

AOS-CX 10.6 Update
November, 2020



BGP ADD-PATH

Aruba Switching TME



ADD-PATH

Agenda



1

Overview

2

Use Cases

3

Details / Caveats

4

Configuration

5

Best Practices

6

Troubleshooting

7

Demo

Overview

aruba

a Hewlett Packard
Enterprise company

Definitions

Acronyms

- MP-BGP **M**ulti-**P**rotocol **B**order **G**ateway **P**rotocol
- AF **A**ddress **F**amily (Ex: IPv4, IPv6 or EVPN address families used in MP-BGP)
- EVPN **E**thernet **V**irtual **P**rivate **N**etwork
- L2VPN **L**ayer2 **V**irtual **P**rivate **N**etwork
- MB-BGP EVPN Refers to the EVPN address family in MP-BGP
- ADD-PATH BGP ADD-PATH capability
- NLRI Network Layer Reachability Information

Overview

Reminder

- BGP, by default, does not provide multiple paths for a given prefix.
- A Route-Reflector will only advertise the **best path** as determined by the BGP route selection process. It will never advertise more than **one path** to a RR client irrespective of whether the RR has one or more paths for the given prefix.
- The BGP path advertisement of a prefix with new attributes replaces the previous announcement of that prefix. This behavior is known as an implicit withdraw, which effectively **prevents the advertisement of multiple paths for the same prefix**.

BGP route selection reminder

Part 1 : commonly used

Step	BGP attribute / criteria	Attribute type	Comment
1	next hop reachability		Next-hop in routing table (ex: in OSPF)
2	highest weight	proprietary	local to router
3	highest local-preference	discretionary	Globally defined within the AS. Default LP=100
4	router originated		route locally originated by itself (like redistribution or route leaking)
5	shortest AS-PATH length	Mandatory	• Can be skipped with bgp bestpath as-path ignore. • Selection can be stopped with bgp bestpath as-path multipath-relax
6	lowest origin type	Mandatory	IGP < EGP < incomplete
7	lowest Multi-Exit-Discriminator	Optional non-transitive	MEDs are compared if routes came from the same remote AS or if bgp always-compare-med is enabled
8	eBGP preferred over iBGP or confederation		confederation paths are treated as iBGP
9	lowest IGP cost to the BGP next-hop		closest IGP neighbor

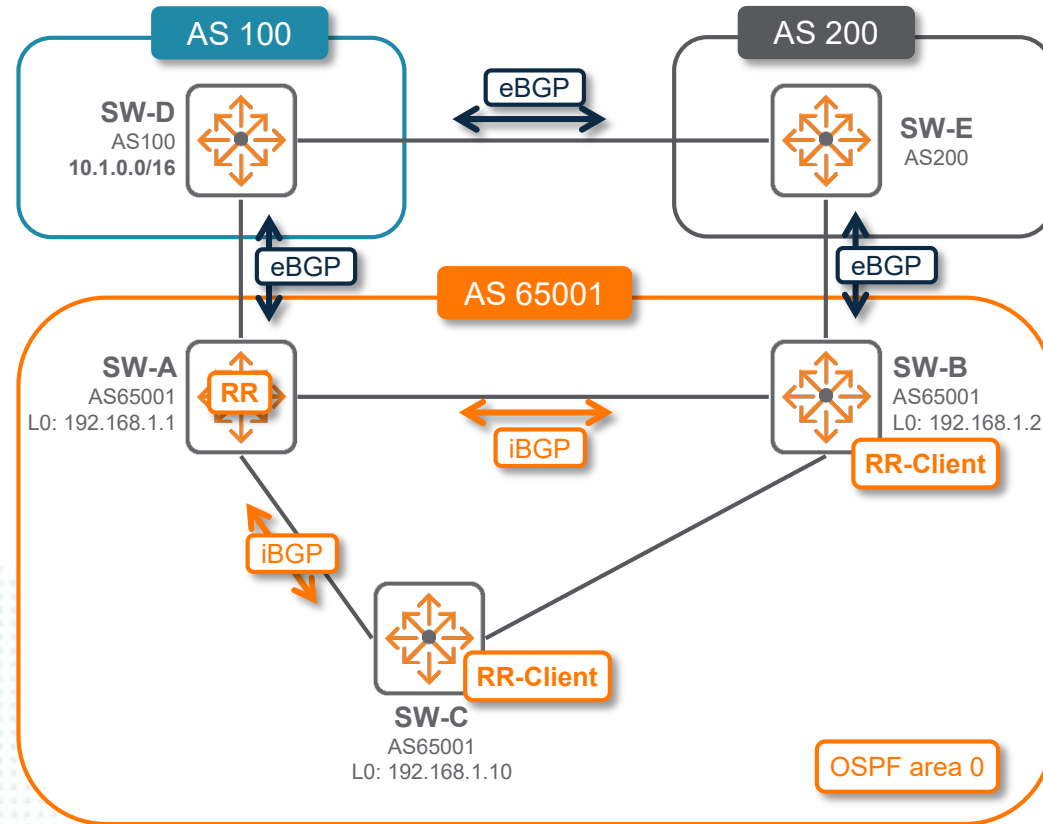
BGP route selection reminder

Part 2 : routing design should avoid usage of these criterias

Step	BGP attribute / criteria	Attribute type	Comment
10	Oldest eBGP route	N/A	most stable path
11	lowest router ID OR lowest originator ID (in case of RR)	Mandatory	
12	lowest cluster list length	Optional non-transitive	
13	lowest BGP peer IP address		IPv6 is preferred over IPv4

