Potential Impacts of Four-Byte AS Numbers in Partial Deployment

Yuncheng Zhu, Junxiu Lu, Maoke Chen
Network Research Center, Tsinghua University, Beijing, P.R.China
{haoyu, ljx, mk}@cernet.edu.cn

Abstract—Inter-domain routing is suffering from the lack of identifiers for the rapidly growing number of Autonomous Systems (AS), and therefore the 4-byte AS number has been proposed along with the transition scenario. In a partial deployment environment, 4-byte-enabled Border Gateway Protocol (BGP) speakers have to exchange AS path vector with the old, 2-byte-only ones, via some way of translation and tunneling. The latter does not recognize and will possible lose the information about the 4-byte AS numbers. In this paper, we present a comprehensive analysis in order to identify the potential risks in deploying 4-byte AS numbers. It is found that 1) the loss of the AS4_PATH attributes will disable loop detection and lead to persistent oscillation in certain configuration; 2) AS_PATH-based match and COMMUNITY-based control may be ambiguously applied, and result in policy violations. Typical examples analyzed in real experiments validate these findings. Furthermore, AS path translation adds resource consumption for the routers, and it is necessary to identify the increment to router overhead of CPU time, memory and link bandwidth. Simulation results show that an increased overhead will impact the scalability of the global inter-domain routing system, substantially.

I. INTRODUCTION

Today’s Internet is made up of a great deal of Autonomous Systems (AS), which are also denoted as domains. Traditionally, a 2-byte AS number is used to identify a dedicated AS. The Border Gateway Protocol (BGP) [1], which is the de facto standard for inter-domain routing protocol, employs AS numbers for both its path-vector algorithm and its policy mechanism. For example, AS numbers in AS_PATH attributes are used for loop detection and AS_PATH-based match, while AS numbers are also used in COMMUNITY attributes to indicate the administrators of the communities.

According to historical and current assignment rates, AS number in 2-byte format is expected to run out in 2010 [2]. Thereby, A draft [3] is submitted to IETF as a proposed standard to support AS numbers in 4-byte format in the way of incremental deployment. The draft first define two types of BGP speakers: 4-byte-enabled speakers are called NEW speakers, while 2-byte-only ones are called OLD speakers. These two types of speakers are expected to co-exist for a significant long time. When a NEW speaker interoperates with another NEW one, there is no change in terms of protocol other than using 4-byte format where AS numbers are used. When a NEW speaker interoperates with an OLD one, the technique of a combination of translation and tunneling is used: the former will encode 4-byte entities into new transitive, optional path attributes such as “AS4_PATH,” and use a special AS number – AS23456, which is denoted as “AS_TRANS,” to take up the place in the original attributes when sending UPDATE messages, and decode them with the same method when receiving messages.

Although the detail of this scheme put much effort into interoperability among NEW and OLD speakers, since the latter does not recognize and will possible lose the information of the 4-byte AS numbers, it is suspected that this partial deployment will impact BGP’s path-vector algorithm and policy mechanism.

Huston [2] presented some implications for Internet Service Providers (ISP) with this AS number transition. He concludes that 1) AS4_PATH and AS4_AGGREGATOR attribute should be preserved by OLD speakers. 2) To avoid unnecessary confusion and potential ambiguity, OLD speakers should be able to store the corresponding AS numbers of their peers using 4-byte-only AS numbers. 3) Using AS numbers as identifiers in various situations like community control and local filtering should be avoided.

While in this paper, our analysis and experiments reveals that there are still many potential impacts of 4-byte AS numbers in partial deployment since OLD speakers have no idea of 4-byte AS numbers. We systematically discuss these problems both qualitatively and quantitatively.

First of all, BGP’s robustness, which is already controversial today [4][5], will be further impaired under the scene of partially deployed 4-byte AS numbers, since the same routing information will have different understanding by different types of speakers.

In terms of BGP’s behaviors of path-vector routing, AS numbers are used for loop detection, as specified in the protocol. However, loss of AS4_PATH attributes will disable such function and may lead to persistent oscillation in some situations. AS numbers are also widely employed in BGP’s policy routing for filtering and ranking routes. But AS_PATH-based match and COMMUNITY-based control are found behaving quite different from current practice and may be misused.

To validate our analysis, we further present concrete examples, showing the consequences of the conflicts between partially deployed 4-byte AS numbers and BGP inter-domain
routing practice.

