BGP属性に関するインシデント事例紹介
Bogonフィルタ未更新問題について

NTT Communications
Tomoya Yoshida
<yoshida@nttv6.jp>

内容

- BGPパス属性関連のインシデント事例
  - 2007年、2009年に発生した4つの事例
- Bogonフィルタに関する問題
  - 新規割り振りアドレスに対するフィルタが更新されず、特定のサイトへ到達性がない

BGPパス属性関連のインシデント事例

パス属性関連のインシデント

	PATH Attribute=0	AS_PATH too long	AS_CONFED_SET/SEQUENCE in AS4_PATH	AS4_PATH 0xE01100
事象	PA=0のAttributeを受信したリモートのルータが隣接Peerをリセット	長いAS_PATHを受信したリモートのルータが隣接Peerをリセット	AS4_PATHにConfede Private ASが混在し、それを解釈の上受信したルータが隣接Peerをリセット	AS4_PATH 0xE01100が発生し、それを受信したリモートルータが隣接Peerをリセット
発生時期	2007年12月	2009年2月	2009年3月	2009年8月
リモートアタックの可能性	あり	あり	あり	あり
ポイント	PA=0という、存在しないattributeを受信した時の動作	長すぎるAS_PATHを受信したときの動作	AS4_PATHには入るべきではないAS_CONFED_*を受信した時の動作	不正AS4_PATH を受信した時の動作
対処方法	該当経路をdiscardするOS等に更改	OS更改、上流ISPに広告抑制を適応してもらう、受信時に抑制	OS更改により該当経路の発生及び受信時にdiscard	OS更改により該当経路の発生及び受信時にdiscard
その他		観測場所によって発生有無が異なる	4byteAS関連	4byteAS関連

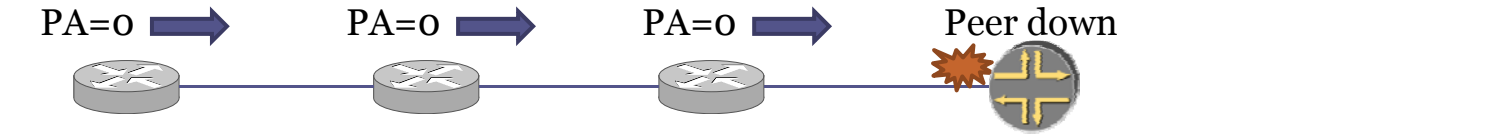
BGPパス属性一覧

Value	Code	Reference
1	ORIGIN	[RFC4271]
2	AS_PATH	[RFC4271]
3	NEXT_HOP	[RFC4271]
4	MULTI_EXIT_DISC	[RFC4271]
5	LOCAL_PREF	[RFC4271]
6	ATOMIC_AGGREGATE	[RFC4271]
7	AGGREGATOR	[RFC4271]
8	COMMUNITY	[RFC1997]
9	ORIGINATOR_ID	[RFC4456]
10	CLUSTER_LIST	[RFC4456]
11	DPA	[Chen, E., Bates, T., "Destination Preference Attribute for BGP", Work in progress, March 1996.]
12	ADVERTISER (Historic)	[RFC1863][RFC4223]
13	RCID_PATH / CLUSTER_ID (Historic)	[RFC1863][RFC4223]
14	MP_REACH_NLRI	[RFC4760]
15	MP_UNREACH_NLRI	[RFC4760]
16	EXTENDED COMMUNITIES	[Eric_Rosen][draft-ramachandra-bgp-ext-communities-00][RFC4360]
17	AS4_PATH	[RFC4893]
18	AS4_AGGREGATOR	[RFC4893]
19	SAFI Specific Attribute (SSA)	[Gargi_Nalawade][draft-kapoor-nalawade-idr-bgp-ssa-00][draft-nalawade-idr-mdt-safi-00][draft-wijnands-mt-discovery-00]
20	Connector Attribute	[Gargi_Nalawade][draft-kapoor-nalawade-idr-bgp-ssa-00][draft-nalawade-idr-mdt-safi-00][draft-wijnands-mt-discovery-00]
21	AS_PATHLIMIT (TEMPORARY - Expires 27 October 2007)	[draft-ietf-idr-as-pathlimit]
22	PMSI_TUNNEL (TEMPORARY - Expires 19 June 2008)	[draft-ietf-l3vpn-2547bis-mcast-bgp]
23	Tunnel Encapsulation Attribute	[RFC5512]
24	Traffic Engineering	[RFC5543]
25	IPv6 Address Specific Extended Community	[RFC-ietf-l3vpn-v6-ext-communities-02]
26-254	Unassigned	
255	Reserved for development	

http://beta.iana.org/assignments/bgp-parameters/bgp-parameters.xml#bgp-parameters-2
http://beta.iana.org/assignments/bgp-parameters/：BGP関連の各種パラメータ