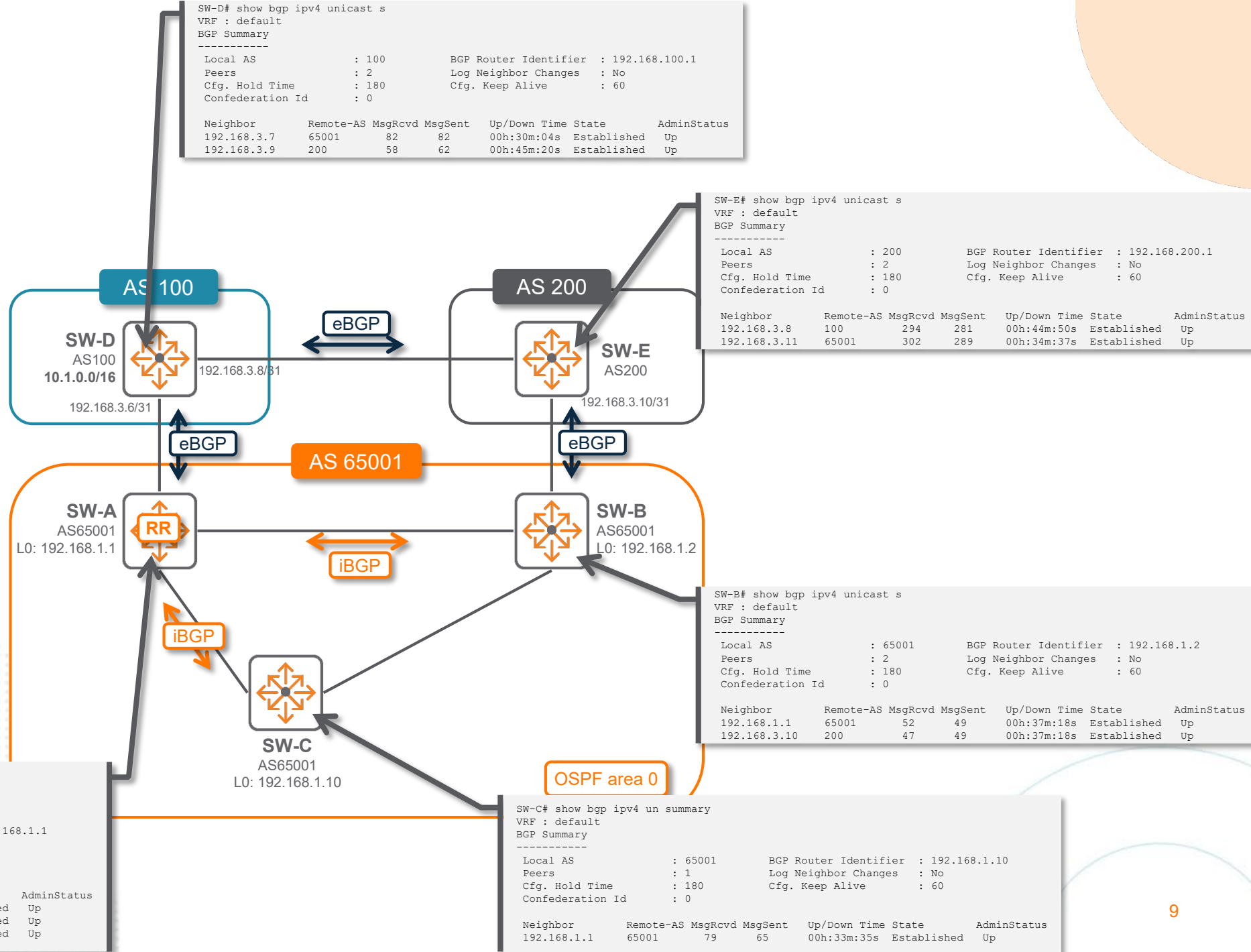
Example of traditional route selection

Topology with RR



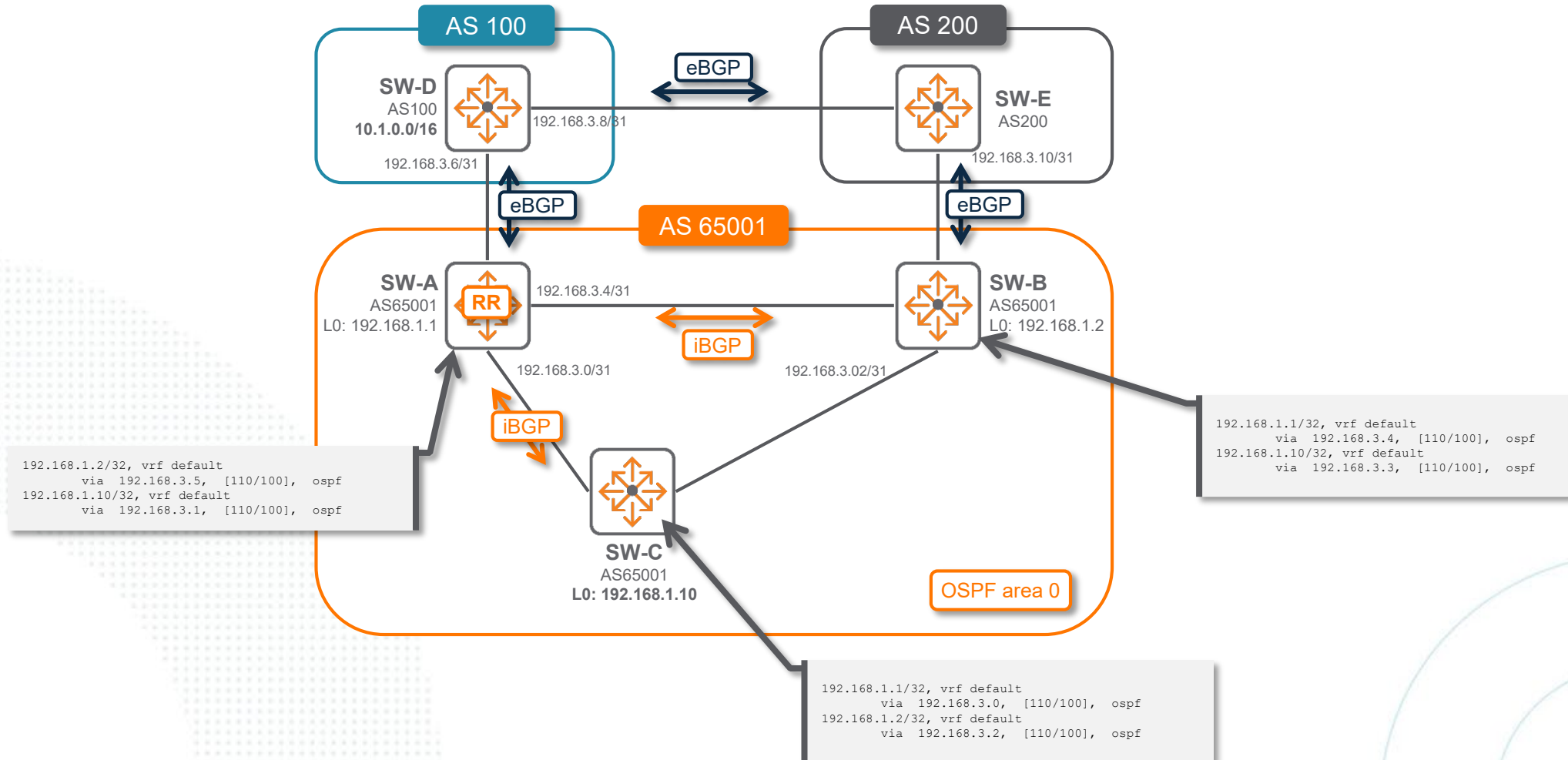
Example

BGP peering



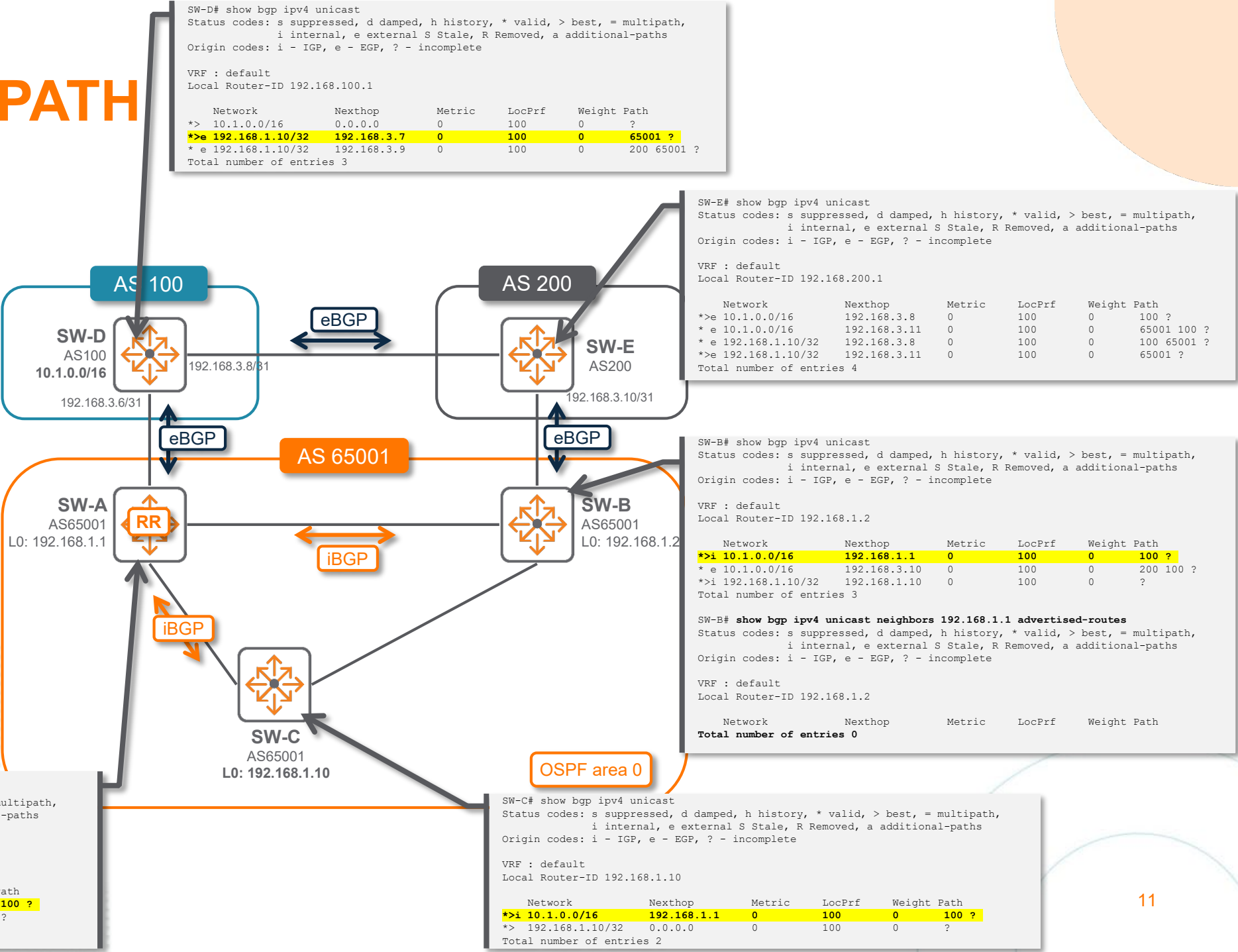
Example

iBGP next-hop learnt from OSPF



Without ADD-PATH

BGP RIB



Detailed Routes Exchange

BGP UPDATE messages

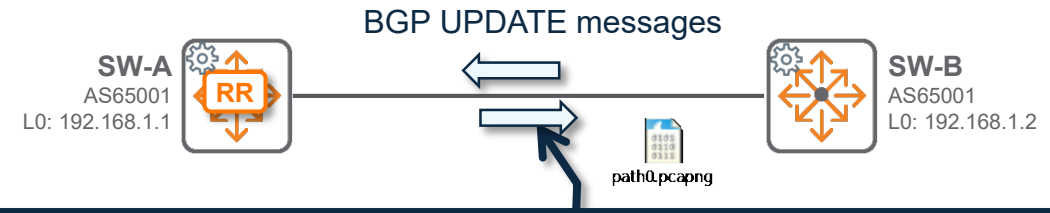
1. SW-B to SW-A: empty update message
2. SW-A to SW-B: empty update message and NLRI update for 192.168.1.10/32
3. SW-B to SW-A: NLRI update for 10.1.0.0/16

```
Border Gateway Protocol - UPDATE Message
Marker: ffffffffffffffffffffffffffff
Length: 57
Type: UPDATE Message (2)
Withdrawn Routes Length: 0
Total Path Attribute Length: 31
Path attributes
  Path Attribute - ORIGIN: INCOMPLETE
  Path Attribute - AS_PATH: 200 100
  Path Attribute - NEXT_HOP: 192.168.1.2
  Path Attribute - LOCAL_PREF: 100
Network Layer Reachability Information (NLRI)
  10.1.0.0/16
    NLRI prefix length: 16
    NLRI prefix: 10.1.0.0
```

4. SW-A to SW-B: NLRI update for 10.1.0.0/16

```
Border Gateway Protocol - UPDATE Message
Marker: ffffffffffffffffffffffffffff
Length: 53
Type: UPDATE Message (2)
Withdrawn Routes Length: 0
Total Path Attribute Length: 27
Path attributes
  Path Attribute - ORIGIN: INCOMPLETE
  Path Attribute - AS_PATH: 100
  Path Attribute - NEXT_HOP: 192.168.1.1
  Path Attribute - LOCAL_PREF: 100
Network Layer Reachability Information (NLRI)
  10.1.0.0/16
    NLRI prefix length: 16
    NLRI prefix: 10.1.0.0
```

