

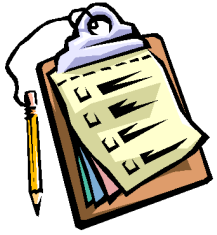


cgi.br nic.br registro.br

Entendendo ASNs BGP de 4 bytes



André Gustavo Albuquerque
aalbuque @ cisco.com
19.Jun.2009



Agenda

- Histórico e Motivações
- Visão Geral
- Detalhamento Operacional
- Impactos Operacionais com a Utilização de 4B ASN

Histórico e Motivações

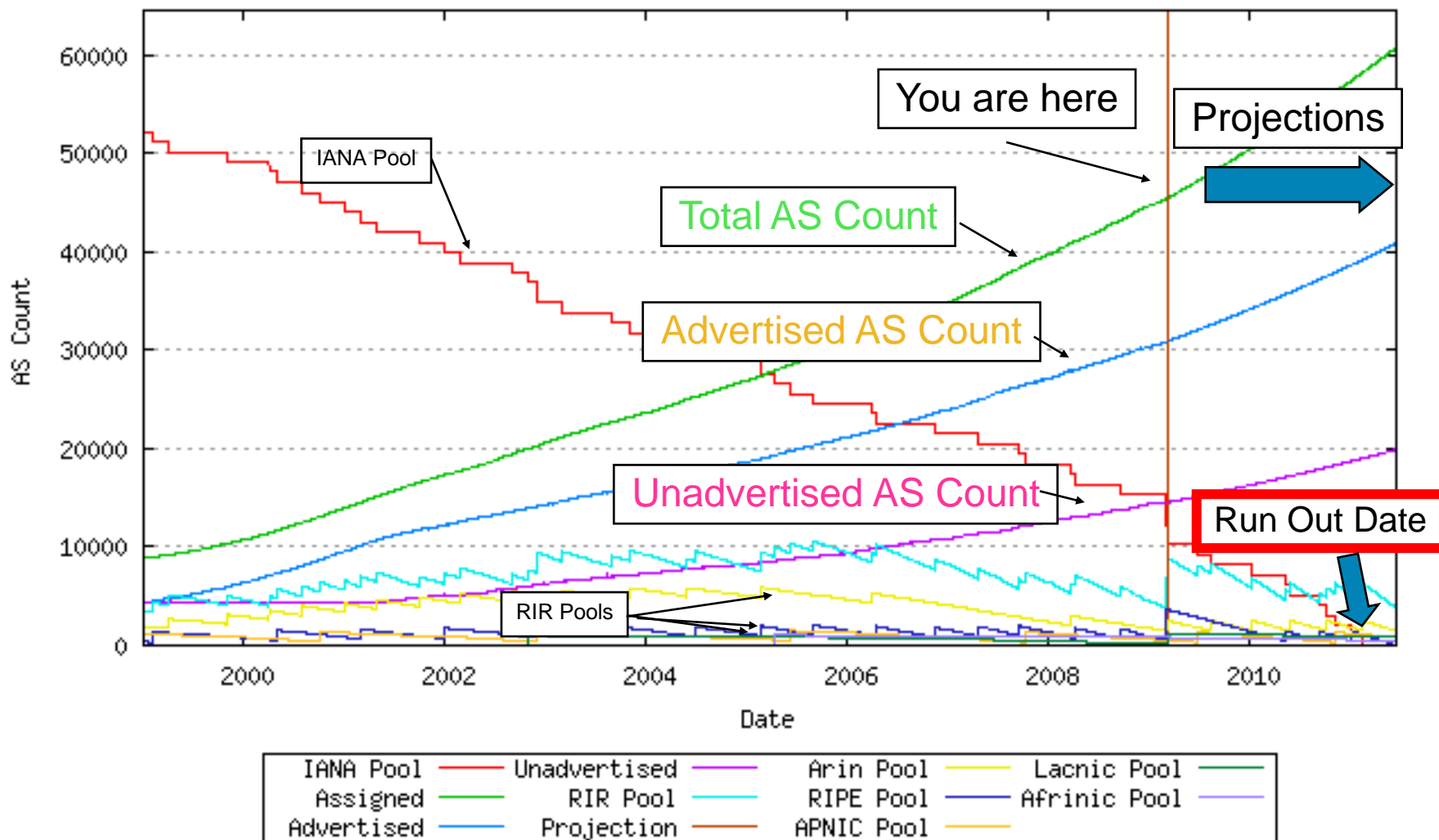


Histórico e Motivação

- Roteadores BGP utilizavam apenas ASNs de 16 bits.
- ASNs reservados pela IANA
 - 64512 a 65534 para utilização privada
 - 0, 55296-64511 e 65535 são reservados e não devem ser utilizados
- Existem 64.494 ASNs disponíveis para alocação, dos quais 55.294 já foram alocados (*).
- O pool de ASN de 2 bytes tem previsão de exaustão para **6 de agosto de 2011** (*).

* Fonte: <http://www.potaroo.net/tools/asns> (13.Jun.2009)

Motivação



* Fonte: <http://www.potaroo.net/tools/asns> (10.Mar.2009)

Política de Alocação de ASN (LACNIC)

- Três etapas para a alocação de ASN:
 1. **01.Jan.2007**: novas alocações com ASNs de 16 bits a menos que o solicitante especifique ASN de 32 bits em sua requisição.
 2. **01.Jan.2009**: novas alocações com ASNs de 32 bits a menos que o solicitante especifique ASN de 16 bits em sua requisição.
 3. **01.Jan.2010**: o registro deixará de fazer distinção entre ASNs de 16 bits e ASNs de 32 bits, e distribuirá ASNs a partir de um espaço não diferenciado de números de AS de 32 bits.

Fonte: <http://www.lacnic.org/pt/politicas/manual4.html> (11.mar.2009)

Visão Geral



RFC 4893 – Support for 4 Byte ASN

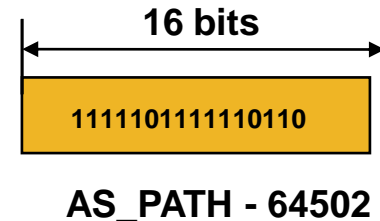
- 4-byte AS support is advertised via BGP capability negotiation
 - Speakers who support 4-byte AS are known as NEW BGP speakers
 - Those who do not are known as OLD BGP speakers
- New Reserved AS#
 - AS_TRANS = AS #23456
 - 2-byte placeholder for a 4-byte AS number
 - Used for backward compatibility between OLD and NEW BGP speakers
- Two new attributes, both are “optional transitive”
 - AS4_AGGREGATOR
 - AS4_PATH

Mappable and Non-Mappable ASNs

OLD SPEAKER



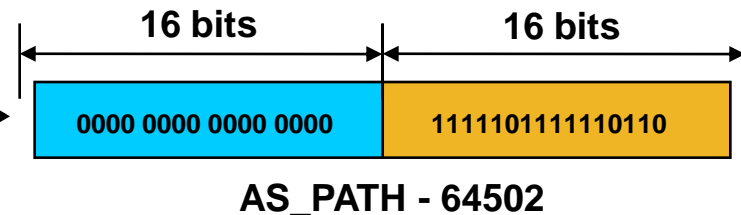
Mappable in 2 octets AS_PATH field



NEW SPEAKER



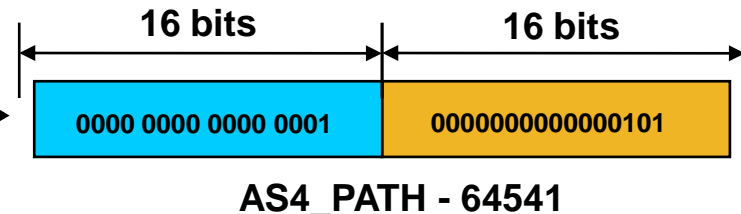
Mappable in 2 octets AS_PATH field
Only the lower order bits are set.



NEW SPEAKER



Non-mappable in 2 octets AS_PATH
field. When peering with an OLD
SPEAKER, non-mappable ASN is
carried by AS4_PATH attribute.



4 Byte AS Number Notation – RFC 5396

- **asplain** - Scheme formats all AS numbers using decimal integer notation. Example 65001 is represented as “65001”, 65546 as “65546”
- **asdot+** - Scheme formats all AS numbers using a notation of two integer values joined by a period character: <high order 16-bit value in decimal>.<low order 16-bit value in decimal. Example 65001 as “0.65001” and 65546 as “1.10”
- **asdot** - Scheme formats all AS numbers less than 65536 as asplain and AS numbers equal to 65536 and greater using asdot+. Example 65001 will be presented as “65001” and 65546 as “1.10”.