On the other hand, as pointed out in [6], today’s Internet routing system is facing serious scaling problems. The ever increasing user population, as well as multiple other factors, has been driving the growth of Default Free Zone (DFZ) routing table size at an alarming rate.

As Huston [2] suggested, mixed-byte AS number in BGP routing will make BGP’s scalability even worse. Since more and more 4-byte AS numbers become visible in global routing table, the Routing Information Base (RIB) size will grow accordingly. Besides, as translation is needed between two types of routers, the burden on their CPU to process BGP updates will become heavier. In this paper, we quantified such overheads with a series of simulations, the results of which show that increased cost of router resource does harm the scalability of global routing system substantially.

The remaining part of the paper is organized as follows. In Section II, we give protocol level analysis on BGP including both path-vector algorithm and policy mechanism. Section III describes our experiment study on our routing testbed, emulating things will happen in real world. In Section IV, we focus on BGP’s scalability as AS are migrating to 4-byte AS numbers. We then conclude and discuss future work in Section V.

II. PROTOCOL LEVEL ANALYSIS

This section presents analysis on the potential conflicts between BGP routing practice and partially deployed 4-byte AS numbers. The first two parts discuss this from the aspects of path-vector routing and policy routing separately, while the last part addresses new problems with route aggregation.

A. Behaviors of Path-Vector Routing

As a path-vector routing protocol, BGP’s major distance-based metric is the length of AS path vector. Besides, any BGP speaker will deny a route if it contains the speaker’s AS number within its AS path vector to prevent routing loop.

When routing with partially deployed 4-byte AS numbers, an AS path vector can be divided into two parts: 2-byte-compatible AS numbers and 4-byte-only AS numbers. Certainly, the former is recognized by all speakers. However, the latter may be dropped since it is transferred in an optional, transitive “AS4_PATH” attribute.

As suggested in the protocol [1], which is the common case that current mainstream BGP routers support, OLD speakers will carry AS4_PATH attributes along. Through reconstructing AS4_PATH with AS4_PATH, where 4-byte-only AS number information is preserved, 4-byte-only AS numbers are recognized by NEW speakers.

However, it is possible that OLD speakers will drop AS4_PATH, which makes things entirely different. Since NEW speakers will not be able to restore 4-byte-only AS number information again, their ability to detect a routing loop is impaired. As previous researches [4][5] show that BGP routing is very fragile in robustness, this may impact global BGP routing seriously.

Without loss of generality, suppose a route has been traversed along an AS path, which end with (the left-most part) \(\langle AS_n, \ldots, AS_2, AS_1 \rangle\), where \(AS_1\) uses NEW speakers with 4-byte-only AS number, and there is one or more OLD speakers among \(AS_2, \ldots, AS_n\) who will drop the AS4_PATH attribute, assume \(AS_i\) is a AS using such OLD speakers.

Thus, loop detection will fail on \(AS_i\)’s speakers because of the loss of 4-byte-only AS number information, when this route pass through \(AS_i\) again. Although this route may be discarded because it has an AS_PATH attribute longer than it traverses to \(AS_1\) the first time, it may also be selected due to a higher local-preference. If this route still propagates from \(AS_i\), loop will be detected when it comes to \(AS_i\) again, but it will not stop this loop permanently. Instead, we will show in evaluation that \(AS_i\)’s loop detection will result in a persistent oscillation.

B. Behaviors of Policy Routing

BGP allows local routing policies to override distance-based metrics with policy-based metrics. Such policy routing can be summarized into route filtering and route ranking. For both filtering and ranking, there are two techniques widely used – AS_PATH-based match and COMMUNITY-based control. Since AS numbers lost their abilities to identify a dedicated AS for OLD speakers, this will make it more difficult to achieve the same routing policy for administrators of both NEW and OLD speakers.

1) AS_PATH-based match: Suppose an AS, using OLD speakers, with BGP peers of different business relationships, e.g. providers, customers and peering-links. This AS will have different policies on different peers. Today, the most common way to do this with AS_PATH-based match is using next-hop judgments: assigning different local-preference values according the last-hop AS numbers in the AS_PATH attributes of routes. But when two or more peers of different types migrate to 4-byte-only AS numbers, this method will not work any more since the last hop AS numbers in AS_PATH attributes of routes from different types will have the same value – AS_TRANS.