5. SW-B to SW-A: Withdraw for 10.1.0.0/16

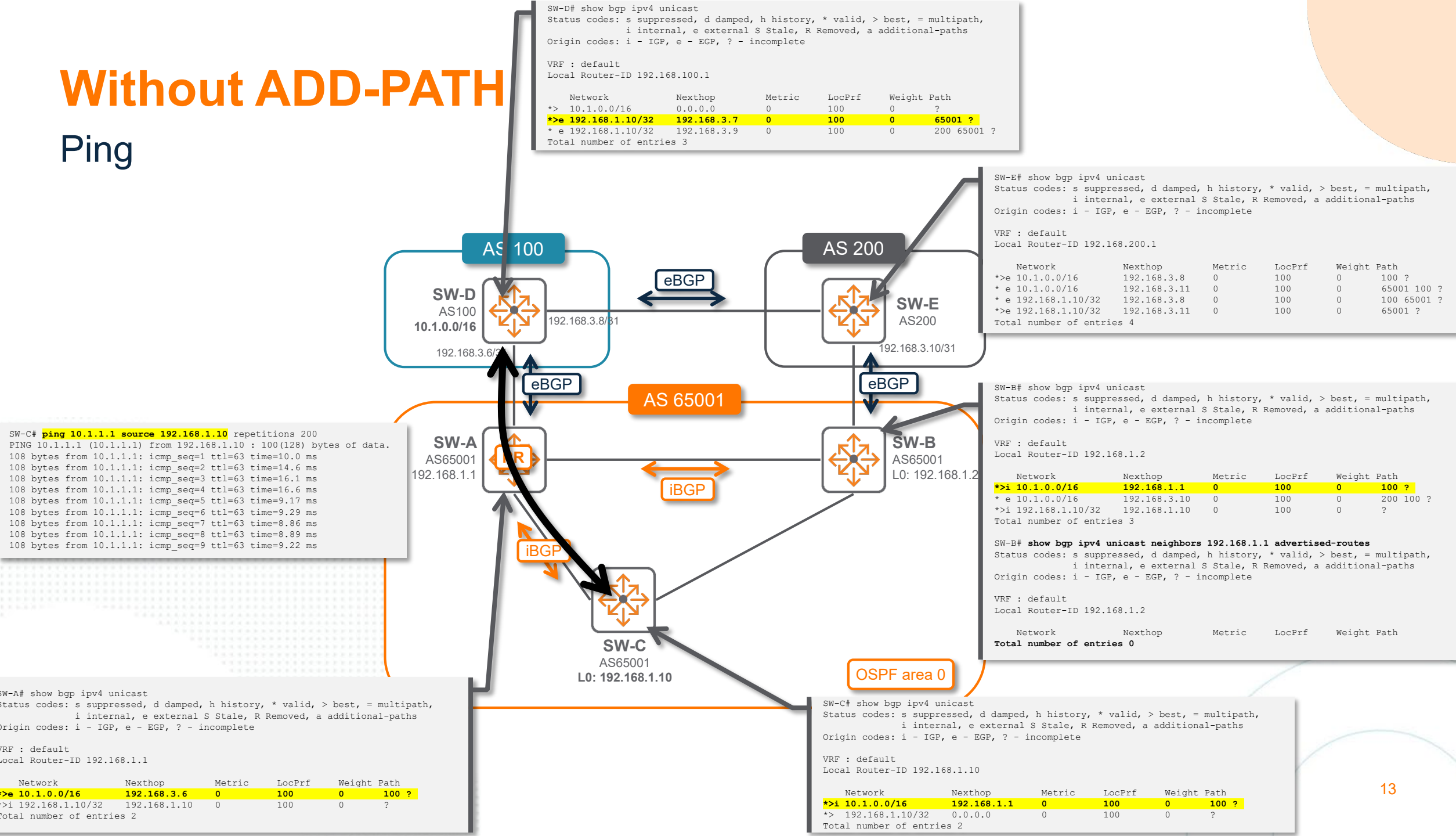


No.	Time	Source	Destination	Protocol	Length	Info
6	13.037144	192.168.1.2	192.168.1.1	TCP	66	36267 → 179 [ACK] Seq=1 Ack=29 Win=30016 Len=0 TSval=28
7	13.054139	192.168.1.1	192.168.1.2	TCP	66	179 → 36267 [FIN, ACK] Seq=29 Ack=1 Win=227 Len=0 TSval=
8	13.062466	192.168.1.1	192.168.1.2	TCP	74	44627 → 179 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_F
9	13.065171	192.168.1.2	192.168.1.1	TCP	74	179 → 44627 [SYN, ACK] Seq=0 Ack=1 Win=28960 Len=0 MSS=
10	13.065828	192.168.1.2	192.168.1.1	TCP	66	36267 → 179 [FIN, ACK] Seq=1 Ack=30 Win=30016 Len=0 TSv
11	13.070395	192.168.1.1	192.168.1.2	TCP	66	44627 → 179 [ACK] Seq=1 Ack=1 Win=29200 Len=0 TSval=101
12	13.071137	192.168.1.1	192.168.1.2	TCP	66	179 → 36267 [ACK] Seq=30 Ack=2 Win=227 Len=0 TSval=101
13	13.075868	192.168.1.2	192.168.1.1	BGP	123	OPEN Message
14	13.081154	192.168.1.1	192.168.1.2	TCP	66	44627 → 179 [ACK] Seq=1 Ack=58 Win=29200 Len=0 TSval=16
15	13.095123	192.168.1.1	192.168.1.2	BGP	123	OPEN Message
16	13.098416	192.168.1.2	192.168.1.1	TCP	66	179 → 44627 [ACK] Seq=58 Ack=58 Win=29056 Len=0 TSval=
17	13.100252	192.168.1.2	192.168.1.1	BGP	85	KEEPALIVE Message
18	13.158036	192.168.1.1	192.168.1.2	BGP	85	KEEPALIVE Message
19	13.158310	192.168.1.1	192.168.1.2	TCP	66	44627 → 179 [ACK] Seq=77 Ack=77 Win=29200 Len=0 TSval=
20	13.164820	192.168.1.2	192.168.1.1	BGP	89	UPDATE Message
21	13.167040	192.168.1.1	192.168.1.2	BGP	152	UPDATE Message, UPDATE Message
22	13.177629	192.168.1.2	192.168.1.1	BGP	123	UPDATE Message
23	13.220317	192.168.1.1	192.168.1.2	TCP	66	44627 → 179 [ACK] Seq=163 Ack=157 Win=29200 Len=0 TSval=
24	13.226789	192.168.1.1	192.168.1.2	BGP	119	UPDATE Message
25	13.269295	192.168.1.2	192.168.1.1	TCP	66	179 → 44627 [ACK] Seq=157 Ack=216 Win=29056 Len=0 TSval=
26	13.786503	192.168.3.4	224.0.0.5	OSPF	82	Hello Packet
27	17.673038	192.168.1.2	192.168.1.1	BGP	92	UPDATE Message
28	17.678748	192.168.1.1	192.168.1.2	TCP	66	44627 → 179 [ACK] Seq=216 Ack=183 Win=29200 Len=0 TSval=
29	19.611001	HewlettP_f9:...	LLDP_Multicast	LLDP	134	MA/08:00:09:f9:41:3c IN/1/1/1 120 SysN=SW-B SysD=Aruba
30	20.001751	192.168.3.5	224.0.0.5	OSPF	82	Hello Packet
31	23.786286	192.168.3.4	224.0.0.5	OSPF	82	Hello Packet

```
> Frame 27: 92 bytes on wire (736 bits), 92 bytes captured (736 bits) on interface -, id 0
> Ethernet II, Src: HewlettP_f9:41:3c (08:00:09:f9:41:3c), Dst: HewlettP_df:10:5f (08:00:09:df:10:5f)
> Internet Protocol Version 4, Src: 192.168.1.2, Dst: 192.168.1.1
> Transmission Control Protocol, Src Port: 179, Dst Port: 44627, Seq: 157, Ack: 216, Len: 26
v Border Gateway Protocol - UPDATE Message
  Marker: ffffffffffffffffffffffffffff
  Length: 26
  Type: UPDATE Message (2)
  Withdrawn Routes Length: 3
  v Withdrawn Routes
    v 10.1.0.0/16
      Withdrawn route prefix length: 16
      Withdrawn prefix: 10.1.0.0
      Total Path Attribute Length: 0
```

Without ADD-PATH

Ping



Without ADD-PATH

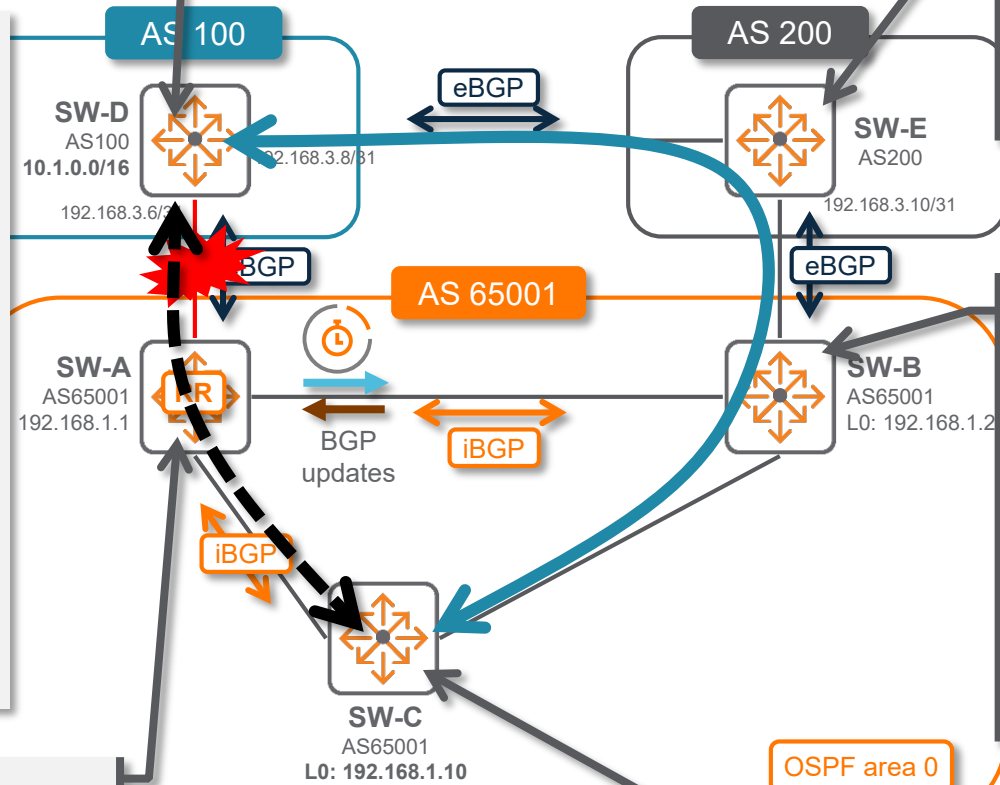
Convergence time
on link failure >3s

```
SW-D# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.100.1

   Network        Nexthop        Metric      LocPrf      Weight Path
*>e 10.1.0.0/16    0.0.0.0            0           100         0       ?
*>e 192.168.1.10/32 192.168.3.9        0           100         0       200 65001 ?
Total number of entries 2
```

```
SW-C# ping 10.1.1.1 source 192.168.1.10 repetitions 200
PING 10.1.1.1 (10.1.1.1) from 192.168.1.10 : 100(128) bytes of data.
 08 bytes from 10.1.1.1: icmp_seq=1 ttl=63 time=10.0 ms
 08 bytes from 10.1.1.1: icmp_seq=2 ttl=63 time=14.6 ms
 08 bytes from 10.1.1.1: icmp_seq=3 ttl=63 time=16.1 ms
 08 bytes from 10.1.1.1: icmp_seq=4 ttl=63 time=16.6 ms
 08 bytes from 10.1.1.1: icmp_seq=5 ttl=63 time=9.17 ms
 08 bytes from 10.1.1.1: icmp_seq=6 ttl=63 time=9.29 ms
 08 bytes from 10.1.1.1: icmp_seq=7 ttl=63 time=8.86 ms
 08 bytes from 10.1.1.1: icmp_seq=8 ttl=63 time=8.89 ms
 08 bytes from 10.1.1.1: icmp_seq=9 ttl=63 time=9.22 ms
 08 bytes from 10.1.1.1: icmp_seq=10 ttl=63 time=10.5 ms
 08 bytes from 10.1.1.1: icmp_seq=11 ttl=63 time=10.9 ms
 08 bytes from 10.1.1.1: icmp_seq=12 ttl=63 time=9.12 ms
 08 bytes from 10.1.1.1: icmp_seq=13 ttl=63 time=10.0 ms
 08 bytes from 10.1.1.1: icmp_seq=14 ttl=63 time=9.60 ms
 08 bytes from 10.1.1.1: icmp_seq=15 ttl=63 time=9.25 ms
 08 bytes from 10.1.1.1: icmp_seq=16 ttl=63 time=10.1 ms
 08 bytes from 10.1.1.1: icmp_seq=17 ttl=63 time=9.29 ms
 08 bytes from 10.1.1.1: icmp_seq=18 ttl=63 time=8.37 ms
 08 bytes from 10.1.1.1: icmp_seq=19 ttl=63 time=9.36 ms
 08 bytes from 10.1.1.1: icmp_seq=20 ttl=63 time=8.03 ms
 08 bytes from 10.1.1.1: icmp_seq=21 ttl=63 time=11.1 ms
ping: sendmsg: Network is unreachable
ping: sendmsg: Network is unreachable
ping: sendmsg: Network is unreachable
ping: sendmsg: Network is unreachable
ping: sendmsg: Network is unreachable
 08 bytes from 10.1.1.1: icmp_seq=27 ttl=62 time=13.0 ms
 08 bytes from 10.1.1.1: icmp_seq=28 ttl=62 time=10.2 ms
 08 bytes from 10.1.1.1: icmp_seq=29 ttl=62 time=11.3 ms
 08 bytes from 10.1.1.1: icmp_seq=30 ttl=62 time=16.6 ms
^C
--- 10.1.1.1 ping statistics ---
30 packets transmitted, 25
```



```
SW-E# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.200.1

   Network        Nexthop        Metric      LocPrf      Weight Path
*>e 10.1.0.0/16    192.168.3.8        0           100         0       100 ?
*>e 192.168.1.10/32 192.168.3.11        0           100         0       65001 ?
Total number of entries 2
```

```
SW-B# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.2

   Network        Nexthop        Metric      LocPrf      Weight Path
*>e 10.1.0.0/16    192.168.3.10        0           100         0       200 100 ?
*>i 192.168.1.10/32 192.168.1.10        0           100         0       ?
Total number of entries 2
```

```
SW-B# show bgp ipv4 unicast neighbors 192.168.1.1 advertised-routes
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.2

   Network        Nexthop        Metric      LocPrf      Weight Path
*>i 10.1.0.0/16    192.168.1.2        0           100         0       200 100 ?
Total number of entries 1
```

```
SW-A# show bgp ipv4 un
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.1

   Network        Nexthop        Metric      LocPrf      Weight Path
*>i 10.1.0.0/16    192.168.1.2        0           100         0       200 100 ?
*>i 192.168.1.10/32 192.168.1.10        0           100         0       ?
Total number of entries 2
```

```
SW-C# show bgp ipv4 un
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.10

   Network        Nexthop        Metric      LocPrf      Weight Path
*>i 10.1.0.0/16    192.168.1.2        0           100         0       200 100 ?
*> 192.168.1.10/32 0.0.0.0            0           100         0       ?
Total number of entries 2
```