1. 2007年12月 PATH Attribute=0

- BGPのPath Attribute Type=0 Prefixを受信した際にPeerのリセットが発生
  - RFC4271では、この事象に対して特定の動作は明記されていないが、Malformedと判断された場合にはNotificationの送信が明記されている
  - 多くの実装ではPeerをリセットせずにそのまま横に伝搬するため、リモートアタックとなり得る



PA=0の経路を送信

PA=0の経路をそのまま伝搬

PA=0の経路をそのまま伝搬

PA=0の経路を受信した際に、Malformed Attributeと判断し隣接Peerをリセット。その後再度Peerが再開され、再び同一経路を受信すると再度Peerリセットが発生。その繰り返しとなる。

- 対処・予防策
  - 該当経路を無視、あるいは取り除く等の対処がされたOSにアップグレードし、Peerリセットを回避することが可能
    - 対処方法は実装によりまちまちなので要確認
    - 本来的には、水際で不正経路を取り除くためにピアがリセットされるのが正しいかもしれないが、伝搬による被害を受けないための防衛手段は必要

2. 2009年2月 AS_PATH too long

- 極めて長いASパスが広報され、リモートのPeerが切断
  - 2009/2/17 1:23~2:23 (JST)

AS47868が
252個

```
94.125.216.0/21     *[BGP/170] 00:31:49, MED 22367, localpref 100
                      AS path: 3356 29113 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868 47868
47868 47868 47868 47868 47868 47868 47868 I
```

本事象の原因：MikroTikルータに起因した問題

- Prependする数量を書くところに、Prependで付加するAS番号を記載してしまい、かつルータが%256の結果として採用した
- Long AS_PATHの事象自体は大昔から発生しており、日常でもしばし観測されている。ここ最近は問題になっていなかった

47868 % 256 = 252
45307 % 256 = 251

AS path length	update time (UTC)	announced prefix	Prepending AS	originAS
257 (251 prepends)Matches Mikrotik Bug	2009-02-09 18:31	[AS path]	AS45307	AS45307
256 (252 prepends)Matches Mikrotik Bug	2009-02-16 17:06	[AS path]	AS47868	AS47868
249 (242 prepends)	2009-03-14 17:59	[AS path]	AS48262	AS48262
248 (242 prepends)	2009-03-14 23:35	[AS path]	AS48262	AS48262
203 (200 prepends)	2009-08-27 04:32	[AS path]	AS39786	AS39786
181 (176 prepends)Matches Mikrotik Bug	2009-02-19 21:02	[AS path]	AS20912	AS20912
149 (145 prepends)Matches Mikrotik Bug	2009-04-03 15:23	[AS path]	AS8337	AS8337
143 (129 prepends)	2009-02-25 19:56	[AS path]	AS44436	AS44436
142 (134 prepends)Matches Mikrotik Bug	2009-04-10 06:34	[AS path]	AS48262	AS48262
141 (134 prepends)Matches Mikrotik Bug	2009-03-29 21:37	[AS path]	AS48262	AS48262
141 (134 prepends)Matches Mikrotik Bug	2009-04-09 11:47	[AS path]	AS48262	AS48262
125 (121 prepends)	2009-07-28 15:52	[AS path]	AS39625	AS39625
118 (108 prepends)	2009-02-13 23:46	[AS path]	AS39625	AS39625
111 (100 prepends)	2009-08-30 02:18	[AS path]	AS38753	AS38753
105 (101 prepends)	2009-07-27 07:58	[AS path]	AS39625	AS39625
104 (101 prepends)	2009-10-27 21:38	[AS path]	AS39625	AS39625
104 (101 prepends)	2009-10-27 23:02	[AS path]	AS39625	AS39625
103 (101 prepends)	2009-10-12 10:17	[AS path]	AS39625	AS39625
103 (101 prepends)	2009-10-13 09:25	[AS path]	AS39625	AS39625
103 (101 prepends)	2009-10-19 19:53	[AS path]	AS39625	AS39625
102 (94 prepends)	2009-07-01 02:20	[AS path]	AS32258	AS32258
94 (91 prepends)	2009-10-14 00:52	[AS path]	AS39625	AS39625
94 (91 prepends)	2009-10-17 02:26	[AS path]	AS39625	AS39625
94 (91 prepends)	2009-10-20 03:37	[AS path]	AS39625	AS39625
94 (91 prepends)	2009-10-20 23:04	[AS path]	AS39625	AS39625
94 (91 prepends)	2009-10-22 14:14	[AS path]	AS39625	AS39625