Motivation Behind “asplain” Format

- Current Systems (provisioning, monitoring etc) support only decimal integer format. Additional work required for these systems to support both formats.
- ISP systems need to support both “asplain” and “asdot” notation for the 4-byte ASN provided by ISP customers
- Impacts filtering policies based on regular expression.
- Internal Scripts used for Network Management and troubleshooting.
- Network Operators prefer “asplain” format (NANOG).

asplain vs asdot(+)

- Problem:

 - ^[0-9]+\$ matches any ASN (16-bit and asplain)

 - This and equivalents extensively used in BGP multihoming configurations for traffic engineering

- Equivalent regexp for asdot is

 - ^([0-9]+)|([0-9]+\.[0-9]+)\$

- Equivalent regexp for asdot+ is

 - ^[0-9]+\.[0-9]+\$

⇒ BGP policy regular expressions would need to be rewritten

IANA Assignments

- Using dot notation for readability
- 0.0 - 0.65535 16-bit ASN block
- 2.0 - 2.1023 APNIC
- 3.0 - 3.1023 RIPE NCC
- 4.0 - 4.1023 LACNIC
- 5.0 - 5.1023 AfriNIC
- 6.0 - 6.1023 ARIN
- Remainder are reserved or held by IANA

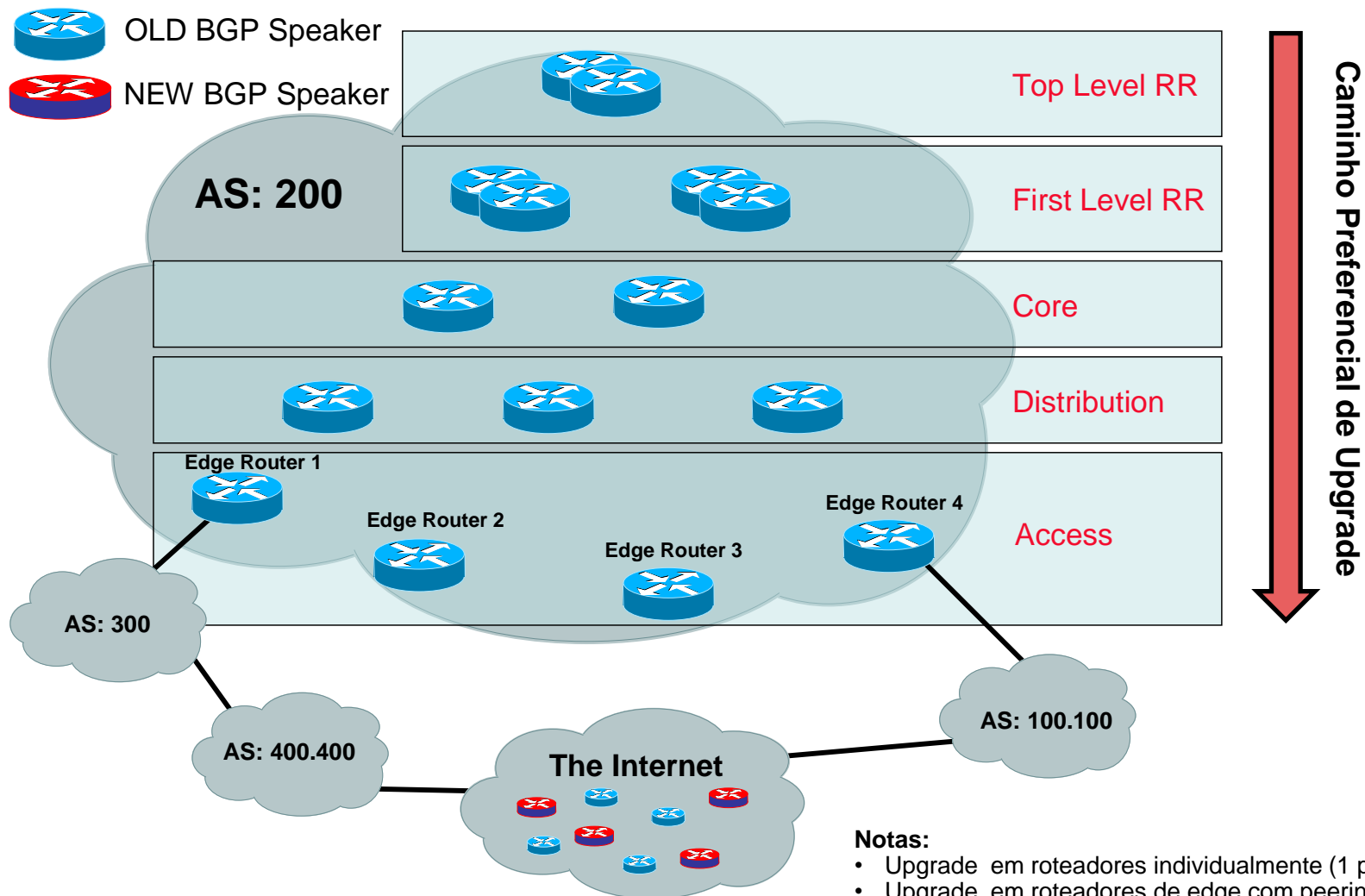
Changes (1)

- 32-bit ASNs are backwardly compatible with 16-bit ASNs
- There is no flag day
- You do NOT need to:
 - Throw out your old routers
 - Replace your 16-bit ASN with a 32-bit ASN

Changes (2)

- You do need to be aware that:
 - Your customers will come with 32-bit ASNs
 - ASN 23456 is not a bogon!
 - You will need a router supporting 32-bit ASNs to use a 32-bit ASN
- If you have a proper BGP implementation, 32-bit ASNs will be transported silently across your network

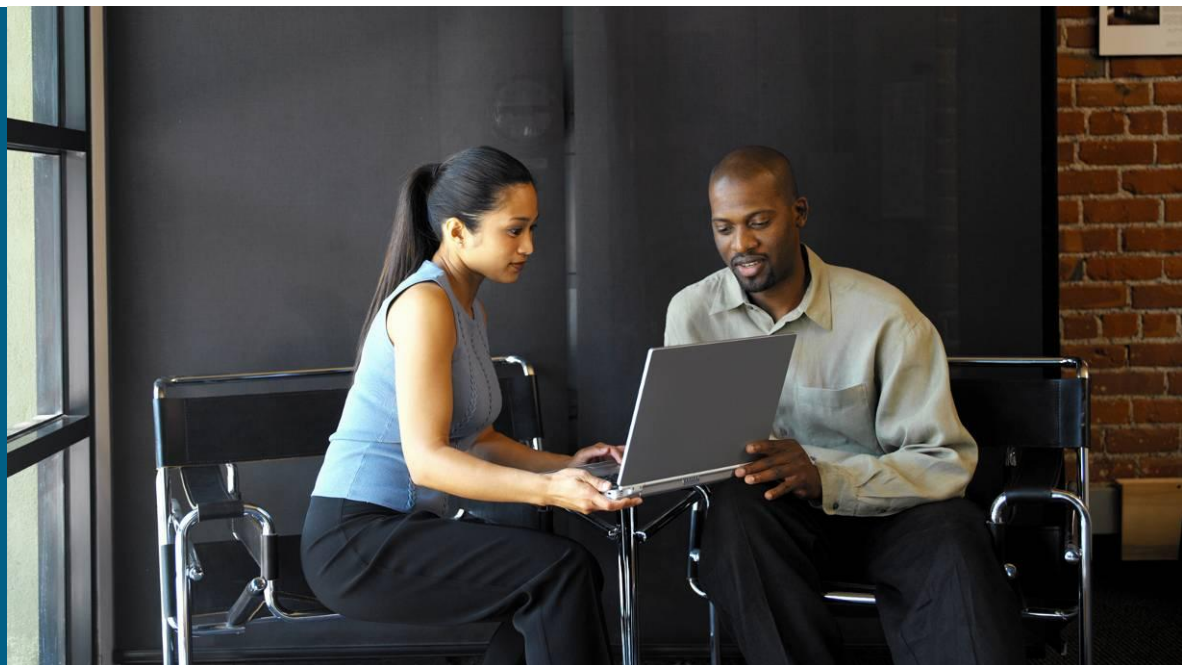
Arquitetura de Referência



Notas:

- Upgrade em roteadores individualmente (1 por vez)
- Upgrade em roteadores de edge com peering com NEW Speakers requer mudanças de configuração

Detalhamento Operacional



NEW Speaker– OPEN MESSAGE Handling

- Sending OPEN Message

To indicate its capability to support 4 byte AS number, a NEW speaker will embed a capability code 65 with value of AS being its own AS in the OPEN message to its BGP Peer.

Encode the “My AS field” in OPEN message with “23456” if NEW Speaker is configured with a non-mappable AS number.

Encode its OWN AS in My AS field in OPEN message if configured with mappable AS number

- Receipt of OPEN Message

Depending on the capability of the OPEN message received from the peer, NEW speaker will interpret if the peer is a NEW speaker or OLD speaker.