ADD-PATH

Overview

- ADD-PATH feature enables RR to send **more than one path** (called additional paths) for the same prefix.
- ADD-PATH is a BGP capability, with capability Code 69.
- With ADD-PATH, a BGP speaker can **send and/or receive** additional paths **to/from** a BGP Peer. It is a BGP extension that allows the advertisement of multiple paths (ADD-PATH) for the same prefix (NLRI).
- The additional paths include the **first (N-1) best paths**. The **total paths** for an address prefix received by a BGP speaker will include **the best path plus the additional paths** determined by its BGP peer.
- A BGP speaker uses the **Address Prefix** and a **Path Identifier** to uniquely identify a path advertised to a neighbor.
- A BGP speaker must generate a route update for the <AFI, SAFI> based on the combination of the **Address Prefix** and the **Path Identifier**, and use the extended NLRI encoding for PATH Identifier.
- Every device that needs to send or receive multiple paths must supports ADD-PATH capability.
- The alternate path's next-hop must be different than the primary path.

ADD-PATH

Not identical to

- AS-PATH multipath-relax (supported on AOS-CX)

```
SW-A(config-bgp)# bgp bestpath as-path
ignore          Do not consider as-path in best-path selection
multipath-relax  Consider routes with different AS-path but same length as ECMP
```

- BGP PIC (Prefix Independent Convergence)

<https://tools.ietf.org/html/draft-ietf-rtgwg-bgp-pic-12>

Instead of having a separate list of next-hops for each destination, all destinations sharing the same list of next-hops can point to a **single copy of this list** thereby allowing fast convergence by making changes to a single shared list of next-hops rather than possibly a large number of destinations.

Use Cases

aruba

a Hewlett Packard
Enterprise company

Use-Cases

Mainly for RR

- **Faster routing convergence**

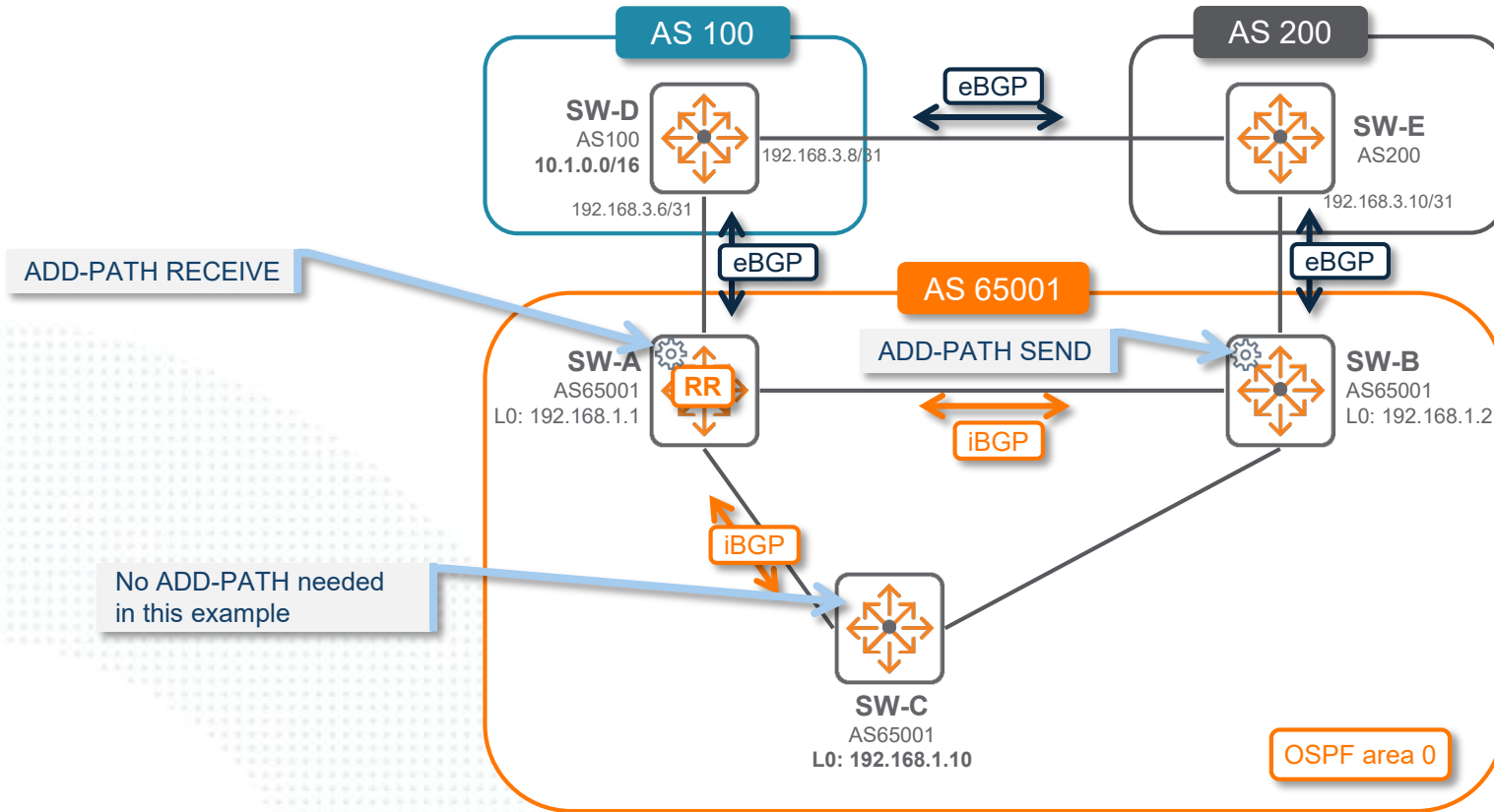
ADD-PATH helps to achieve faster re-convergence, as an alternative path is immediately available when a primary path fails.

- **Load-Balancing**

It provides also alternative path for support of load balancing.

Faster Convergence

Required ADD-PATH setting



Faster Convergence

BGP RIB

```
SW-D# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.100.1

   Network        Nexthop        Metric      LocPrf      Weight Path
*> 10.1.0.0/16     0.0.0.0              0           100         0      ?
*>e 192.168.1.10/32 192.168.3.7          0           100         0      65001 ?
* e 192.168.1.10/32 192.168.3.9          0           100         0      200 65001 ?
Total number of entries 3
```

```
SW-E# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.200.1

   Network        Nexthop        Metric      LocPrf      Weight Path
*>e 10.1.0.0/16     192.168.3.8          0           100         0      100 ?
* e 10.1.0.0/16     192.168.3.11         0           100         0      65001 100 ?
* e 192.168.1.10/32 192.168.3.8          0           100         0      100 65001 ?
*>e 192.168.1.10/32 192.168.3.11         0           100         0      65001 ?
Total number of entries 4
```

```
SW-B# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.2

   Network        Nexthop        Metric      LocPrf      Weight Path
*>i 10.1.0.0/16     192.168.1.1          0           100         0      100 ?
* e 10.1.0.0/16     192.168.3.10         0           100         0      200 100 ?
*>i 192.168.1.10/32 192.168.1.10         0           100         0      ?
Total number of entries 3
```

```
SW-B# sh bgp ipv4 unicast neighbors 192.168.1.1 advertised-routes
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.2

   Network        Nexthop        Metric      LocPrf      Weight Path
*>i 10.1.0.0/16     192.168.1.2          0           100         0      200 100 ?
Total number of entries 1
```

```
SW-C# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.10

   Network        Nexthop        Metric      LocPrf      Weight Path
*>i 10.1.0.0/16     192.168.1.1          0           100         0      100 ?
*> 192.168.1.10/32 0.0.0.0              0           100         0      ?
Total number of entries 2
```

```
SW-A# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.1

   Network        Nexthop        Metric      LocPrf      Weight Path
*>ai 10.1.0.0/16     192.168.1.2          0           100         0      200 100 ?
*>e 10.1.0.0/16     192.168.3.6          0           100         0      100 ?
*>i 192.168.1.10/32 192.168.1.10         0           100         0      ?
Total number of entries 3
```