There exists a way to avoid this problem: using “route-map” command to identify routes from different peers and then applying different policies on them. The drawback of this way is obvious – administrators will need to configure the “route-map” command at the granularity of link other than AS. This unnecessary granularity refining will lead to lots of reconfiguration burden and may increase possibility of misconfiguration.

Although configuring at the granularity of link makes us be able to deal with next-hop routing policies, there are still other common ways of current routing policies with AS_PATH-based match, which are based on multi-hop AS numbers, e.g. AS_PATH regular expression, can not be settled with similar ways. With the ambiguity of AS numbers with OLD speakers, AS_PATH-based multi-hop filtering and ranking will not work for them any more.
2) COMMUNITY-based control: Communities provide a way to simplify routing decision by defining a group of destinations [7]. Extended attributes further simplifies this information labeling by the addition of a type field [8]. As Kammer pointed out [9], AS number based extended communities exhibit some co-operation and transition problems with partially deployed 4-byte AS numbers. Since it is not specified in the draft, different types of solutions may be provided. A coherent solution is needed on every 4-byte AS number aware routers.

For the best backward compatibility, Kammer had proposed a solution, which suggests to make a 2-byte AS number version for every 4-byte AS number specific extended community [10], so that OLD speakers can recognize such attribute once they are aware about extended community. However, we argue that this solution provides little help in many situations. When OLD speakers are receiving 4-byte AS number specific extended community attributes from many AS, they will recognize 2-byte AS number version of such attributes. Thus, OLD speakers will encounter previously mentioned “AS number ambiguity” problem, since they will mistake these attributes from different AS for attributes from the same AS – AS TRANS. Besides, this solution also gives speakers a mass of overhead in using extended communities since they will generate, store, propagate and calculate with two versions of the same extended community.

With the lack of standardization, we cannot assume any NEW speakers’ behavior on 4-byte AS number based extended community control with OLD speakers. IPv4-based extended community may be a solution for this problem. However, since IPv4-based extended community is not widely deployed, detailed study on this solution is required. Since COMMUNITY-based control is widely used in today’s BGP routing policy, problem with the usability of COMMUNITY-based control is severe to administrators.

C. Route Aggregation Concerns

Route aggregation summarizes routes so there are fewer routes to advertise across the Internet, thereby helps scaling the Internet’s routing table. When speakers are configured to aggregate a specific route, they will generate a new AS \_PATH attribute according to their knowledge of AS \_PATH attributes of those routes to be aggregated. However, with partially deployed 4-byte AS numbers, if those routes mostly traverse through different AS using 4-byte-only AS numbers, this new AS \_PATH generated by OLD speaker, including its length, will be quite different from which is generated by NEW speakers.

Once again, it is mainly because OLD speakers will regard different 4-byte-only AS numbers as the same AS number – AS TRANS. They will take different AS \_PATH segments as an identical one. In this case, the new AS \_PATH generated by OLD speaker will contain an AS-Sequence of multiple instances of AS TRANS instead of an AS-Set of different AS numbers, so that the length of the AS \_PATH will be longer than expected. Thus, with the same local-preference and other conditions, an AS will always prefer the route aggregated by a NEW speaker to one aggregated by an OLD speaker, which might hurt the interest of AS using OLD speakers. This problem will be worse when 4-byte-only AS numbers is prevalent in practice.

In one word, the major problems with BGP routing on protocol level with partially deployed 4-byte AS numbers is caused by AS numbers losing their abilities to identify a dedicated AS. Moreover, our study shows that partially deployed 4-byte AS number has no direct consequence on other issues like AS number pre-pending. The effects of AS number related issues are summarized in Table I, in which “Y” means attention must be paid to using corresponding speakers for that issue.

III. Experiment Study

To validate our analysis on potential problems above, we carried out several concrete experiments on a routing testbed which we built over a network experiment platform similar to PlanetLAB [11].

Since this transition just considers EBGP relationship, we can abstract an AS with one BGP speaker only, so connections between nodes are all EBGP sessions. We install two copies of Quagga routing software [12] version 0.99.6 onto about 50 nodes of the platform, one with the “4-byte AS number support” patch [13] and one without it. That is, every node of them can imitate either an AS with OLD speakers or an AS with NEW ones.

This routing testbed has a web-based central control node. EBGP sessions are activated on the demand of the experiment by the control node. Without manually configuring every experiment node, users can simply submit short scripts to construct experiment topology and employ some typical routing policies, like those in the relationships of customer-provider and peering, onto the topology. The control node will interpret such scripts into configurations and then deploy them for all selected nodes. Finally, it will periodically retrieve BGP routing information related to current experiment from experiment nodes and put them up onto a result page.