http://www.bgpmon.net/maxASpath.php

影響と原因

- 一部のPeerがダウン
  - BGP Peer xx.xx.xx.xx DOWN (Malformed AS_PATH)

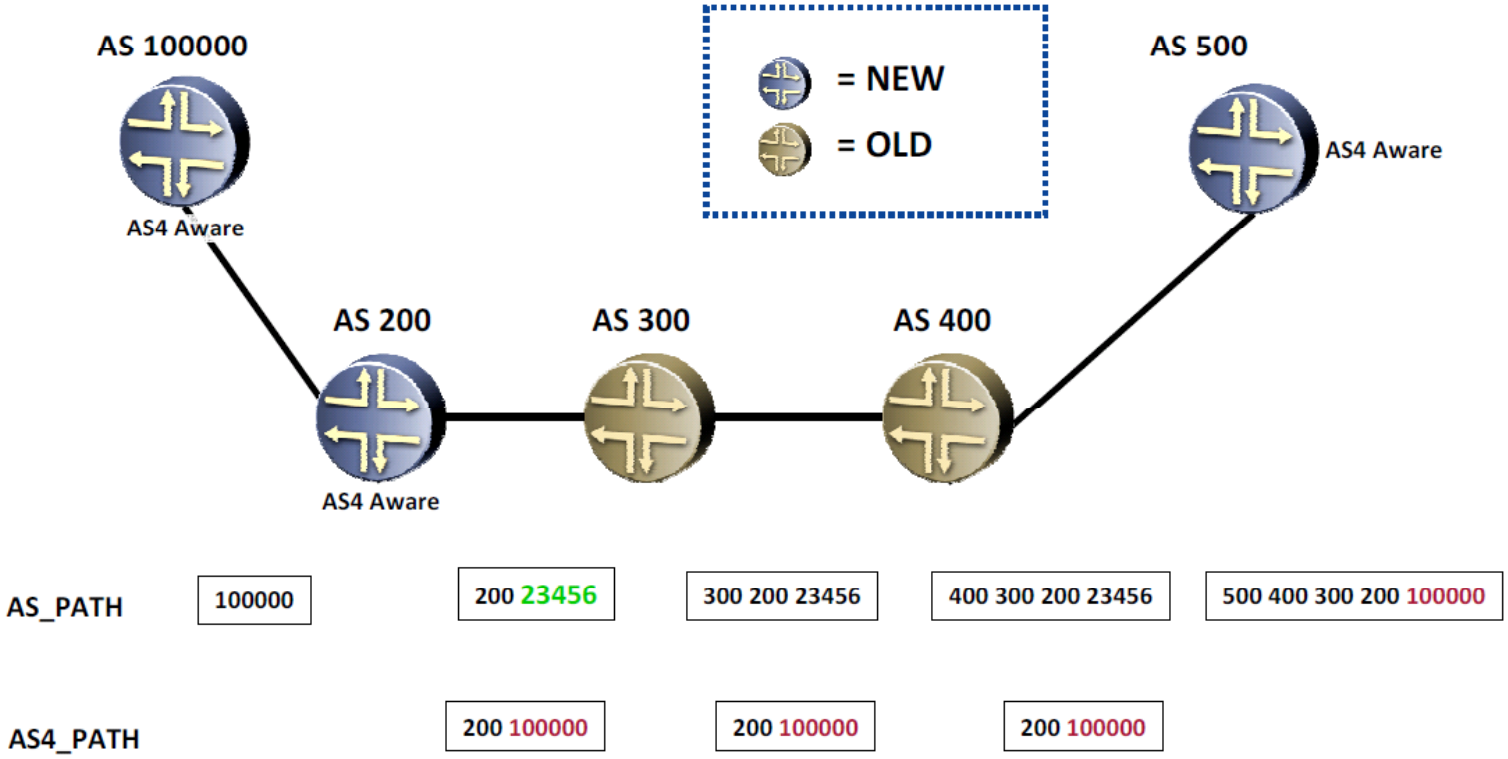
- 受信側の問題
  - 255以上のAS_PATH長のUpdateを受信した際にPeerがリセット
- 送信側の問題
  - 256以上のAS_PATH長としてUpdateを送信する際に不正なメッセージを送信してしまい、対向側でPeerをリセット