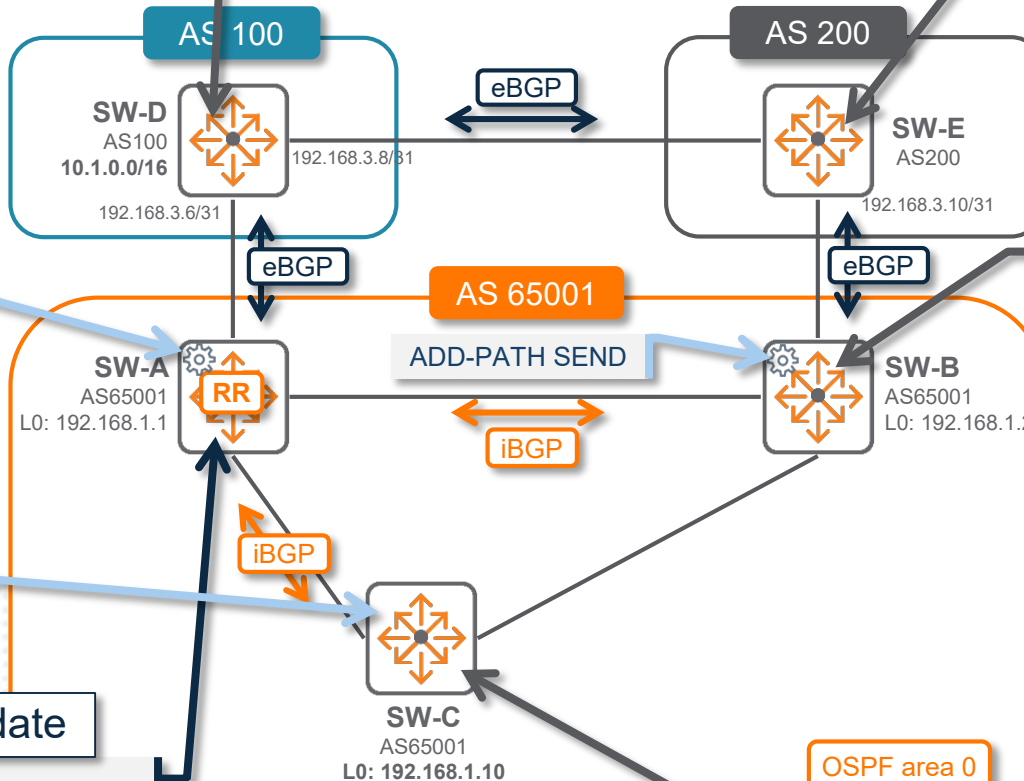
Alternate route is already candidate

No ADD-PATH needed in this example

ADD-PATH RECEIVE

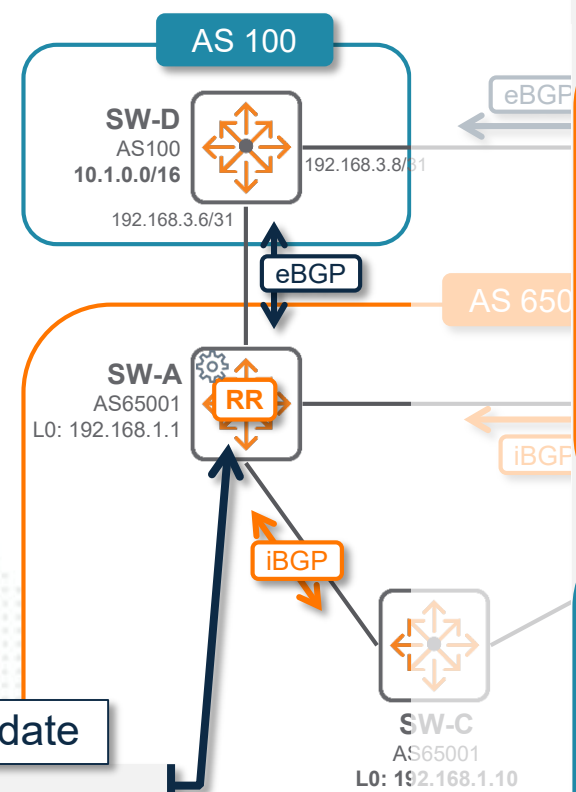
ADD-PATH SEND

OSPF area 0



Faster Convergence

BGP RIB



Alternate route is already candidate

```
SW-A# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.1
```

Network	Nexthop	Metric	LocPrf	Weight	Path
*ai 10.1.0.0/16	192.168.1.2	0	100	0	200 100 ?
*>e 10.1.0.0/16	192.168.3.6	0	100	0	100 ?
*>i 192.168.1.10/32	192.168.1.10	0	100	0	?

Total number of entries 3

```
SW-A# show bgp ipv4 unicast 10.1.0.0/16

VRF : default
BGP Local AS 65001          BGP Router-id 192.168.1.1
```

Network : 10.1.0.0/16
Peer : 192.168.1.2
Metric : 0
Weight : 0
Best : **No**
Type : internal
Originator ID : 0.0.0.0
Aggregator ID :
Aggregator AS :
Atomic Aggregate :
RFD Flaps : 0

Nexthop : 192.168.1.2
Origin : incomplete
Local Pref : 100
Calc. Local Pref : 100
Valid : Yes
Stale : No
Path ID : 294913

RFD Penalty : 0

AS-Path : 200 100

Cluster List :
Communities :
Ext-Communities :

Alternate route

Network : 10.1.0.0/16
Peer : 192.168.3.6
Metric : 0
Weight : 0
Best : **Yes**
Type : external
Originator ID : 0.0.0.0
Aggregator ID :
Aggregator AS :
Atomic Aggregate :
RFD Flaps : 0

Nexthop : 192.168.3.6
Origin : incomplete
Local Pref : 100
Calc. Local Pref : 100
Valid : Yes
Stale : No
Path ID : 0

RFD Penalty : 0

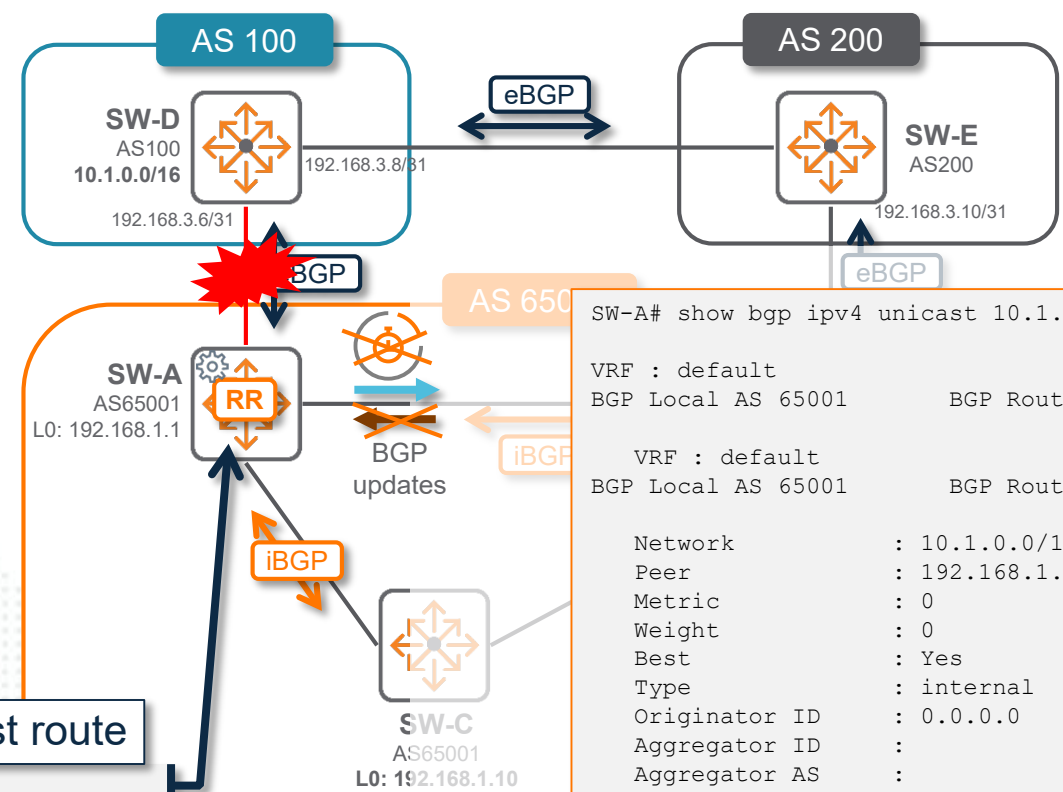
AS-Path : 100

Cluster List :
Communities :
Ext-Communities :

Best route

Faster Convergence

Reduced number of BGP updates



Alternate Route becomes Best route

The alternate route is now the best route

```
SW-A# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.1

  Network        Nexthop        Metric    LocPrf    Weight Path
*>i 10.1.0.0/16   192.168.1.2      0         100      0      200 100 ?
*>i 192.168.1.10/32 192.168.1.10    0         100      0      ?
Total number of entries 3
```

```
SW-A# show bgp ipv4 unicast 10.1.0.0/16

VRF : default
BGP Local AS 65001          BGP Router-id 192.168.1.1

VRF : default
BGP Local AS 65001          BGP Router-id 192.168.1.1

Network      : 10.1.0.0/16      Nexthop      : 192.168.1.2
Peer         : 192.168.1.2     Origin       : incomplete
Metric       : 0               Local Pref   : 100
Weight       : 0               Calc. Local Pref : 100
Best         : Yes             Valid        : Yes
Type         : internal        Stale        : No
Originator ID : 0.0.0.0        Path ID      : 0
Aggregator ID :
Aggregator AS :
Atomic Aggregate :
RFD Flaps    : 0               RFD Penalty  : 0

AS-Path      : 200 100

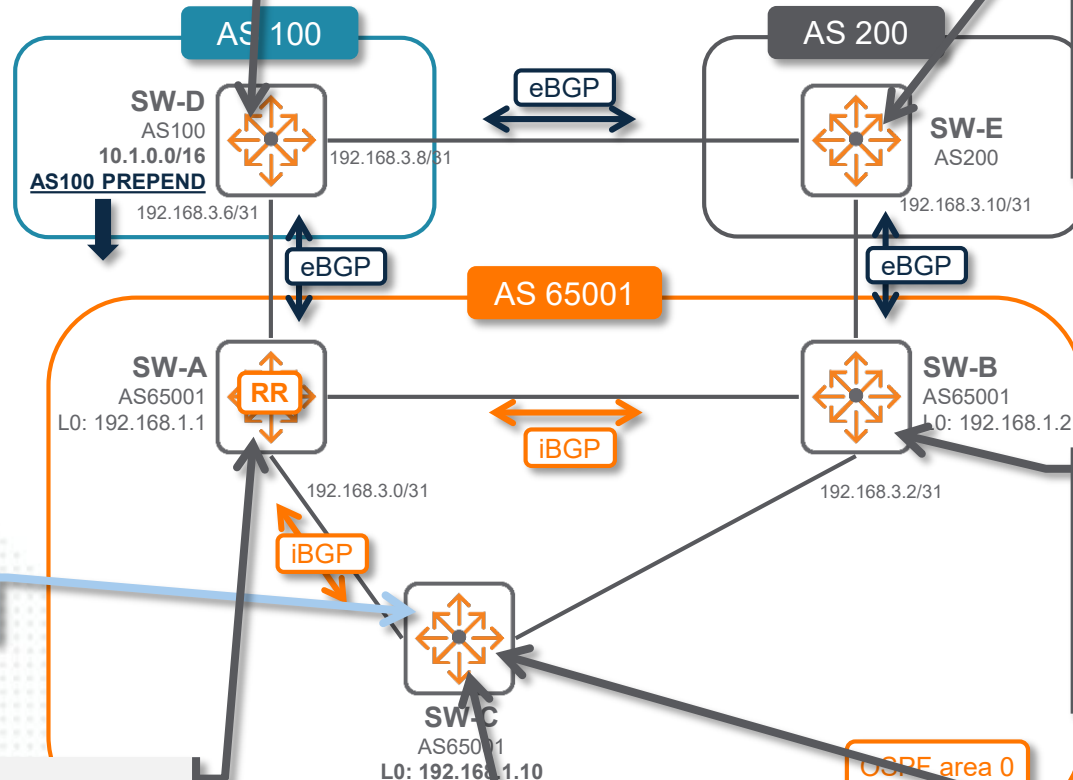
Cluster List :
Communities  :
Ext-Communities :
```