NEW Speaker– UPDATE MESSAGE Handling

- When Sending an Update message

Peering with NEW Speaker: 4 bytes AS_PATH and AS_AGGREGATOR in the Update message.

Peering with OLD Speaker,

If the AGGREGATOR/ASPATH does not contain a non-2-byte mappable 4-byte AS we are fine

If it does, substitute AS_TRANS (AS #23456) for each 4-byte AS
AS4_AGGREGATOR and/or AS4_ASPATH will contain a 4-byte encoded copy of the attribute if needed

OLD Speaker will blindly pass along AS4_AGGREGATOR and AS4_ASPATH attributes

NEW Speaker– UPDATE MESSAGE Handling (2)

- On Receipt of an Update message

Peering with NEW Speaker:

It will receive an update message with 4 byte AS_PATH and, eventually, AS_AGGREGATOR attributes

Peering with OLD Speaker:

AS4_PATH and AS_PATH must be merged to form the correct as-path

AS4_AGGREGATOR will override AGGREGATOR if AGGREGATOR is AS_TRANS

AGGREGATOR will override AS4_AGGREGATOR if AGGREGATOR is not AS_TRANS

- Merging AS4_PATH and AS_PATH

AS_PATH – 275 250 225 23456 23456 200 23456 175

AS4_PATH – 100.1 100.2 200 100.3 175

Merged as-path – 275 250 225 100.1 100.2 200 100.3 175

OLD Speaker– OPEN and UPDATE Message Handling

- Ignore the 4 byte capability received in the NEW speakers open message.
- When Sending an Update message, to a NEW speaker or an OLD speaker, it will construct an update message by pre-pending AS_PATH attribute with its OWN AS and will pass AS4_PATH and AS4_AGGREGATOR attributes as is.
- On Receipt of an Update message, it will construct an AS path information using the AS_PATH attributes.

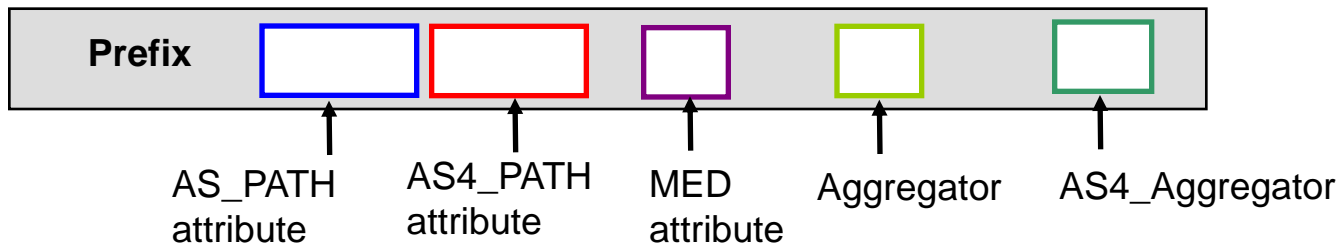
Convention Used



2 Byte capable Router



4 Byte capable Router



BGP Update Message



BGP OPEN Message



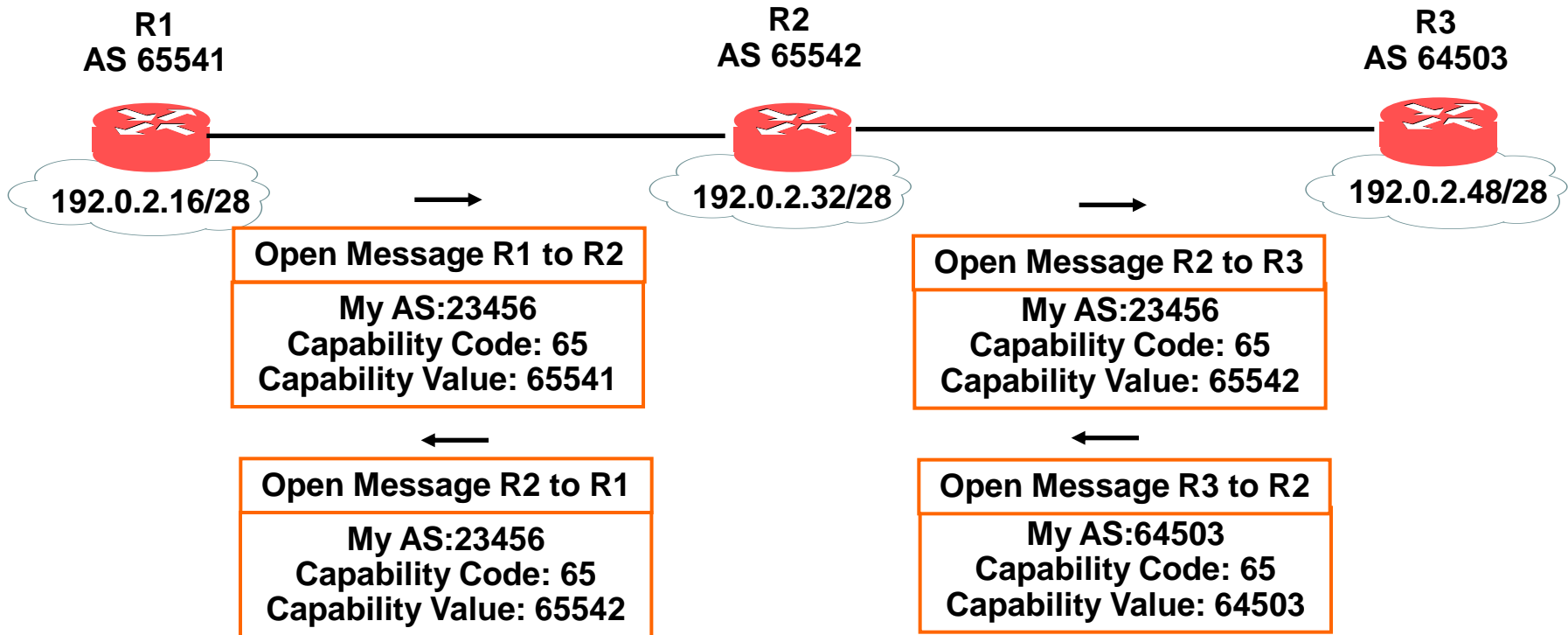
BGP Table



RIB

- IPv4 address range used – 192.0.2.0/24 as per RFC 3330
- 2 byte ASN range used - 64496 – 64511 as per RFC 5398
- 4 byte ASN range used - 65536 – 65551 as per RFC 5398