対策・予防策

- 根本的にはOSのバージョンアップで対処
- 予防策もあわせて実施
  - 上流ISPに一定の長さのAS_PATHを広告しないようにしてもらう
    - max-limitは受信時に適応しても有効とならないので注意
- 標準化へのフィードバック
  - 該当経路のみを遮断するなど
  - 現在IETFでも検討中

3. 2009年3月 AS_CONFED_SET/SEQUENCE in AS4_PATH

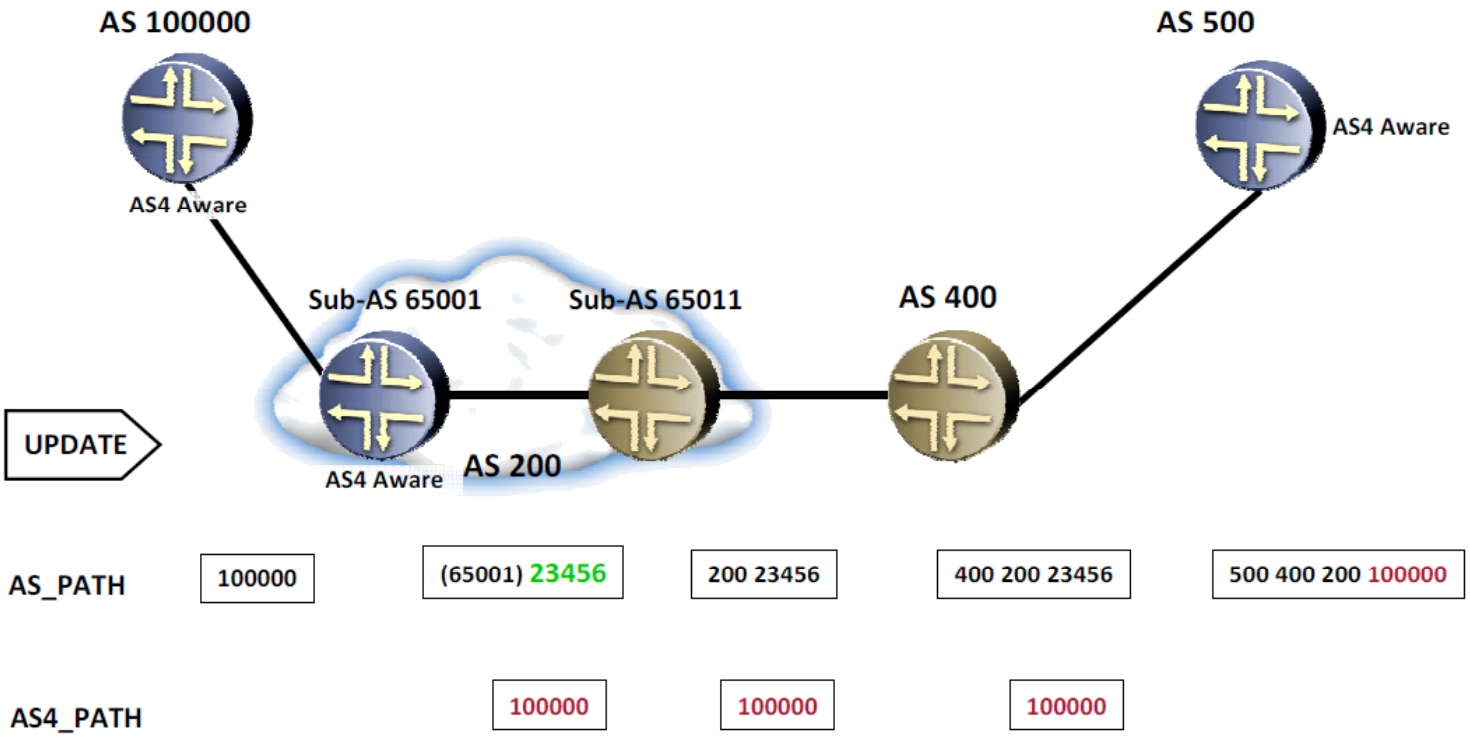
おさらい : OLD BGP speakerとNEW BGP Speaker



http://irs.ietf.to/past/docs_20090218/irs19_MK_IRS_AS4_PATH.pdf

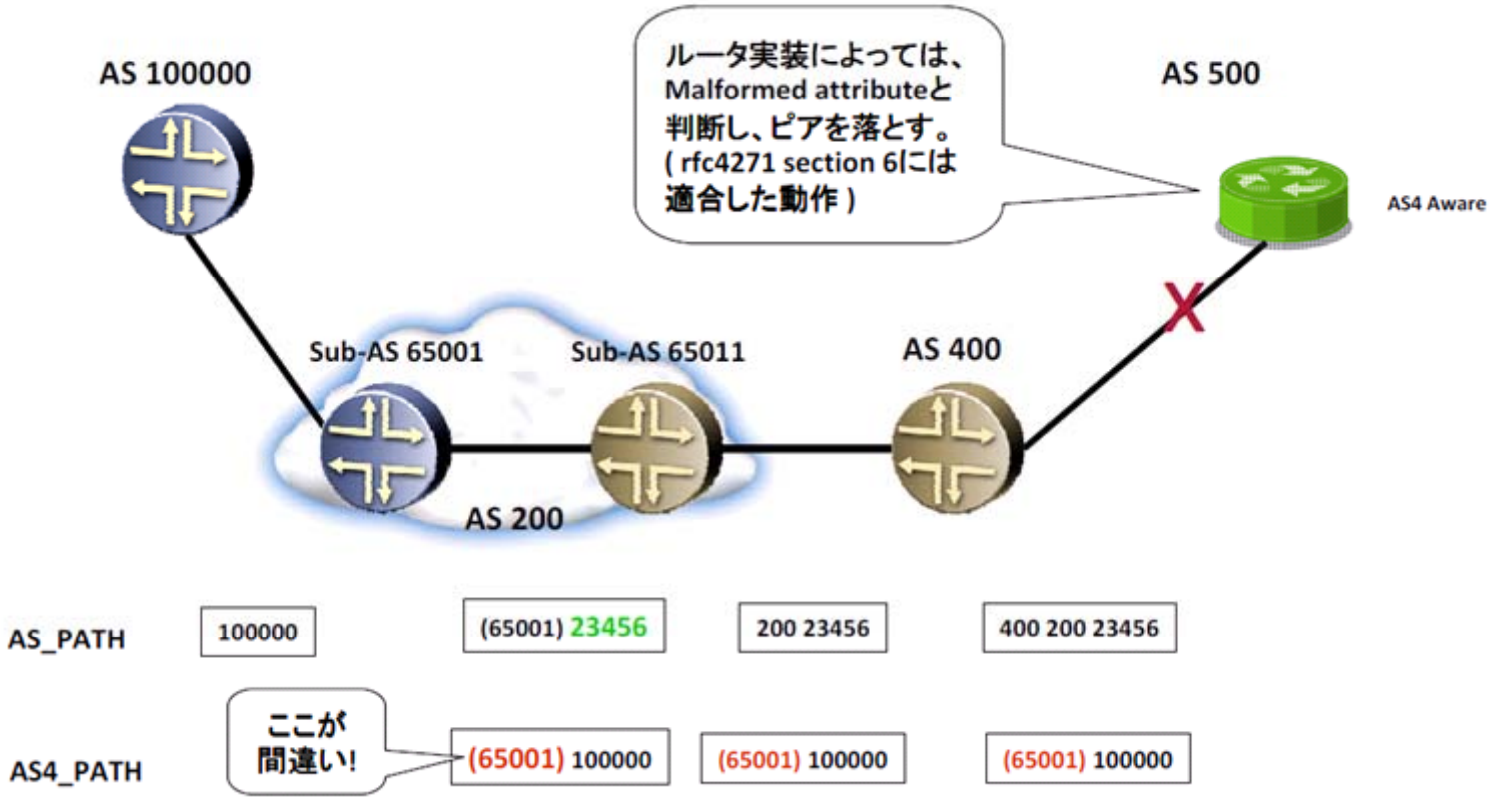
3. 2009年3月 AS_CONFED_SET/SEQUENCE in AS4_PATH