No Load-Balancing Without ADD-PATH

```
SW-D# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.100.1

   Network        Nexthop        Metric      LocPrf      Weight Path
*> 10.1.0.0/16     0.0.0.0              0           100         ?
*>e 192.168.1.10/32 192.168.3.7          0           100         0 65001 ?
* e 192.168.1.10/32 192.168.3.9          0           100         0 200 65001 ?
Total number of entries 3
```



```
SW-E# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.200.1

   Network        Nexthop        Metric      LocPrf      Weight Path
*>e 10.1.0.0/16     192.168.3.8          0           100         0 100 ?
* e 192.168.1.10/32 192.168.3.8          0           100         0 100 65001 ?
*>e 192.168.1.10/32 192.168.3.11         0           100         0 65001 ?
Total number of entries 3
```

```
SW-B# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.2

   Network        Nexthop        Metric      LocPrf      Weight Path
* i 10.1.0.0/16     192.168.1.1          0           100         0 100 100 ?
*>e 10.1.0.0/16     192.168.3.10         0           100         0 200 100 ?
*>i 192.168.1.10/32 192.168.1.10         0           100         0 ?
Total number of entries 3
```

```
SW-A# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.1

   Network        Nexthop        Metric      LocPrf      Weight Path
* i 10.1.0.0/16     192.168.1.2          0           100         0 200 100 ?
*>e 10.1.0.0/16     192.168.3.6          0           100         0 100 100 ?
*>i 192.168.1.10/32 192.168.1.10         0           100         0 ?
Total number of entries 3
```

```
SW-C# show ip route 10.1.0.0/16
Displaying ipv4 routes selected for forwarding

'[x/y]' denotes [distance/metric]

10.1.0.0/16, vrf default, tag 0
    via 192.168.3.0, [200/0], bgp
```

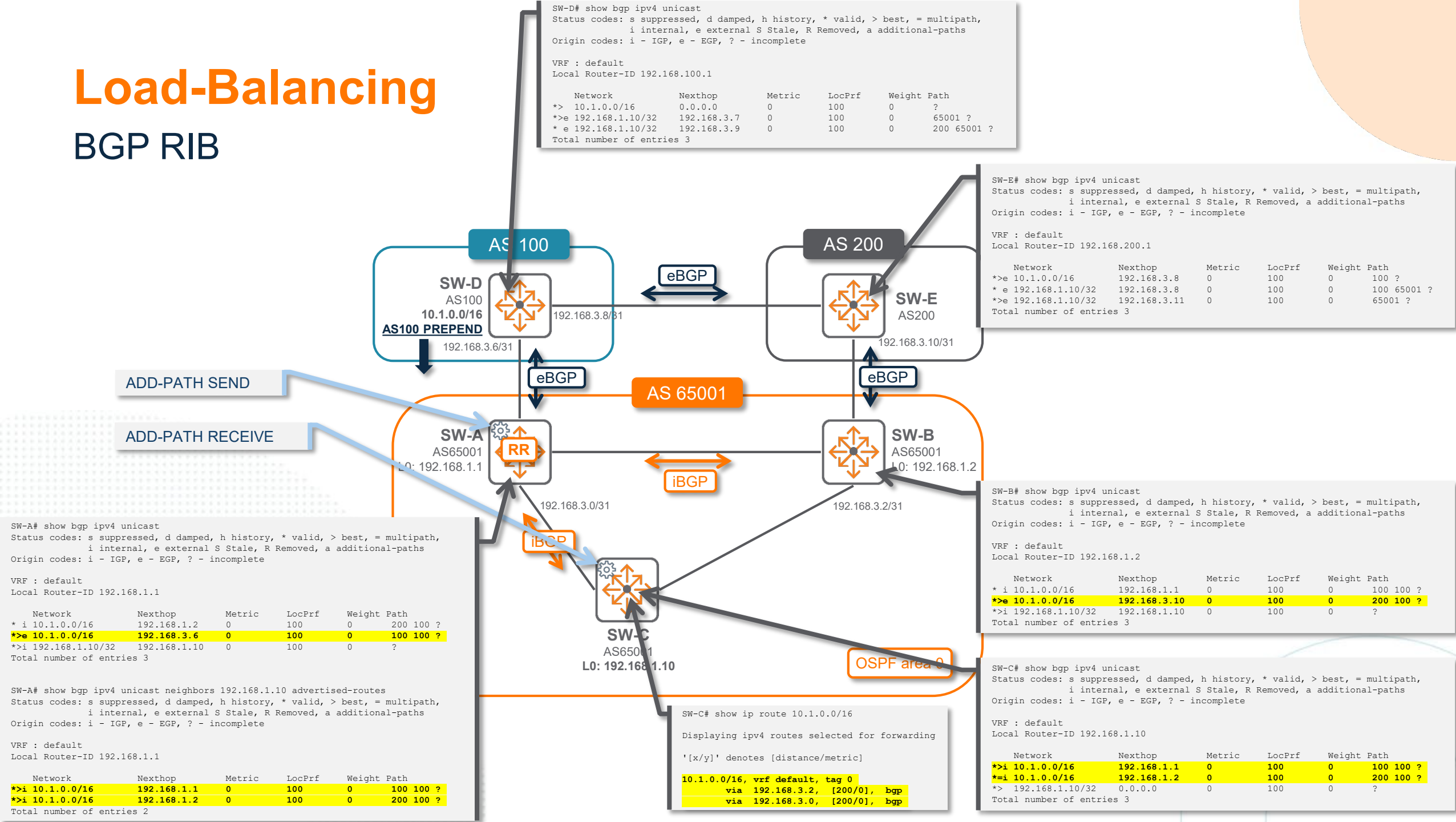
```
SW-C# show bgp ipv4 un
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 192.168.1.10

   Network        Nexthop        Metric      LocPrf      Weight Path
*>i 10.1.0.0/16     192.168.1.1          0           100         0 100 ?
*> 192.168.1.10/32 0.0.0.0              0           100         0 ?
Total number of entries 2
```

Load-Balancing

BGP RIB



Details / Caveats

Platform Support

10.6 – ADD-PATH support

ADD-PATH support	6200	6300	6400	8360	8320	8325	8400
MP-BGP IPv4 AF	N/A	Yes	Yes	Yes	Yes	Yes	Yes
MP-BGP IPv6 AF		Yes	Yes	Yes	Yes	Yes	Yes
MP-BGP EVPN AF		No	No	No	No	No	No

▪ Scale:

- BGP additional path feature increases processing load and memory utilization since the additional backup paths increase the BGP table size.
- This is purely software consumption (no ASIC programming).
- Time to load the BGP table will be impacted due to the number of additional paths to compute.
- The number of candidate additional paths selected for each destination should be kept small.
A value of 2 is sufficient to calculate the best path and a backup path for each destination.

ADD-PATH Details

OPEN message from ADD-PATH RECEIVER

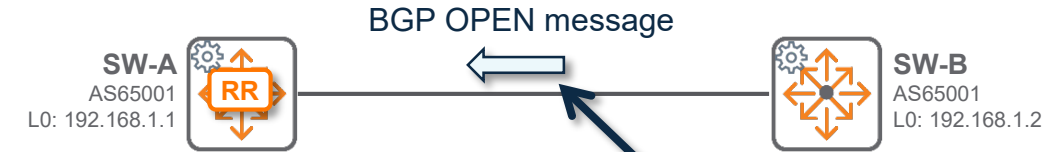
- RR should be capable of “**receiving**” the additional paths.