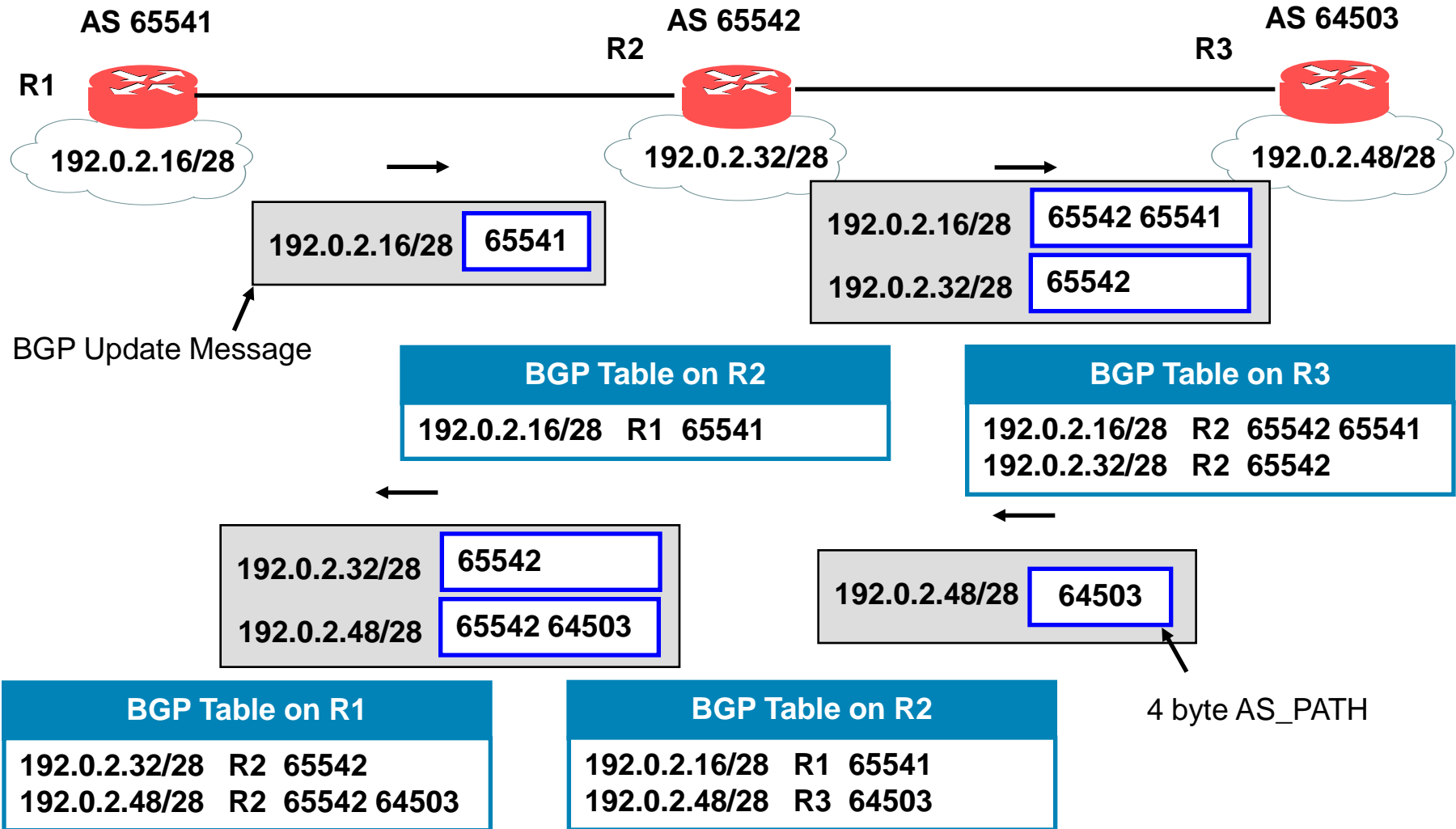
NEW Speaker – OPEN Message Exchange



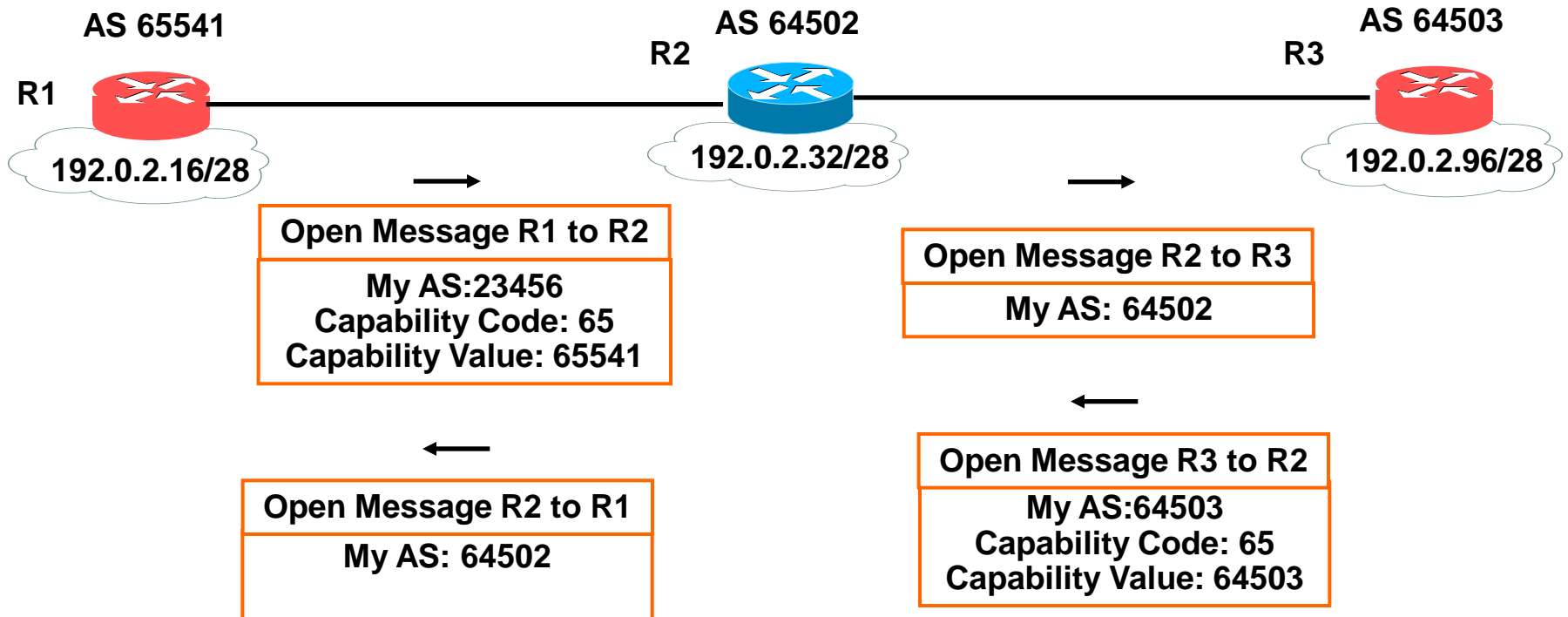
On R1

```
BGP: 192.0.2.2 sending OPEN, version 4, my as: 23456, holdtime 180 seconds
BGP: 192.0.2.2 rcv OPEN w/ OPTION parameter len: 24
BGP: 192.0.2.2 OPEN has CAPABILITY code: 65, length 4
BGP: 192.0.2.2 OPEN has 4-byte ASN CAP for: 1.5
BGP: 192.0.2.2 rcvd OPEN w/ remote AS 23456, 4-byte remote AS 65542
```

