf-idr-as4bytes.potaroo.net>>

Disclaimer

The above views do not necessarily represent the views or positions of the Asia Pacific Network Information Centre.

About the Author

GEOFF HUSTON holds a B.Sc. and a M.Sc. from the Australian National University. He has been closely involved with the development of the Internet for many years, particularly within Australia, where he was responsible for the initial build of the Internet within the Australian academic and research sector. He is author of a number of Internet-related books, and is currently the Chief Scientist at APNIC, the Regional Internet Registry serving the Asia Pacific region. He was a member of the Internet Architecture Board from 1999 until 2005, and served on the Board of the Internet Society from 1992 until 2001.

<http://www.potaroo.net>