ADD-PATH Details

OPEN message from ADD-PATH SENDER

- RR-client for alternate path should be capable of “**sending**” the additional paths.



```
No.      Time      Source      Destination      Protocol Length Info
  14  9.242847    192.168.1.2    192.168.1.1      BGP         129    OPEN Message

Frame 14: 129 bytes on wire (1032 bits), 129 bytes captured (1032 bits) on interface -, id 0
Ethernet II, Src: HewlettP_f9:41:3c (08:00:09:f9:41:3c), Dst: HewlettP_df:10:5f (08:00:09:df:10:5f)
Internet Protocol Version 4, Src: 192.168.1.2, Dst: 192.168.1.1
Transmission Control Protocol, Src Port: 179, Dst Port: 40167, Seq: 1, Ack: 64, Len: 63
Border Gateway Protocol - OPEN Message
  Marker: ffffffffffffffffffffffffffffffff
  Length: 63
  Type: OPEN Message (1)
  Version: 4
  My AS: 65001
  Hold Time: 180
  BGP Identifier: 192.168.1.2
  Optional Parameters Length: 34
  Optional Parameters
    Optional Parameter: Capability
      Parameter Type: Capability (2)
      Parameter Length: 32
      Capability: Multiprotocol extensions capability
        Type: Multiprotocol extensions capability (1)
        Length: 4
        AFI: IPv4 (1)
        Reserved: 00
        SAFI: Unicast (1)
      Capability: Route refresh capability
        Type: Route refresh capability (2)
        Length: 0
      Capability: Route refresh capability (Cisco)
        Type: Route refresh capability (Cisco) (128)
        Length: 0
      Capability: Enhanced route refresh capability
        Type: Enhanced route refresh capability (70)
        Length: 0
      Capability: Graceful Restart capability
        Type: Graceful Restart capability (64)
        Length: 6
        Restart Timers: 0x0078
          0... .. = Restart state: No
          .0... .. = Graceful notification: No
          ... 0000 0111 1000 = Time: 120
        AFI: IPv4 (1)
        SAFI: Unicast (1)
        Flag: 0x00
          0... .. = Preserve forwarding state: No
      Capability: Support for 4-octet AS number capability
        Type: Support for 4-octet AS number capability (65)
        Length: 4
        AS Number: 65001
      Capability: Support for Additional Paths
        Type: Support for Additional Paths (69)
        Length: 4
        AFI: IPv4 (1)
        SAFI: Unicast (1)
        Send/Receive: Send (2)
```

Detailed Routes Exchange

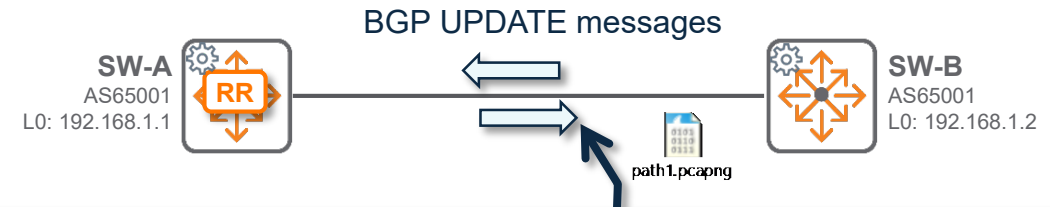
BGP UPDATE messages

1. SW-B to SW-A: NLRI update for 10.1.0.0/16 and empty update message

```
Border Gateway Protocol - UPDATE Message
Marker: ffffffffffffffffffffffff
Length: 61
Type: UPDATE Message (2)
Withdrawn Routes Length: 0
Total Path Attribute Length: 31
Path attributes
  Path Attribute - ORIGIN: INCOMPLETE
  Path Attribute - AS_PATH: 200 100
  Path Attribute - NEXT_HOP: 192.168.1.2
  Path Attribute - LOCAL_PREF: 100
Network Layer Reachability Information (NLRI)
  10.1.0.0/16 PathId 360449
  NLRI path id: 360449
  Prefix Length: 16
  NLRI prefix: 10.1.0.0
```

2. SW-A to SW-B: NLRI update for 192.168.1.10/32 and empty update message
3. SW-A to SW-B: NLRI update for 10.1.0.0/16

```
Border Gateway Protocol - UPDATE Message
Marker: ffffffffffffffffffffffff
Length: 53
Type: UPDATE Message (2)
Withdrawn Routes Length: 0
Total Path Attribute Length: 27
Path attributes
  Path Attribute - ORIGIN: INCOMPLETE
  Path Attribute - AS_PATH: 100
  Path Attribute - NEXT_HOP: 192.168.1.1
  Path Attribute - LOCAL_PREF: 100
Network Layer Reachability Information (NLRI)
  10.1.0.0/16
  NLRI prefix length: 16
  NLRI prefix: 10.1.0.0
```



No.	Time	Source	Destination	Protocol	Length	Info
5	9.219268	192.168.1.1	192.168.1.2	TCP	66	179 → 36663 [FIN, ACK] Seq=29 Ack=1 Win=227 Len=0
6	9.220917	192.168.1.2	192.168.1.1	TCP	66	36663 → 179 [ACK] Seq=1 Ack=30 Win=30016 Len=0
7	9.223467	192.168.1.2	192.168.1.1	TCP	66	36663 → 179 [FIN, ACK] Seq=1 Ack=30 Win=30016 Len=0
8	9.226201	192.168.1.1	192.168.1.2	TCP	66	179 → 36663 [ACK] Seq=30 Ack=2 Win=227 Len=0
9	9.226957	192.168.1.1	192.168.1.2	TCP	74	40167 → 179 [SYN] Seq=0 Win=29200 Len=0 MSS=1460
10	9.229806	192.168.1.2	192.168.1.1	TCP	74	179 → 40167 [SYN, ACK] Seq=0 Ack=1 Win=28960 Len=0
11	9.232124	192.168.1.1	192.168.1.2	TCP	66	40167 → 179 [ACK] Seq=1 Ack=1 Win=29200 Len=0
12	9.239971	192.168.1.1	192.168.1.2	BGP	129	OPEN Message
13	9.242037	192.168.1.2	192.168.1.1	TCP	66	179 → 40167 [ACK] Seq=1 Ack=64 Win=29056 Len=0
14	9.242847	192.168.1.2	192.168.1.1	BGP	129	OPEN Message
15	9.245068	192.168.1.1	192.168.1.2	TCP	66	40167 → 179 [ACK] Seq=64 Ack=64 Win=29200 Len=0
16	9.267580	192.168.1.2	192.168.1.1	BGP	85	KEEPALIVE Message
17	9.269183	192.168.1.1	192.168.1.2	BGP	85	KEEPALIVE Message
18	9.312193	192.168.1.1	192.168.1.2	TCP	66	40167 → 179 [ACK] Seq=83 Ack=83 Win=29200 Len=0
19	9.312838	192.168.1.2	192.168.1.1	TCP	66	179 → 40167 [ACK] Seq=83 Ack=83 Win=29056 Len=0
20	9.314099	192.168.1.2	192.168.1.1	BGP	150	UPDATE Message, UPDATE Message
21	9.315235	192.168.1.1	192.168.1.2	BGP	152	UPDATE Message, UPDATE Message
22	9.315938	192.168.1.1	192.168.1.2	TCP	66	40167 → 179 [ACK] Seq=169 Ack=167 Win=29200 Len=0
23	9.317930	192.168.1.2	192.168.1.1	TCP	66	179 → 40167 [ACK] Seq=167 Ack=169 Win=29056 Len=0
24	9.349245	192.168.1.1	192.168.1.2	BGP	119	UPDATE Message
25	9.353577	192.168.1.2	192.168.1.1	TCP	66	179 → 40167 [ACK] Seq=167 Ack=222 Win=29056 Len=0
26	10.001363	192.168.3.4	224.0.0.5	OSPF	82	Hello Packet
27	13.090071	HewlettP_f9:...	LLDP_Multicast	LLDP	134	MA/08:00:09:f9:41:3c IN/1/1/1 120 SysN=SW-B SysID=1
28	15.725576	192.168.3.5	224.0.0.5	OSPF	82	Hello Packet
29	20.002297	192.168.3.4	224.0.0.5	OSPF	82	Hello Packet

```
> Frame 24: 119 bytes on wire (952 bits), 119 bytes captured (952 bits) on interface -, id 0
> Ethernet II, Src: HewlettP_df:10:5f (08:00:09:df:10:5f), Dst: HewlettP_f9:41:3c (08:00:09:f9:41:3c)
> Internet Protocol Version 4, Src: 192.168.1.1, Dst: 192.168.1.2
> Transmission Control Protocol, Src Port: 40167, Dst Port: 179, Seq: 169, Ack: 167, Len: 53
> Border Gateway Protocol - UPDATE Message
  Marker: ffffffffffffffffffffffff
  Length: 53
  Type: UPDATE Message (2)
  Withdrawn Routes Length: 0
  Total Path Attribute Length: 27
  Path attributes
  > Network Layer Reachability Information (NLRI)
    > 10.1.0.0/16
      NLRI prefix length: 16
      NLRI prefix: 10.1.0.0
```

Configuration

aruba

a Hewlett Packard
Enterprise company

ADD-PATH configuration

2 steps per IPv4 or IPv6 AF

- This feature consists in two commands:

1. **Enable the ADD-PATH capability:** the router notifies its neighbor about its ADD-PATH capability:

- Routing scenario analysis is required to identify where to set-up ADD-PATH **recv** or **send**. It might be useful to configure **both** in some cases.
- Depending on use-cases, a RR is configured to receive additional paths.
- Depending on use-cases, the RR-Client hosting the alternate path is configured to send the corresponding additional paths.

2. **Configure the total number of best-paths to advertise:**

- Configure the number of best paths to send to the neighbor.
- This command is **REQUIRED on the sender**.
- The default value is 2 (absolute best path + next best path).

ADD-PATH RECEIVER

```
SW-A(config-bgp-ipv4-uc)# neighbor 192.168.1.2
  add-paths      Configure the additional paths capability for the neighbor.
SW-A(config-bgp-ipv4-uc)# neighbor 192.168.1.2 add-paths
  advertise-best Configure the number of best paths to advertise to the neighbor.
  both           Configure to advertise and receive additional paths for the neighbor.
  recv           Configure to only receive additional paths from the neighbor.
  send           Configure to only send additional paths to the neighbor.
SW-A(config-bgp-ipv4-uc)# neighbor 192.168.1.2 add-paths recv
```

capability

ADD-PATH SENDER

```
SW-B(config-bgp-ipv4-uc)# neighbor 192.168.1.1
  add-paths      Configure the additional paths capability for the neighbor.
SW-B(config-bgp-ipv4-uc)# neighbor 192.168.1.1 add-paths
  advertise-best Configure the number of best paths to advertise to the neighbor.
  both           Configure to advertise and receive additional paths for the neighbor.
  recv           Configure to only receive additional paths from the neighbor.
  send           Configure to only send additional paths to the neighbor.
SW-B(config-bgp-ipv4-uc)# neighbor 192.168.1.1 add-paths send
```

Number of additional paths

```
SW-B(config-bgp-ipv4-uc)# neighbor 192.168.1.1 add-paths advertise-best
  [<2-4>] Set the number of best paths to send to the neighbor. (Default: 2)
SW-B(config-bgp-ipv4-uc)# neighbor 192.168.1.1 add-paths advertise-best 2
```



Baseline

Without ADD-PATH

```
router bgp 65001
  bgp router-id 192.168.1.1
  bgp cluster-id 1.1.1.1
  neighbor 192.168.1.2 remote-as 65001
  neighbor 192.168.1.2 update-source loopback 0
  neighbor 192.168.1.10 remote-as 65001
  neighbor 192.168.1.10 update-source loopback 0
  neighbor 192.168.3.6 remote-as 100
  address-family ipv4 unicast
    neighbor 192.168.1.2 activate
    neighbor 192.168.1.2 next-hop-self
    neighbor 192.168.1.2 route-reflector-client
    neighbor 192.168.1.10 activate
    neighbor 192.168.1.10 next-hop-self
    neighbor 192.168.1.10 route-reflector-client
    neighbor 192.168.3.6 activate
  exit-address-family
```