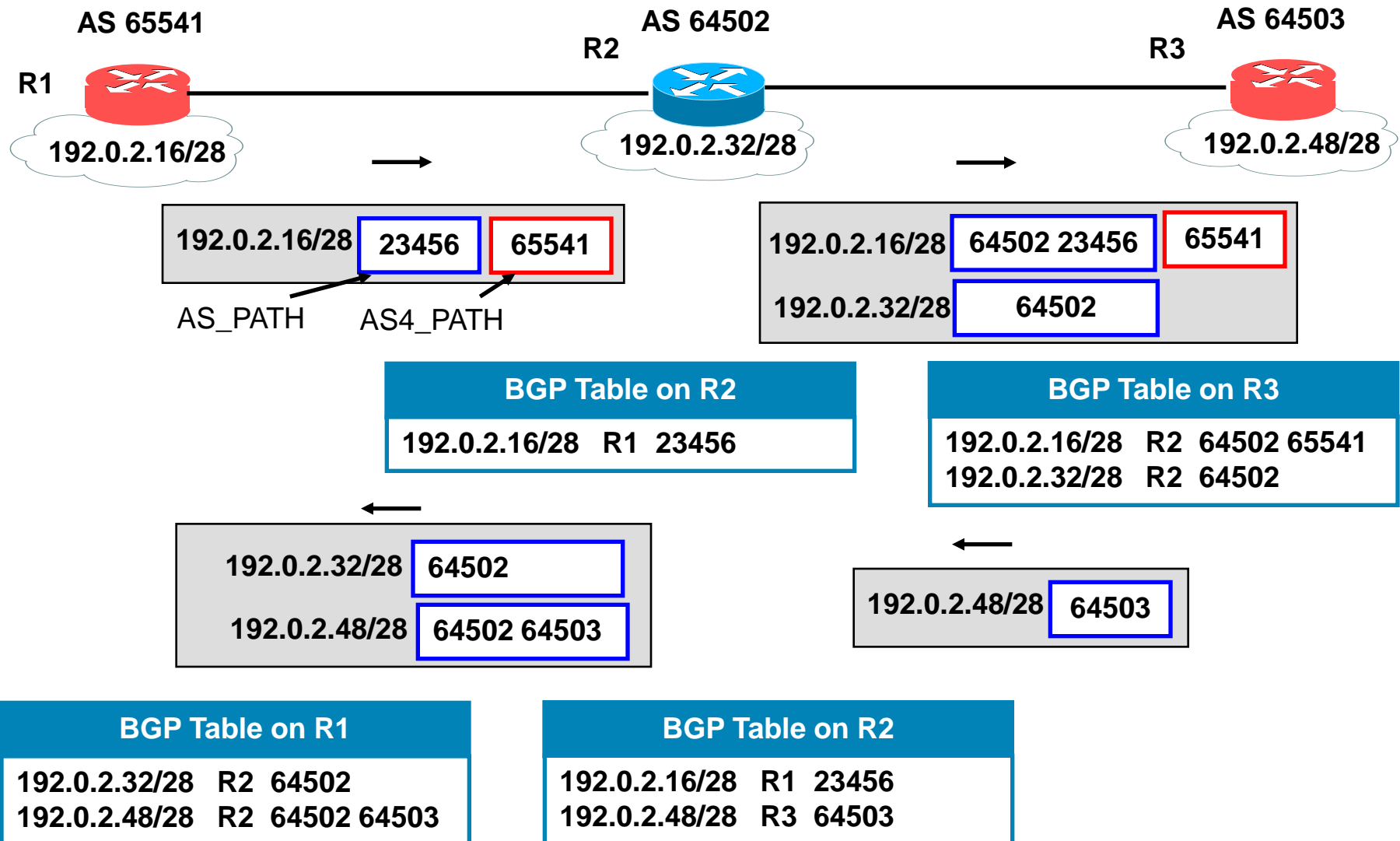
NEW Speaker – BGP Update Exchange



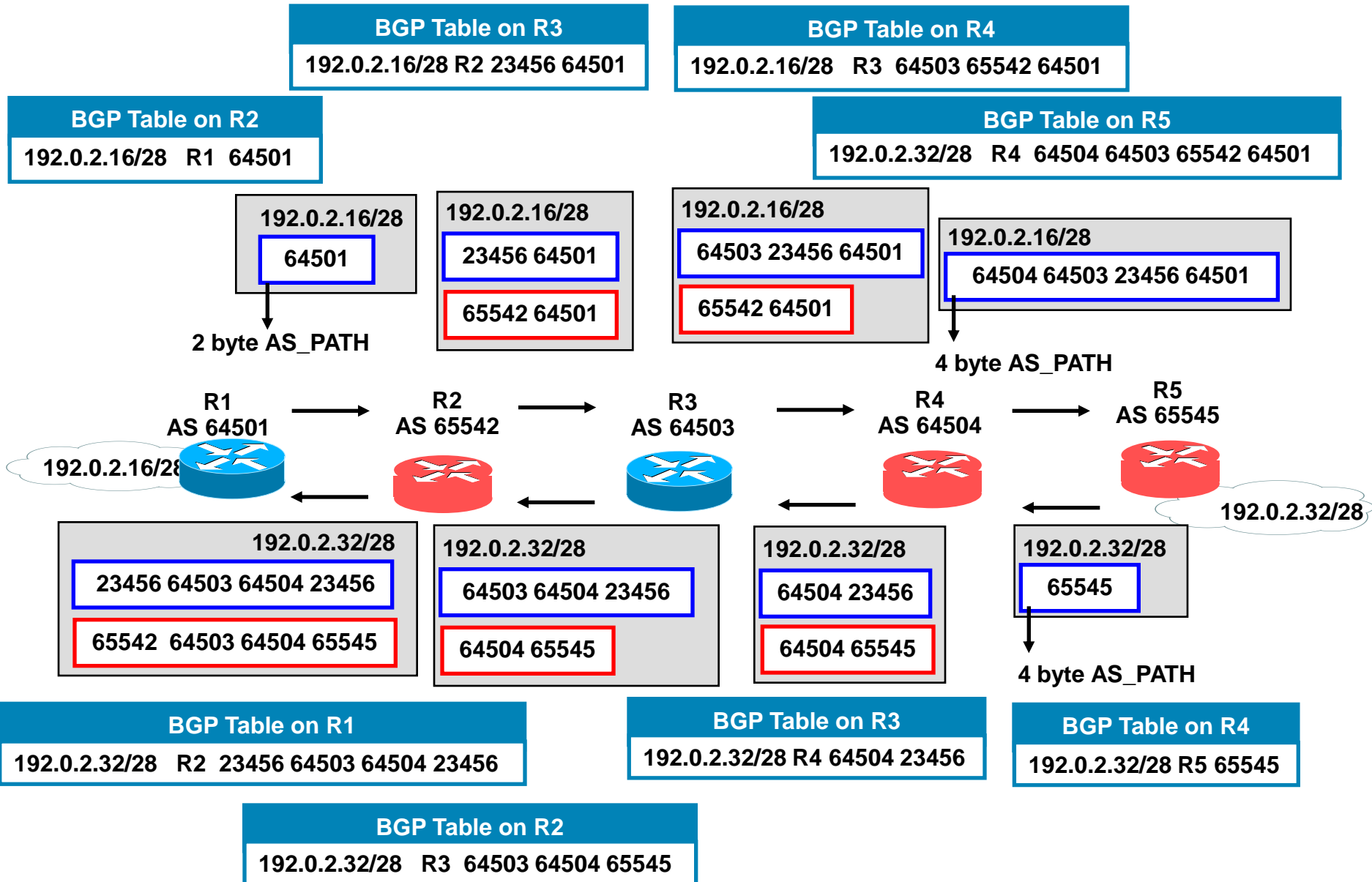
NEW and OLD Speaker – OPEN Message Exchange



OLD Speaker – BGP Update Exchange



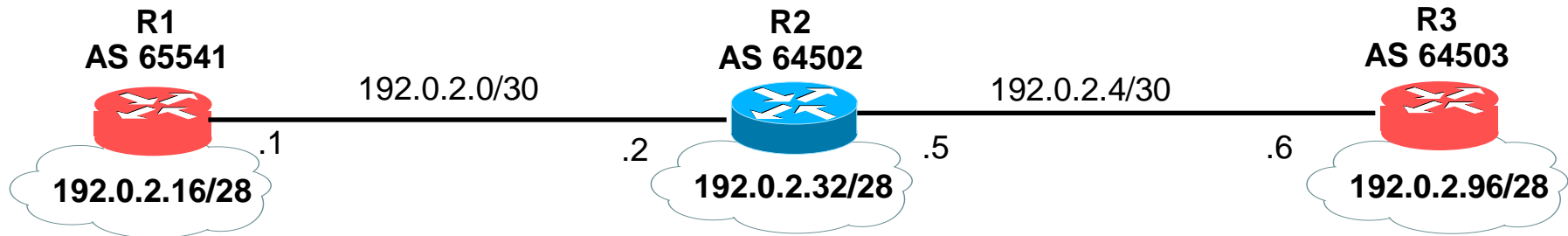
Construction of AS Path Information



NEW Speaker Configuration (asplain)



For your reference



```
hostname R1
!  
router bgp 65541  
no synchronization  
bgp router-id 192.0.2.11  
bgp log-neighbor-changes  
network 192.0.2.16 mask 255.255.255.240  
neighbor 192.0.2.2 remote-as 64502  
no auto-summary
```

```
hostname R3  
!  
router bgp 64503  
no synchronization  
bgp router-id 192.0.2.13  
bgp log-neighbor-changes  
network 192.0.2.96 mask 255.255.255.240  
neighbor 192.0.2.5 remote-as 64502  
no auto-summary
```

```
hostname R2  
!  
router bgp 64502  
no synchronization  
bgp router-id 192.0.2.12  
bgp log-neighbor-changes  
network 192.0.2.32 mask 255.255.255.240  
neighbor 192.0.2.1 remote-as 23456  
neighbor 192.0.2.6 remote-as 64503  
no auto-summary
```



For your
reference

Configuration (asdot)

```
router bgp 4.4
  bgp log-neighbor-changes
  neighbor 134.0.0.3 remote-as 3.3
```

Neighbor Configuration

R4#sh ip bgp 1.1.1.0

```
BGP routing table entry for 1.1.1.0/24, version 2
Paths: (1 available, best #1, table default)
Flag: 0x820
  Not advertised to any peer
```

3.3 2 1.1

```
  134.0.0.3 from 134.0.0.3 (134.0.0.3)
```

```
    Origin IGP, localpref 100, valid, external, best
```

R4#sh ip bgp sum

```
BGP router identifier 134.0.0.4, local AS number 4.4
```

```
BGP table version is 2, main routing table version 2
```

```
1 network entries using 124 bytes of memory
```

```
1 path entries using 52 bytes of memory
```

```
2/1 BGP path/bestpath attribute entries using 184 bytes of memory
```

```
1 BGP AS-PATH entries using 40 bytes of memory
```

```
0 BGP route-map cache entries using 0 bytes of memory
```

```
0 BGP filter-list cache entries using 0 bytes of memory
```

```
BGP using 400 total bytes of memory
```

```
BGP activity 1/0 prefixes, 1/0 paths, scan interval 60 secs
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
134.0.0.3	4	3.3	28	27	2	0	0	00:25:33	1

BGP Show Command



For your
reference

Configuration (asdot)

```
R3#sh ip rout | include B
```

Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP

```
B      2.2.2.0 [20/0] via 123.0.0.2, 00:11:01
```

```
B 192.0.0.0/24 [20/0] via 123.0.0.2, 00:11:01
```

```
R3#sh ip route 192.0.0.0
```

Routing entry for 192.0.0.0/24

Known via "**bgp 3.3**", distance 20, metric 0

Tag 2, type external

Redistributing via ospf 1

Advertised by ospf 1

Last update from 123.0.0.2 00:12:14 ago

Routing Descriptor Blocks:

* 123.0.0.2, from 123.0.0.2, 00:11:09 ago

Route metric is 0, traffic share count is 1

AS Hops 1

Route tag 2

Routing Show Commands