おさらい：Confederation時の正しい動作



http://irs.ietf.to/past/docs_20090218/irs19_MK_IRS_AS4_PATH.pdf

3. 2009年3月 AS_CONFED_SET/SEQUENCE in AS4_PATH



http://irs.ietf.to/past/docs_20090218/irs19_MK_IRS_AS4_PATH.pdf

3. 2009年3月 AS_CONFED_SET/SEQUENCE in AS4_PATH

- RFC4271では、Peerをresetすべきと記載されている
  - が、本事象ではResetしたほうがある種悪者扱い
  - VendorがRFCを無視してくれて救われたとも言われている
  - 本来は水際でPeerが切断されれば、transitされずに済むはずだが…
- 議論
  - 該当経路のみをignoreする実装が望ましいのでは？
  - 一部の経路不具合によりPeer全体がリセットされる必要はないのでは？
- 今後の方向性
  - Error handlingに関して詳細に定義する方向に
  - draft-ietf-idr-rfc4893bis-01.txt

> A NEW BGP speaker that receives a malformed AS4_PATH attribute in an UPDATE message from an OLD BGP speaker MUST discard the attribute, and continue processing the UPDATE message. The error SHOULD be logged locally for analysis.

・ DiscardしてPeerは継続すべきと明記

4. 2009年8月 AS4_PATH 0xE01100

- AS4_PATH「0xE01100」の経路が発生し、該当経路を受信した複数のリモートASのPeerがup/downを繰り返す
  - bgp_read_v4_message: NOTIFICATION received from xxx.xxx.xxx.xxx (ExternalAS 65400):
  - code 3 (Update Message Error) subcode 11 (AS path attribute problem), Data: e0 11 00

Value	Code	Reference
1	ORIGIN	[RFC4271]
2	AS_PATH	[RFC4271]
3	NEXT_HOP	[RFC4271]
4	MULTI_EXIT_DISC	[RFC4271]
5	LOCAL_PREF	[RFC4271]
6	ATOMIC_AGGREGATE	[RFC4271]
7	AGGREGATOR	[RFC4271]
8	COMMUNITY	[RFC1997]
9	ORIGINATOR_ID	[RFC4456]
10	CLUSTER_LIST	[RFC4456]
11	DPA	[Chen, E., Bates, T., "Destination Preference Attribute for BGP", Work in progress, March 1996.]
12	ADVERTISER (Historic)	[RFC1863][RFC4223]
13	RCID_PATH / CLUSTER_ID (Historic)	[RFC1863][RFC4223]
14	MP_REACH_NLRI	[RFC4760]
15	MP_UNREACH_NLRI	[RFC4760]
16	EXTENDED COMMUNITIES	[Eric_Rosen][draft-ramachandra-bgp-ext-communities-00][RFC4360]
17	AS4_PATH	[RFC4893]
18	AS4_AGGREGATOR	[RFC4893]
19	SAFI Specific Attribute (SSA)	[Gargi_Nalawade][draft-kapoor-nalawade-idr-bgp-ssa-00][draft-nalawade-idr-mdt-safi-00][draft-wijnands-mt-discovery-00]
20	Connector Attribute	[Gargi_Nalawade][draft-kapoor-nalawade-idr-bgp-ssa-00][draft-nalawade-idr-mdt-safi-00][draft-wijnands-mt-discovery-00]
21	AS_PATHLIMIT (TEMPORARY - Expires 27 October 2007)	[draft-ietf-idr-as-pathlimit]
22	PMSI_TUNNEL (TEMPORARY - Expires 19 June 2008)	[draft-ietf-l3vpn-2547bis-mcast-bgp]
23	Tunnel Encapsulation Attribute	[RFC5512]
24	Traffic Engineering	[RFC5543]
25-254	Unassigned	
255	Reserved for development	

e:
1:optional
1 transitive
1 partial
0 attribute length=1octet
0:
0x11=17
00: path attribute length

nanog@nanog.org 8/18

From: "Ballard, Eric" <Eric.Ballard@suddenlink.com>

Date: Mon, 17 Aug 2009 17:21:08 -0500

Subject: RE: Anyone else seeing "(invalid or corrupt AS path) 3 bytes E01100" ?

