頭の片隅においておきたい 4-octet ASのお話

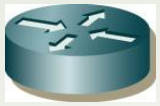
JANOG25
NEC BIGLOBE, Ltd.
川村 聖一

4-octet ASとBGPの簡単なおさらい

- RFC4893 “BGP Support for Four-octet AS Number Space”
- NEW BGP Speaker（4-octetに対応したルータ）とOLD BGP Speaker（4-octetに対応していないルータ）のやり取りに注目。
  - NEW-OLD間のPeerは2-octet AS間のみ（当然）
  - 経路広告はNEW側がAS_TRANS、AS4_PATHを使って上手くやってくれる。

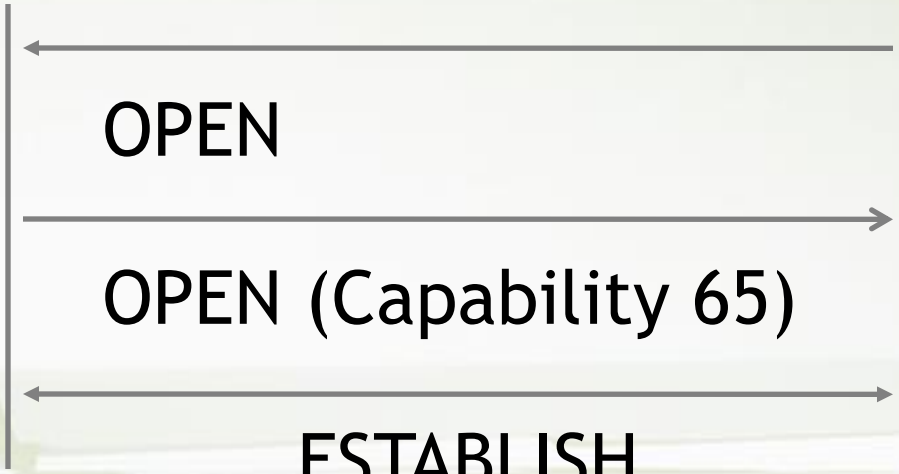
- NEW BGP Speakerは、相手が誰であろうとCapability Code 65 (Support for 4-octet AS number capability) を送信します。

NEW BGP と OLD BGPのPeer



AS 65001
(4-octet対応)

AS 65002
(4-octet未対応)



Neighbor capabilities:
Four-octets ASN
Capability: advertised

Neighbor capabilities:
＜記述なし＞

Capabilityの取扱

- RFC 3392 “Capabilities Advertisement with BGP-4”にこんな記載があります。Section3

> If a BGP speaker that supports a certain capability determines that its peer doesn't support this capability, the speaker MAY send a NOTIFICATION message to the peer, and terminate peering (see Section "Extensions to Error Handling" for more details). The Error Subcode in the message is set to Unsupported Capability. The message SHOULD contain the capability (capabilities) that causes the speaker to send the message. The decision to send the message and terminate peering is local to the speaker. If terminated, such peering SHOULD NOT be re-established automatically.

Capabilityの取扱

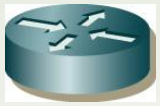
- RFC 3392 “Capabilities Advertisement with BGP-4”にこんな記載があります。Section3

■ 特定のBGP Capability を実装しているルータが、
■ 相手側ルータがそのケーパビリティをサポートしていないと判断した場合、
■ NOTIFICATIONを送信した上でCapabilityを実装している側がセッションを切断することができる。
■ 切断する際のError Subcodeは、“Unsupported Capability” (Subcode 7)。Dataフィールドに、原因となったCapabilityを記載するべき。

OLD Speakerとの親和性

- NEW Speakerは、相手がOLD Speakerだとわかるので、AS_TRANS、AS4_PATHを使ってうまくやってくる。BGPをTeardownすることはしない。
- RFC4893はbackward compatibilityをしっかり考慮して書いてある。
- NEW Speakerは、相手がOLD Speakerだからと言ってUnsupported Capabilityを出すことは絶対ない。
- ところが・・・

ところがこんな事が・・・



AS 65001
(4-octet対応)

AS 65002
(4-octet未対応)

OPEN

OPEN (Capability 65)

NOTIFICATION 2/7
(Unsupported Capability)

Peerがあがらない！

2/7 (Unsupported Capability)は、「新しい側」が出すメッセージ！BGPの実装ミスだ！！！

実際に起きますか？

- 複数国産メーカーでこの挙動が確認されています。
- では知らないCapability来たらどうするべきか？
  - 無視して、Sessionを確立させて、そのまま何もなかったように日常をすごすべき。
  - 実はRFC3392に明確には書いていない。
- RFC 5492でここが明確化される。
- If a BGP speaker receives from its peer a capability that it does not itself support or recognize, it MUST ignore that capability. In particular, the Unsupported Capability NOTIFICATION message MUST NOT be generated and the BGP session MUST NOT be terminated in response to reception of a capability that is not supported by the local speaker.

ミスインプリが多い理由

- RFC3392のSection5だけ読むと、そうとれなくもない。

- 5. Extensions to Error Handling

  This document defines new Error Subcode - Unsupported Capability. The value of this Subcode is 7. The Data field in the NOTIFICATION message SHOULD list the set of capabilities that cause the speaker to send the message. Each such capability is encoded the same way as it would be encoded in the OPEN message.

- Section3をちゃんと読んでいれば、2/7を出すのは、Capabilityをサポートしている側ですと書いてあります。

何が問題？

- 昨今のOSは、バージョンアップすれば4-octet ASに対応している場合が多い。
  - 4-octet ASに対応しているつもりでなくても、Capability-65をデフォルトで出してしまう。
- アップデート頻度が低い相互接続ネットワークで・・・
  - IP-VPN
  - 法人の接続
  - 閉域(もしくは社内)網
  - ホールセール網
- ユーザがある日突然バージョンアップ/機種リプレースすると、xSP側が古いOSだったりして、BGPがつながらなくなってしまう。(逆もそう)

Workaround

- dont-capability-negotiate
  - quagga, Ciscoなど。古くからあるコマンド。IOSではhidden
  - 全Capabilityおくらない
  - イマイチ
- dont-capability-negotiate four-octets-as
  - quaggaは、0.99.12ではなかった。Ciscoは12.2(33)SRE, 12.2(33)XNE, 15.0(1)Mでhiddenとして実装
  - Capability65だけおくらない
  - う〜ん。どうだろ。最終手段か

- ↑は新しい側のworkaround。
- 知らないCapabilityは無視するのが正しい実装。
- 今使ってるBGPルータが、どういう実装なのかはしっかり把握しておきましょう。

Special Thanks

- Cisco 土屋さん
- Juniper　國岡さん
- Tools Team　谷津さん
- コードをなおしてくれたベンダーのみなさま
- 対応してくれたPeerのみなさま