```
ip as-path access-list 1 permit ^1\\.4$  
router bgp 1  
neighbor 4.4.4.4 remote-as 1.4  
neighbor 4.4.4.4 route-map foo in
```

Note that the "." must be
escaped from the
regular expression with a "\

```
route-map foo permit 10  
match as-path 1
```

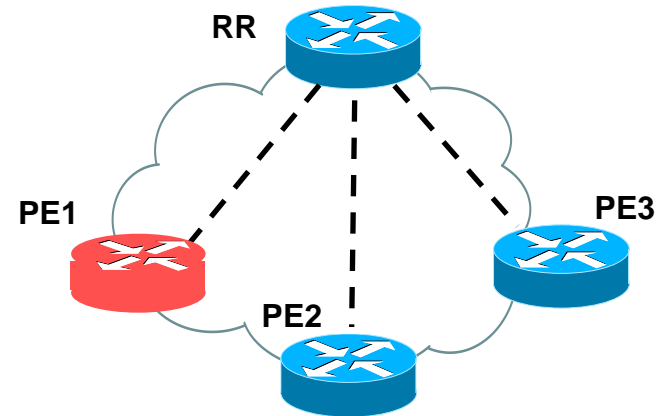
AS-PATH Filter for ASDOT notation

Impactos Operacionais com a Utilização de 4B ASN

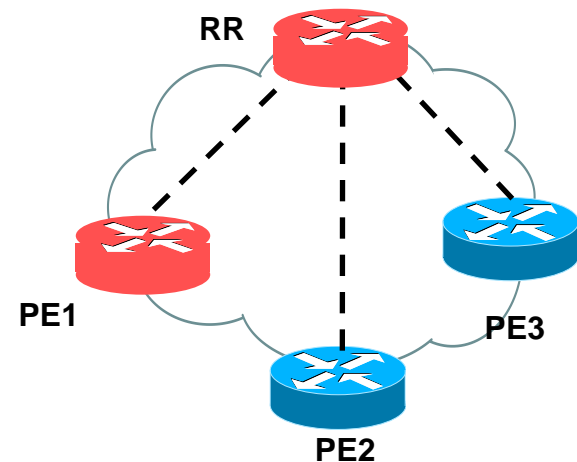


Update Groups

- Update group groups all the BGP peers with similar outbound policies
- Builds a single update message for all the grouped peer members in an update group and then replicate the formatted message to the tcp stream of all grouped peer members in an update group.
- Formatting a single update message using an update group allow BGP to optimize the CPU cycles and its transient memory usage.
- In a mixed environment with 4B Route Reflector, 4 byte and 2 byte capable PE's are grouped in different update-groups.



Update Group 1 – PE1, PE2, PE3



Update Group 1 – PE1
Update Group 2 – PE2, PE3

BGP Multipath

- Allows installation of multiple BGP path for the same destination into the IP routing table.
- Installed in the table together with the best path for load sharing.
- Need to have the same attributes as the best path to be a candidate for multipath.
- If AS-PATH length and AS-PATH content of candidate path and bestpath is the same, then the candidate path is considered as a multipath.
- In a mixed provider network (with RR) with NEW and OLD speaker and multipath enabled, an ingress PE will utilize one of the paths as primary and other as backup/secondary except in the case of add-path capability (currently a draft at IETF)
- In a mixed VPN provider network with unique RD's with NEW and OLD speaker and multipath enabled, an ingress PE will load balance traffic instead of utilizing one of the paths as primary and secondary path as backup.

BGP Selection Process Refresher

1. Highest WEIGHT (Cisco specific)
2. Highest LOCAL_PREF
3. Locally originated via a network or aggregate BGP subcommand or through redistribution from an IGP.
4. Shortest AS_PATH.