A. Routing Oscillation

As mentioned above, if there are some OLD speakers dropping transitive, optional path attributes they do not recognize, i.e. AS \_PATH, loop detection on NEW speakers may fail due to information lost.

<table>
<thead>
<tr>
<th>Speakers</th>
<th>OLD</th>
<th>NEW (^a)</th>
<th>NEW (^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loop detection</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>AS _PATH-based match</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>COMMUNITY-based control</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Route Aggregation</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>AS number pre-pending</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

\(^a\) NEW speakers using 2-byte-compatible AS numbers.

\(^b\) NEW speakers using 4-byte-only AS numbers.
In this experiment, we present such a situation when four AS form the AS graph shown in Fig. 1. Assume there is only one route destination \( d \), generated by AS1.1, who is the provider of AS2.1. AS2.1, AS3.1, and AS1001 happens to form a customer-provider ring, in which the former is the customer of the latter, e.g. AS2.1 is the customer of AS3.1. In this case, AS2.1 will not propagate routes received from AS1.1 to AS3.1. Besides, AS2.1 prefers AS3.1 to AS1.1 between its two providers. Finally, AS1001, which uses OLD speakers, drops AS4\_PATH as it propagates route information.\(^1\)

In this situation, due to information loss at AS1001, AS2.1 is always switching between the route from AS1.1 and the one from AS3.1, as the state transition graph shown in Fig. 2, which is obviously an oscillation that never converges on a stable routing. If AS2.1 is the only peer of AS1.1, Then \( d \) will be never reachable. Moreover, when AS2.1 has other peers accepting and propagating \( d \), the oscillation will form a severe update packet flood on global BGP routing, if mitigation techniques like route flap damping [14] is absent.

Though the relationship of customer-provider ring is seldom observed, policy settings leading to route loop is possible, especially when they may need to cooperate with AS using OLD speakers.

B. Community Control

Current solution of COMMUNITY-based control, although having already heavily considered the compatibility with OLD speakers, still will not work in some condition, especially when OLD speakers are confronted with AS number ambiguity.

Suppose six AS are inter-connected as Fig. 3 shows, where AS1.1 generates the single destination \( d \). AS3.1 and AS1001 collaborates with each other, so they both want AS1001’s traffic to \( d \) traverse through AS3.1. Administrator of AS1001 found that AS3.1 was using 4-byte-only AS number, so AS\_PATH based match cannot be employed. They want to solve this with extended community, therefore AS3.1 attaches extended community “soo 3.1:100” when propagating \( d \) while AS1001 sets a high local-preference when the translated extended community “soo 23456:100” is matched. However, AS1001 does not notice that AS4.1 also attaches extended community “soo 4.1:100” when propagating routes received from AS2.1 for some reason.

Result of our experiment shows that, under such condition, the two routes will get the same local-preference since AS1001 will match both extended communities with “soo 23456:100.” Consequently, AS1001 will choose the route according to the BGP identifiers of AS4.1 and AS5.1, and the collaboration is broken.

Although COMMUNITY-based control is, at some extent, compatible with OLD speakers. However, it is not very reliable for practical purpose. People should be extremely careful when using 4-byte AS number specific extended communities, especially when they may need to cooperate with AS using OLD speakers.

C. Route Aggregation

In this experiment, we show that route aggregation, if executed on OLD speakers suffering from AS number ambiguity, will generate wrong AS\_PATH information and make this aggregated route less competitive.

In the topology shown in Fig. 4, AS1.1 and AS2.1 generate a pair of adjacent network address prefixes, saying “10.0.0.0/24” and “10.0.1.0/24.” These prefixes are aggregated

\(^{1}\)Since Quagga does not support dropping optional, transitive attributes, we modified its code to achieve that.
Although BGP updates do not consume considerable bandwidth, since BGP control sessions are transferred along with other data streams, increased update packet size will make bandwidth more saturated. Since network congestion may result in BGP session failure [16], this problem will impair the stability of BGP routing.