With the help from our transit providers and Cisco TAC the issues looks to be that ASXXXX is sending AS0 and causing the corruption when processed in our Cisco CRS routers.

....以下略

4. 2009年8月 AS4_PATH 0xE01100

BGP Update @ AS38639

Aug 18 02:07:25.304209 BGP RECV message type 2 (Update) length 61
Aug 18 02:07:25.304223 BGP RECV flags 0x40 code Origin(1): IGP
Aug 18 02:07:25.304236 BGP RECV flags 0x40 code ASPath(2) length 6: 4713 9354
Aug 18 02:07:25.304247 BGP RECV flags 0x40 code NextHop(3): x.x.x.x
Aug 18 02:07:25.304256 BGP RECV flags 0x80 code MultiExitDisc(4): 0
Aug 18 02:07:25.304266 BGP RECV flags 0xe0 code AS4Path(17) length 0: <null>
Aug 18 02:07:25.304287 BGP RECV xxx.xxx.192.0/20 , xxx.xxx.112.0/20
Aug 18 02:07:25.304368 bgp_rcv_nlri: xxx.xxx.192.0/20
Aug 18 02:07:25.304398 bgp_rcv_nlri: xxx.xxx.192.0/20 duplicate update
Aug 18 02:07:25.304409 bgp_rcv_nlri: xxx.xxx.112.0/20
Aug 18 02:07:25.304425 bgp_rcv_nlri: xxx.xxx.112.0/20 duplicate update

日常でBGPパケットをdumpしておくことが肝要です

BGPパス属性関連のインシデント事例まとめ

- 事象の把握・情報収集
  - xNOG等のMLで情報収集
  - 手元でBGPパケットをdumpしておくと便利
- 心構え
  - 想定外の経路が流れる可能性を認識しておく
- 対策
  - OSの入れ替えで対応できるケースもある
  - 上流ISPに依頼すれば対応可能なケースも
- より良い解決に向けての取り組み
  - **draft-ietf-idr-optional-transitive-01.txt**
    - 各ルータでの認識や解釈の違いを統一、より詳細に動作を定義

Bogonフィルタに関する問題

Bogonフィルタ

- Bogonルート
  - bogus(偽りの) という言葉から派生したもので、Bogonルートとは、本来広告されないPrefixのこと
- Bogonフィルタ
  - BogonルートをGWルータ等でフィルタし、本来広告されるべきではない経路をフィルタすること

用途	Prefix
プライベートアドレス	10.0.0.0/8、172.16.0.0/12、192.168.0.0/16
ループバックアドレス	127.0.0.0/8
リンクローカルアドレス	169.254.0.0/16
TEST-NET	192.0.2.0/24
ベンチマークテスト	198.18.0.0/15
マルチキャストアドレス	224.0.0.0/3
IANA Reserve	**現在 /8 x26**

IRR: fltr-unallocated object

$whois –h jpirr.nic.ad.jp fltr-unallocated

見割り振りの/8を表すIRRのObject

```
filter-set: fltr-unallocated
descr:      Unallocated (by IANA) IPv4 prefixes.
filter:     {1.0.0.0/8^+,
            5.0.0.0/8^+,
            14.0.0.0/8^+,
            23.0.0.0/8^+,
            27.0.0.0/8^+,
            31.0.0.0/8^+,
            36.0.0.0/8^+,
            37.0.0.0/8^+,
            39.0.0.0/8^+,
            42.0.0.0/8^+,
            49.0.0.0/8^+,
            50.0.0.0/8^+,
            100.0.0.0/8^+,
            101.0.0.0/8^+,
            102.0.0.0/8^+,
            103.0.0.0/8^+,
            104.0.0.0/8^+,
            105.0.0.0/8^+,
            106.0.0.0/8^+,
            107.0.0.0/8^+,
            176.0.0.0/8^+,
            177.0.0.0/8^+,
            179.0.0.0/8^+,
            181.0.0.0/8^+,
            185.0.0.0/8^+,
            223.0.0.0/8^+}
```