5. Lowest origin type (IGP < EGP < INCOMPLETE)
6. Lowest multi-exit discriminator (MED).
7. eBGP over iBGP paths.
8. Lowest IGP metric to the BGP next hop (continue, even if bestpath is already selected)
9. Determine if multiple paths require installation in the routing table for BGP Multipath.
10. When both paths are external, prefer the path that was received first (the oldest one).
11. Prefer the route that comes from the BGP router with the lowest router ID.
12. If the originator or router ID is the same for multiple paths, prefer the path with the minimum cluster list length (RR environments)
13. Prefer the path that comes from the lowest neighbor address.

BGP Multipath Refresher

- Candidates must have equal characteristics:

Weight

Local preference

AS-PATH length

Origin

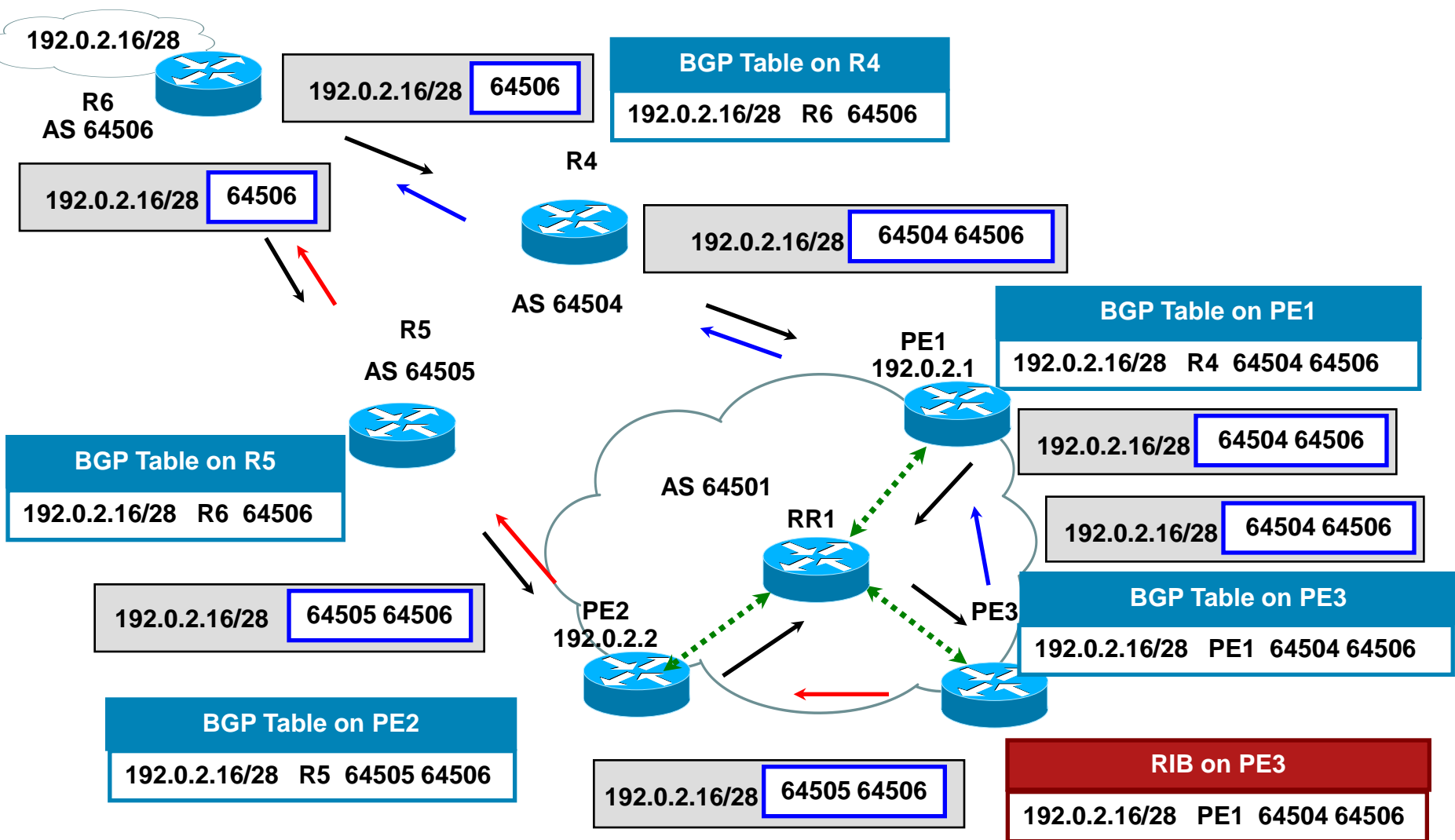
MED

One of these:

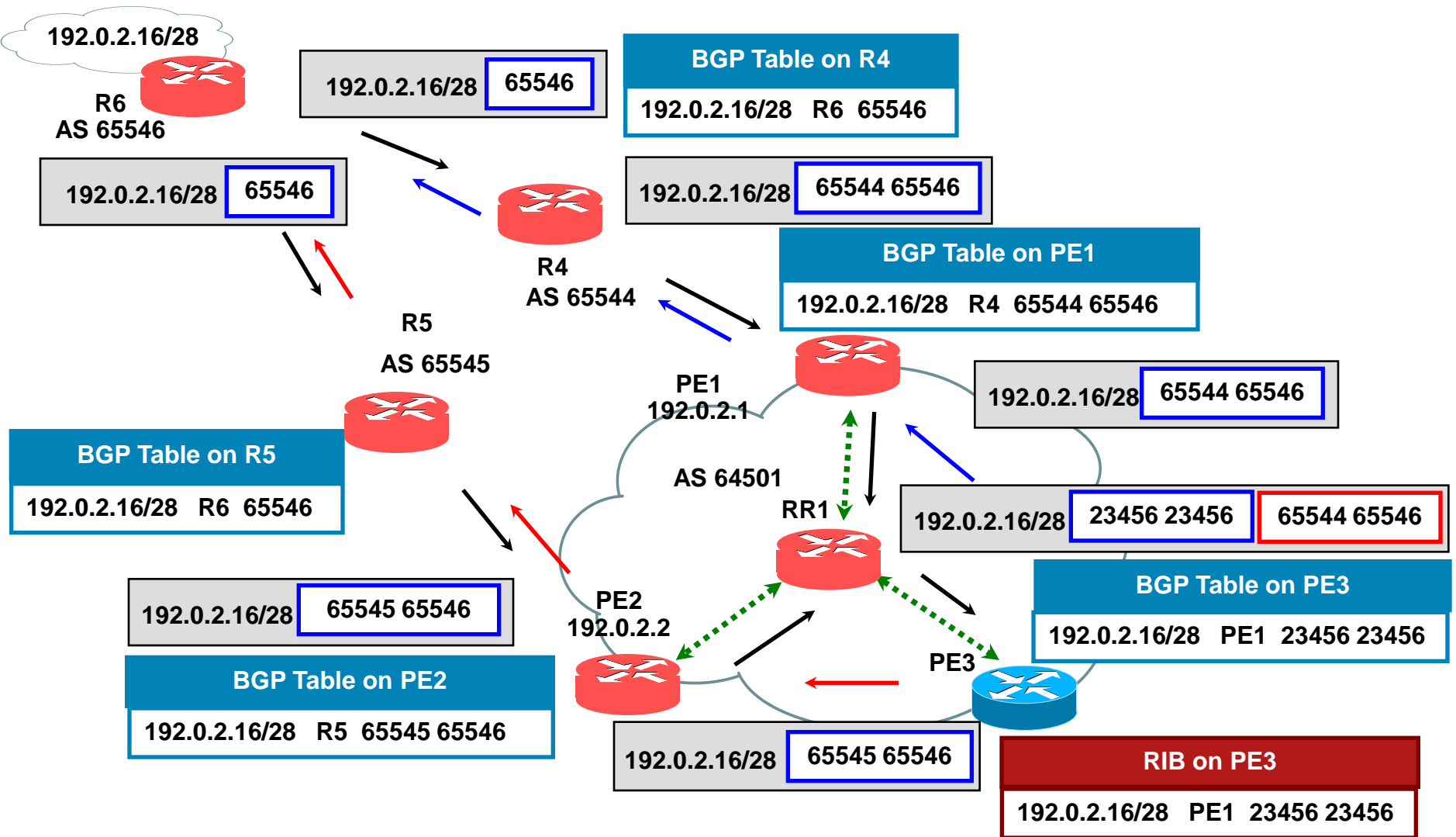
Neighboring AS or sub-AS (before eiBGP Multipath feature)

AS-PATH (after eiBGP Multipath feature)

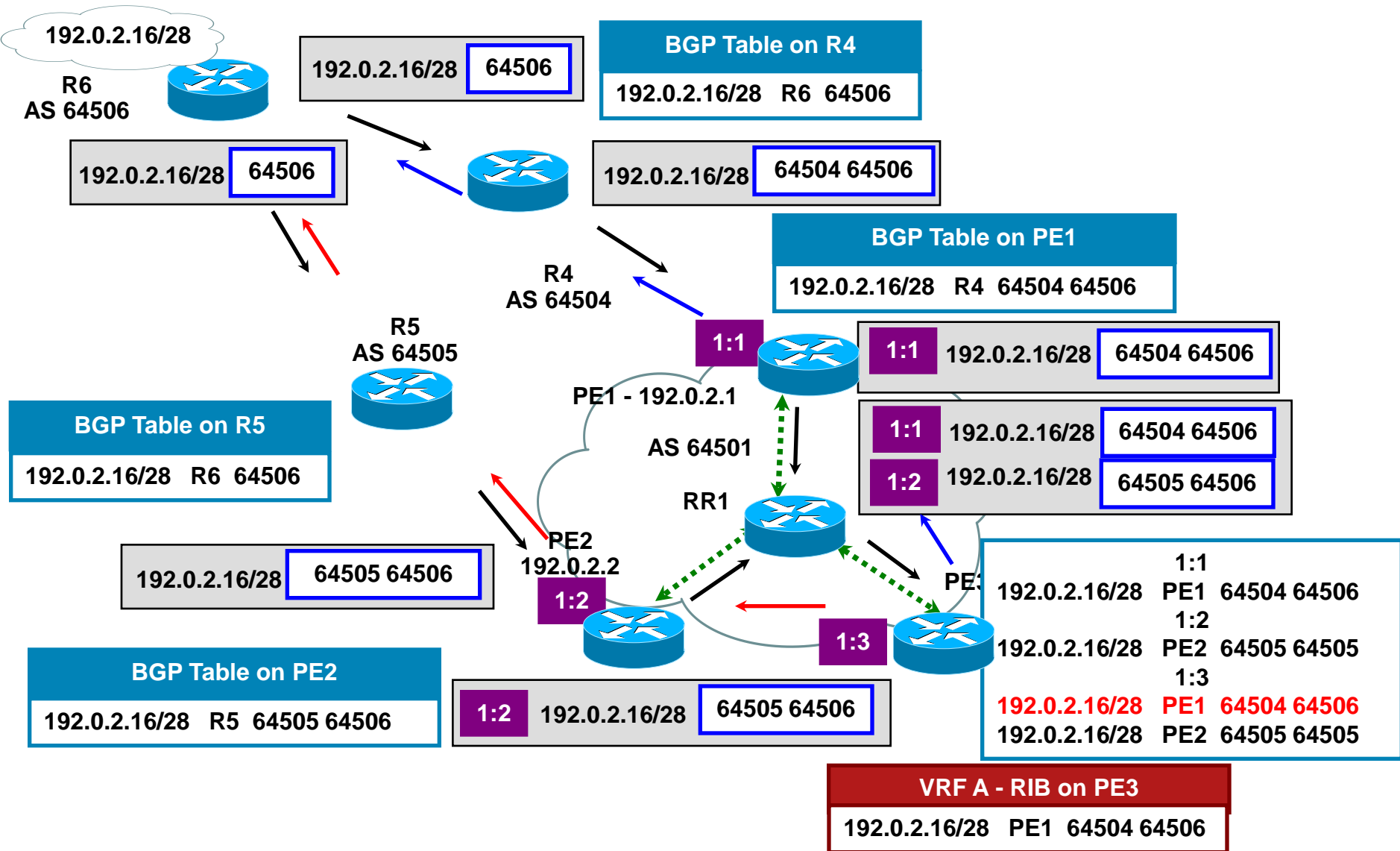
BGP Multipath - OLD Speaker Network



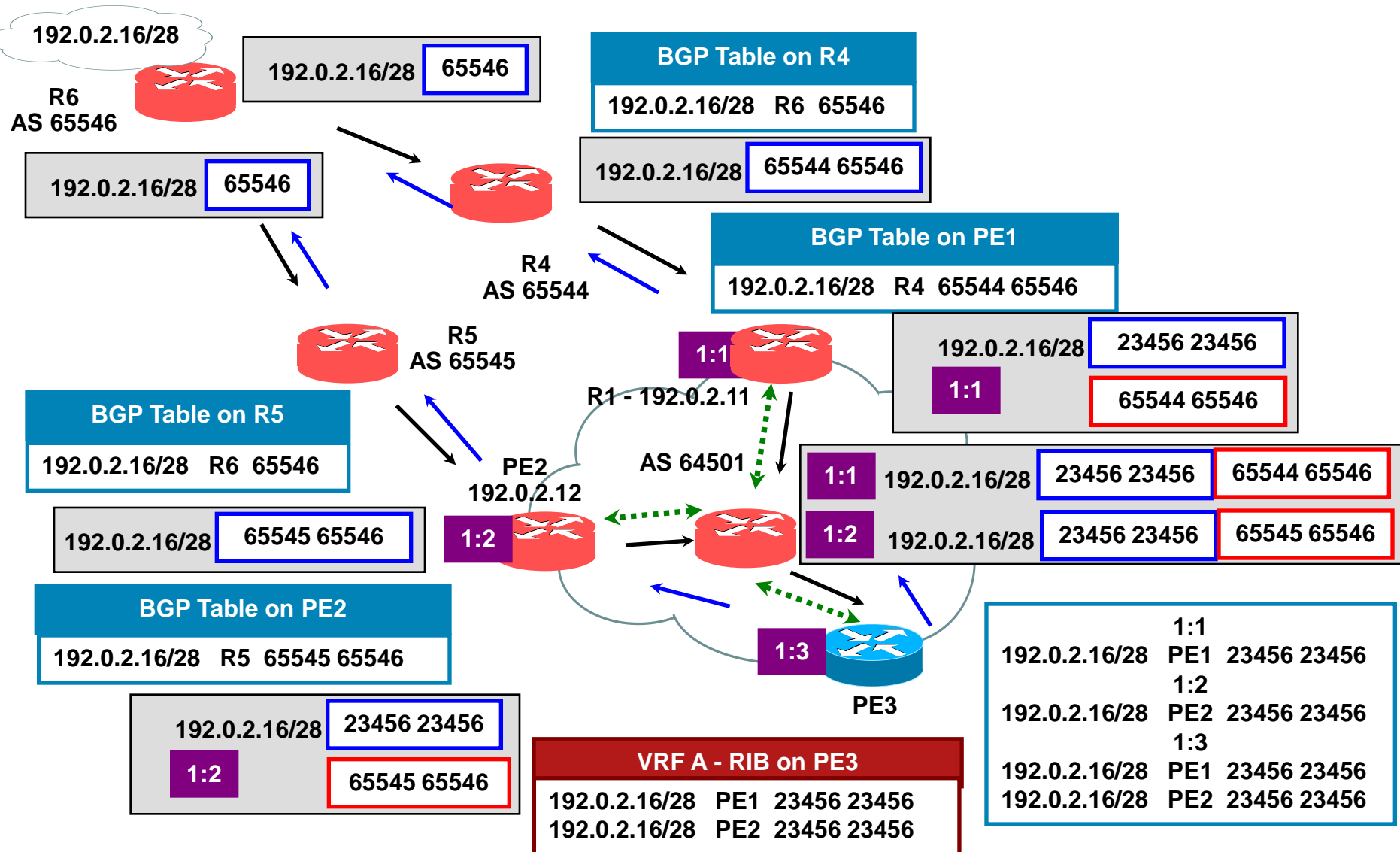
BGP Multipath - Mixed Network NEW & OLD Speaker



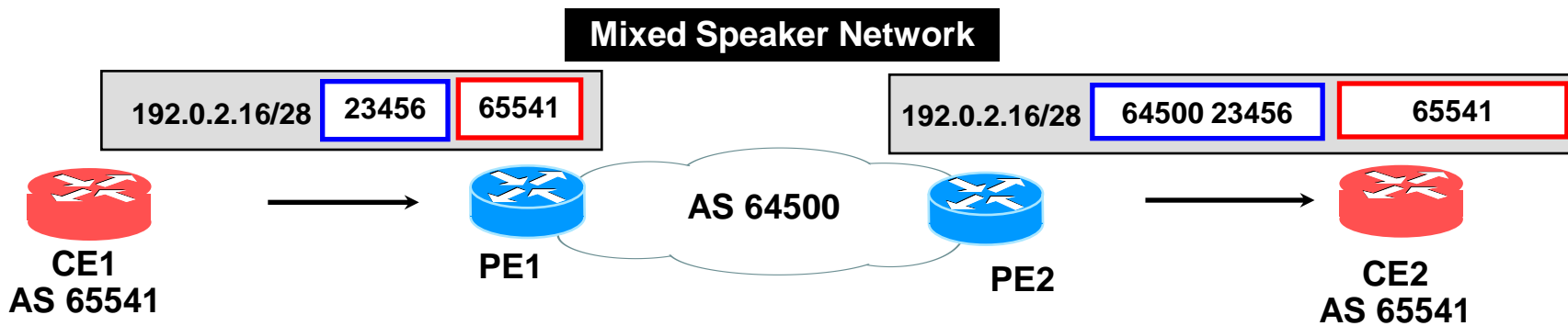
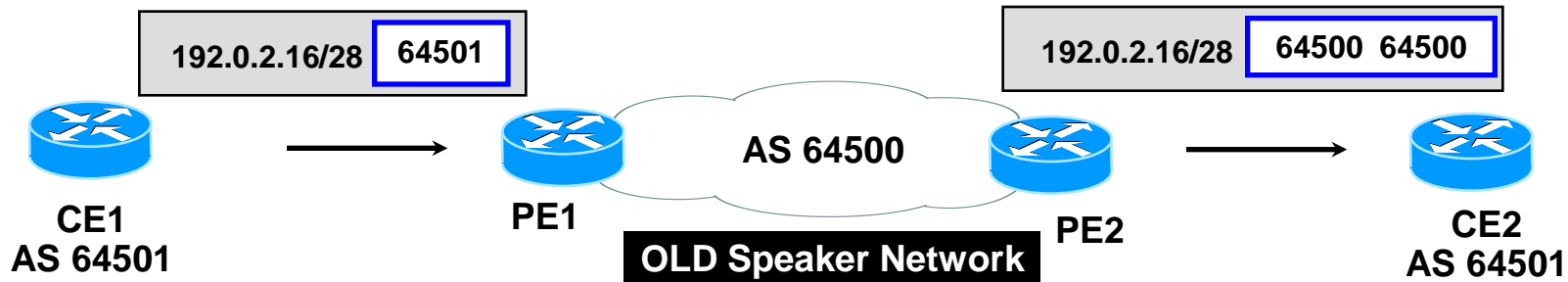
BGP Multipath - OLD Speaker VPN (unique RDs)



BGP Multipath - Mixed NEW & OLD Speaker VPN (unique RDs)



AS-Override



CE2 construct AS PATH information from the AS_PATH and AS4_PATH received from PE2.
 CE2 however drops the UPDATE as it finds 65541 in the path (loop prevention).
 AS Override function fails in an Old Speaker to replace ASN in AS4_PATH, as it doesn't recognize it.

Key Takeaways and Guidelines

- 4 byte ASN influences AS-PATH handling. Examine areas or services within your network that could be potentially influenced.
- Reserved AS Number - AS_TRANS i.e 23456 is a number reserved by IANA. The number is used by a NEW speaker in the AS-PATH to map non-mappable 4 octet AS numbers. This AS number is also used to configure a BGP peer on an OLD speaker. Misconfiguring a 4 byte peer connected to a OLD speaker could still result in the BGP session being formed