To quantify these overheads, we simulated how routers’ resource usages increase along with the transition. In the simulation, we abstract each AS as one speaker since only EBGP sessions are considered. First, we obtain a set of current BGP data of core Internet, including a RIB dump and some update packet dump, from Oregon Route Views [17]. We then extract the topology of current BGP routing according to these data. On the topology, we randomly choose different percentage of AS to migrate to 4-byte AS numbers, without the change of amount and relationship of AS. Simulation is conducted 100 times for every parameter setting and for OLD and NEW speaker, respectively. For each data point, the confidence interval half-width is less than 1% with an alpha level of 0.05.

### A. Processing Overhead

NEW speakers need to do more calculation for sending and receiving updates to and from their OLD speaker peers, to translate and reconstruct 4-byte AS number information. This simulation uses a 7-day update packets dump2, supposing these data are gathered from a NEW speaker. As AS are migrating to 4-byte AS number and using NEW speakers, peers of this speaker may migrate, too. For every update packet, we judge that it needs additional calculation if it contains AS_PATH attribute and come from a peer using OLD speaker.

According to our result shown in Fig. 5(a), when none except this speaker is NEW speaker, this update process overhead reaches a maximum of about 93% of total update entries, with remaining 7% update packets contain only “withdraw” information. This percentage gradually decreases as more neighbors of it migrate to 4-byte AS number and use NEW speakers.

### B. Memory Consumption

The size of RIBs on both NEW and OLD speakers will increase as 4-byte AS numbers are stored in RIBs, instead of the 2-byte one on NEW speakers, and along with the 2-byte one on OLD speakers. To simulate the size of RIB, we use a snapshot of MRT format RIB dump file3, which is close to linear in size with actual RIB. Result of this simulation is shown in Fig. 5(b).

Increase of RIB on NEW speakers is a constant during transition because they save everything in 4-byte format. We translated the MRT file into 4-byte AS number format, and the increase is approx. 15%.

---

1. We use the dump files from Apr. 1, 2007 to Apr. 7, 2007, which have 25,583,892 entries of BGP update messages.
2. We use the dump file with timestamp 20070501.0048, which has 9,543,174 entries of routing information.
On the other hand, size of RIB on OLD speakers will increase as they will save AS4_PATH along with AS_PATH. We simulated this and compared result files to the original one, getting that size of RIB on OLD speakers will increase with the percentage of AS migrated, reaching a maximum of about 32% if all AS except itself use 4-byte-only AS numbers.

C. Bandwidth Consumption

As update packets containing 4-byte AS numbers, they will consume more bandwidth for sending and receiving update packets than before. Using the same data as the simulation for CPU processing time, we also simulated the increase of average size of update packets and result shows in Fig. 5(c).

For NEW speakers, they will communicate with NEW ones in 4-byte AS number format, and will communicate with OLD speakers with 2-byte AS number format including AS4_PATH information, whose size is larger than the former one. Thus, along with AS migrating 4-byte AS number, more 4-byte format AS number need to be encoded in AS4_PATH for peers using OLD speaker, which will augment the overhead; meanwhile, less peers will use OLD speakers, which will result in less overhead. Hence, the bandwidth overhead climbs to approx. 18% and then decreases to about 15% eventually.

Since OLD speakers are not aware about 4-byte AS numbers, they just transmit AS4_PATH without knowledge. Bandwidth overhead will continuously increase and the maximum overhead is about 33% when all AS except itself use 4-byte-only AS numbers.

The result of this simulation also confirms the result of simulation with memory overhead.

V. CONCLUSION

We have presented the potential impacts of 4-byte AS numbers in partial deployment, from both the protocol level issues and the scaling of the protocol. First, most of the problems with the practice of the protocol, either as a path-vector routing algorithm or as a policy routing mechanism, impairing the robustness of routing, are attributable to AS numbers lost their abilities to identify a dedicated AS in many condition, while some are caused by new formats introduced. Second, the problems with the scalability, making the efficiency of global routing even worse, which is severer to OLD speakers than NEW ones in the long term. This is mainly due to the nature of doubled size of the format of AS numbers.

Both vendors and network operators working with BGP speakers should be careful when dealing with routing with partially deployed 4-byte AS numbers, especially with AS_PATH-based match, COMMUNITY-based control and route aggregation. Therefore, we expect detailed guidelines will be brought forward in future, from research as well as practice, for different people to figure out these problems, including those multihoming issues which are mainly concerned by administrators of customer networks and the scalability problems which will bother administrators of transit network. Furthermore, the detailed evaluation of routing security with 4-byte AS numbers in partial deployment may be required.

REFERENCES