26個

```
admin-c:    TY6070JP
tech-c:     TY6070JP
remarks:    For the complete set of bogons, please see:
            fltr-martian - special use and reserved prefixes.
            fltr-bogons - fltr-unallocated + fltr-martian.
            http://www.cymru.com/Documents/bogon-list.html
notify:     irr-admin@nic.ad.jp
mnt-by:     MAINT-JPIRR
changed:    irr-admin@nic.ad.jp 20060712
changed:    irr-admin@nic.ad.jp 20060831  #RIPEx3
changed:    irr-admin@nic.ad.jp 20061011  #ARINx4
changed:    irr-admin@nic.ad.jp 20070118  #APNICx5
changed:    irr-admin@nic.ad.jp 20070330  #RIPEx2
changed:    irr-admin@nic.ad.jp 20070731  #RIPEx2
changed:    irr-admin@nic.ad.jp 20071001  #LACNICx2
changed:    irr-admin@nic.ad.jp 20071030  #APNICx2
changed:    irr-admin@nic.ad.jp 20080215  #ARINx2
changed:    irr-admin@nic.ad.jp 20080215  #add 014/8
changed:    irr-admin@nic.ad.jp 20080529  #APNIC 112/8 113/8
changed:    irr-admin@nic.ad.jp 20081104  #AfriNIC 197/8
changed:    irr-admin@nic.ad.jp 20081113  #APNIC 110/8 111/8
changed:    irr-admin@nic.ad.jp 20081224  #ARIN 108/8 184/8
changed:    irr-admin@nic.ad.jp 20090204  #RIPE NCC 109/8 178/8
changed:    irr-admin@nic.ad.jp 20090501  #APNIC 180/8 183/8
changed:    irr-admin@nic.ad.jp 20090804  #APNIC 175/8 182/8
changed:    irr-admin@nic.ad.jp 20090922  #RIPE 2/8 46/
source:     JPIRR
```

Allocationされていない/8を記載したIRRのオブジェクトもの
現在JPNICにて本オブジェクトを随時更新

最近特にbogonフィルタ問題が顕著に

- 過去のIANAリザーブ空間を元にフィルタを実施しているため、新規アドレス空間払い出し後に該当経路がきちんと利用できない
- IPv4アドレスが残り少なくなってきているため、IANA->RIRへの新規/8の割り振り後、それほど時間が経過せずにLIRへアドレスが配布
  - フィルタ更新時間の猶予が徐々に短くなってきており、残りが少なくなるに従いより顕著になる可能性がある
  - /8の残りが少なくなると、フィルタ自体の意味がなくなってくる
- フィルタを逆にかけすぎる形になってしまうので、到達性に問題が発生する。実施の際には適切なタイミングでの更新が必要。

対応策

- 自分側のフィルタ更新
  - 新規割り振り時にはきちんと更新する
- 相手側のフィルタ更新問題
  - 到達出来ないサイトに直接コンタクトしてみる
  - xNOGのML等に投稿
  - 大抵他の人も同じ問題に遭遇している

NANOGへのIANAからのリマインダ

On 09/10/2009 4:22, "Matthew Walster" <matthew at walster.org> wrote:

> A customer of mine is reporting that there are a large number of addresses
> he can not reach with his addresses in the 109/8 range. This was
> declassified as a BOGON and assigned by IANA to RIPE in January 2009.
>
> If you have a manually updated BOGON list, can I please ask that you review
> it and update it as soon as possible please? His addresses in 89/8 and 83/8
> work just fine, hence this presumption of BOGON filtering.

This might be a good moment to list all the /8s allocated so far this year.

046/8	RIPE NCC	2009-09	whois.ripe.net	ALLOCATED
002/8	RIPE NCC	2009-09	whois.ripe.net	ALLOCATED
182/8	APNIC	2009-08	whois.apnic.net	ALLOCATED
175/8	APNIC	2009-08	whois.apnic.net	ALLOCATED
183/8	APNIC	2009-04	whois.apnic.net	ALLOCATED
180/8	APNIC	2009-04	whois.apnic.net	ALLOCATED
178/8	RIPE NCC	2009-01	whois.ripe.net	ALLOCATED
109/8	RIPE NCC	2009-01	whois.ripe.net	ALLOCATED

Also, I'd like to mention that if you ever want to check your filters
against the registry, we have made the columns sortable. It's now nice and
easy to identify newly allocated /8s.

http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xml

Regards,

Leo Vegoda