- Increase in Memory and CPU Processing – Due to additional attributes AS4_PATH and AS4_AGGREGATOR, an increase in memory usage as well as CPU processing cycle may be observed.
- Unable to construct AS path information - AS4_PATH attribute would be lost during route aggregation, or both the AS_PATH attribute and the AS4_PATH attribute would contain valid, partial information that cannot be combined seamlessly, resulting in incomplete AS path information in these cases
- Netflow – Due to usage of AS_TRANS, Netflow may not be able to provide accurate data as prefixes from 4 byte ASN will all be summarized under AS 23456.

References (1)

- Q.Vohra and E.Chen, “BGP Support for Four-octet AS Number Space” RFC 4893, May 2007
- “Four-Octet AS Specific BGP Extended Community.” Internet Draft [draft-ietf-l3vpn-as4octet-ext-community-02.txt](#), work in progress, May 2009
- G. Huston and G. Michaelson, “Textual Representation of Autonomous System (AS) Numbers” RFC 5396, December 2008
- “The 32-bit AS Number Report.” <http://www.potaroo.net/tools/asn32/>
- Y.Rekter and T.Li , “ A Border Gateway Protocol (BGP-4)” RFC 1771, March 1995
- S. Sangli, D. Tappan and Y. Rekter, “BGP Extended Communities Attribute” RFC 4360, February 2006.
- “IBGP Multipath”.
http://www.cisco.com/en/US/docs/ios/12_2t/12_2t2/feature/guide/ftbgpls.html
- AS4 Wiki. http://as4.cluepon.net/index.php/Main_Page
- APNIC Wiki.
<http://wiki.icons.apnic.net/display/ASN/Operational+Resources>

References (2)

- Migration Guide for Explaining 4-Byte Autonomous System.
http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6554/ps6599/datasheet_c78_516825.html
- Explaining 4-octet Autonomous System (AS) Numbers for Cisco IOS.
http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6554/ps6599/white_paper_C11_516823.html
- Network Considerations for a Network Which Does Not Support 4-octet AS.
http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6554/ps6599/white_paper_c11-516826.html
- Explaining 4-Byte Autonomous System (AS) ASPLAIN and ASDOT Notation for Cisco IOS.
http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6554/ps6599/white_paper_c11_516829.html