SW-A
AS65001
L0: 192.168.1.1



SW-B
AS65001
L0: 192.168.1.2



SW-C
AS65001
L0: 192.168.1.10



```
router bgp 65001
  bgp router-id 192.168.1.10
  neighbor 192.168.1.1 remote-as 65001
  neighbor 192.168.1.1 update-source loopback 0
  address-family ipv4 unicast
    neighbor 192.168.1.1 activate
    redistribute local loopback
  exit-address-family
```

```
router bgp 65001
  bgp router-id 192.168.1.2
  neighbor 192.168.1.1 remote-as 65001
  neighbor 192.168.1.1 update-source loopback 0
  neighbor 192.168.3.10 remote-as 200
  address-family ipv4 unicast
    neighbor 192.168.1.1 activate
    neighbor 192.168.1.1 next-hop-self
    neighbor 192.168.3.10 activate
  exit-address-family
```


Baseline

Without ADD-PATH

```
SW-A# sh bgp ipvr neighbors 192.168.1.1
Codes: ^ Inherited from peer-group
```

VRF : default

BGP Neighbor 192.168.1.1 (Internal)

Description	:		
Peer-group	:		
Remote Router Id	:	192.168.1.2	Local Router Id : 192.168.1.1
Remote AS	:	65001	Local AS : 65001
Remote Port	:	179	Local Port : 37939
State	:	Established	Admin Status : Up
Conn. Established	:	1	Conn. Dropped : 0
Passive	:	No	Update-Source : loopback0
Cfg. Hold Time	:	180	Cfg. Keep Alive : 60
Neg. Hold Time	:	180	Neg. Keep Alive : 60
Up/Down Time	:	00h:37m:18s	Alt. Local-AS : 0
Local-AS Prepend	:	No	
BFD	:	Disabled	
Password	:		
Last Err Sent	:	No Error	
Last SubErr Sent	:	No Error	
Last Err Rcvd	:	No Error	
Last SubErr Rcvd	:	No Error	

Graceful-Restart	:	Enabled	Gr. Restart Time : 120
Gr. Stalepath Time	:	300	Remove Private-AS : No
TTL	:	255	Local Cluster-ID : 1.1.1.1
Weight	:	0	Fall-over : No
Confederation-Peers	:	No	

Message statistics	Sent	Rcvd
Open	1	1
Notification	0	0
Updates	5	4
Keepalives	42	43
Route Refresh	0	0
Total	48	48

Capability	Advertised	Received
Route Refresh	Yes	Yes
Graceful Restart	Yes	Yes
Add-Path	No	No
Four Octet ASN	Yes	Yes
Address family IPv4 Unicast	Yes	Yes
Address family IPv6 Unicast	No	No
Address family L2VPN EVPN	No	No

Address Family : IPv4 Unicast

Rt. Reflect. Client	:	Yes	Send Community	:	
Allow-AS in	:	0	Advt. Interval	:	30
Max. Prefix	:	64000	Soft Reconfig In	:	
NextHop-Self	:	Yes	Default-Originate	:	
Cfg. Add-Path	:				
Neg. Add-Path	:	Disable			

Routemap In	:	
Routemap Out	:	
ORF type	:	Prefix-list
ORF capability	:	

SW-A
AS65001
L0: 192.168.1.1



SW-C
AS65001
L0: 192.168.1.10

iBGP



SW-B
AS65001
L0: 192.168.1.2

```
SW-B# sh bgp ipv4 unicast neighbors 192.168.1.1
Codes: ^ Inherited from peer-group
```

VRF : default

BGP Neighbor 192.168.1.1 (Internal)

Description	:		
Peer-group	:		
Remote Router Id	:	192.168.1.1	Local Router Id : 192.168.1.2
Remote AS	:	65001	Local AS : 65001
Remote Port	:	37939	Local Port : 179
State	:	Established	Admin Status : Up
Conn. Established	:	2	Conn. Dropped : 1
Passive	:	No	Update-Source : loopback0
Cfg. Hold Time	:	180	Cfg. Keep Alive : 60
Neg. Hold Time	:	180	Neg. Keep Alive : 60
Up/Down Time	:	00h:41m:13s	Alt. Local-AS : 0
Local-AS Prepend	:	No	
BFD	:	Disabled	
Password	:		
Last Err Sent	:	No Error	
Last SubErr Sent	:	No Error	
Last Err Rcvd	:	No Error	
Last SubErr Rcvd	:	No Error	

Graceful-Restart	:	Enabled	Gr. Restart Time : 120
Gr. Stalepath Time	:	300	Remove Private-AS : No
TTL	:	255	Local Cluster-ID : 0.0.0.0
Weight	:	0	Fall-over : No
Confederation-Peers	:	No	

Message statistics	Sent	Rcvd
Open	2	2
Notification	0	0
Updates	8	9
Keepalives	48	47
Route Refresh	0	0
Total	58	58

Capability	Advertised	Received
Route Refresh	Yes	Yes
Graceful Restart	Yes	Yes
Add-Path	No	No
Four Octet ASN	Yes	Yes
Address family IPv4 Unicast	Yes	Yes
Address family IPv6 Unicast	No	No
Address family L2VPN EVPN	No	No

Address Family : IPv4 Unicast

Rt. Reflect. Client	:	No	Send Community	:	
Allow-AS in	:	0	Advt. Interval	:	30
Max. Prefix	:	64000	Soft Reconfig In	:	
NextHop-Self	:	Yes	Default-Originate	:	
Cfg. Add-Path	:				
Neg. Add-Path	:	Disable			

Routemap In	:	
Routemap Out	:	
ORF type	:	Prefix-list
ORF capability	:	

Baseline

Without ADD-PATH

```
SW-A# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete
```

```
VRF : default
Local Router-ID 192.168.1.1
```

Network	Nexthop	Metric	LocPrf	Weight	Path
*>e 10.1.0.0/16	192.168.3.6	0	100	0	100 ?
*>i 192.168.1.10/32	192.168.1.10	0	100	0	?

Total number of entries 2

```
SW-A# show bgp ipv4 unicast 10.1.0.0/16
```

```
VRF : default
BGP Local AS 65001      BGP Router-id 192.168.1.1
```

Network	: 10.1.0.0/16	Nexthop	: 192.168.3.6
Peer	: 192.168.3.6	Origin	: incomplete
Metric	: 0	Local Pref	: 100
Weight	: 0	Calc. Local Pref	: 100
Best	: Yes	Valid	: Yes
Type	: external	Stale	: No
Originator ID	: 0.0.0.0	Path ID	: 0
Aggregator ID	:		
Aggregator AS	:		
Atomic Aggregate	:		
RFD Flaps	: 0	RFD Penalty	: 0

```
AS-Path      : 100
```

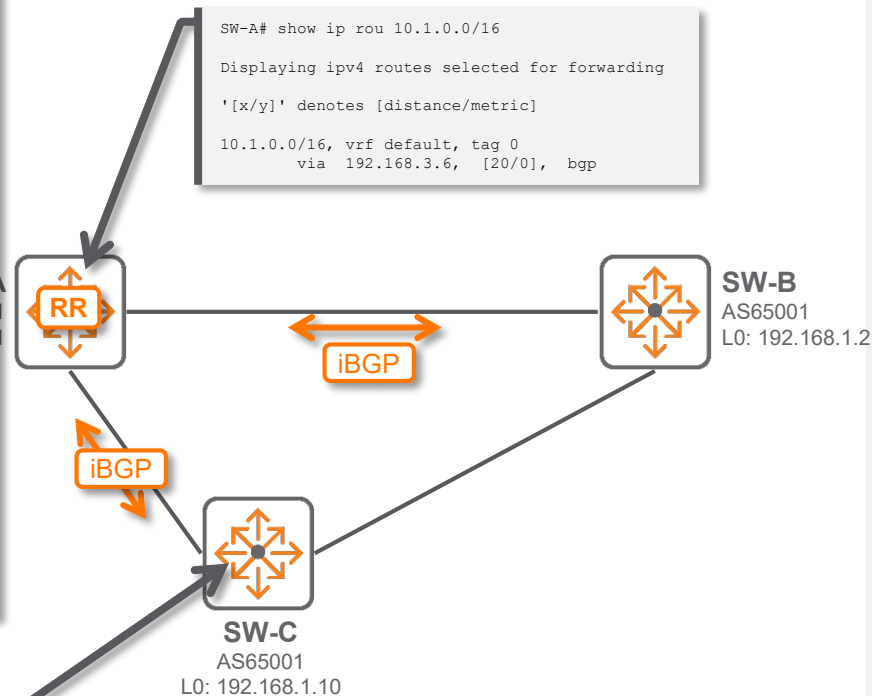
```
Cluster List :
Communities  :
Ext-Communities :
```

```
SW-A# show ip rou 10.1.0.0/16
```

Displaying ipv4 routes selected for forwarding

'[x/y]' denotes [distance/metric]

```
10.1.0.0/16, vrf default, tag 0
via 192.168.3.6, [20/0], bgp
```



```
SW-C# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete
```

```
VRF : default
Local Router-ID 192.168.1.10
```

Network	Nexthop	Metric	LocPrf	Weight	Path
*>i 10.1.0.0/16	192.168.1.1	0	100	0	100 ?
*> 192.168.1.10/32	0.0.0.0	0	100	0	?

Total number of entries 2

```
SW-B# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete
```

```
VRF : default
Local Router-ID 192.168.1.2
```

Network	Nexthop	Metric	LocPrf	Weight	Path
*>i 10.1.0.0/16	192.168.1.1	0	100	0	100 ?
* e 10.1.0.0/16	192.168.3.10	0	100	0	200 100 ?
*>i 192.168.1.10/32	192.168.1.10	0	100	0	?

Total number of entries 3

```
SW-B# sh bgp ipv4 unicast 10.1.0.0/16
```

```
VRF : default
BGP Local AS 65001      BGP Router-id 192.168.1.2
```

Network	: 10.1.0.0/16	Nexthop	: 192.168.1.1
Peer	: 192.168.1.1	Origin	: incomplete
Metric	: 0	Local Pref	: 100
Weight	: 0	Calc. Local Pref	: 100
Best	: Yes	Valid	: Yes
Type	: internal	Stale	: No
Originator ID	: 0.0.0.0	Path ID	: 0
Aggregator ID	:		
Aggregator AS	:		
Atomic Aggregate	:		
RFD Flaps	: 0	RFD Penalty	: 0

```
AS-Path      : 100
```

```
Cluster List :
Communities  :
Ext-Communities :
```

Network	: 10.1.0.0/16	Nexthop	: 192.168.3.10
Peer	: 192.168.3.10	Origin	: incomplete
Metric	: 0	Local Pref	: 100
Weight	: 0	Calc. Local Pref	: 100
Best	: No	Valid	: Yes
Type	: external	Stale	: No
Originator ID	: 0.0.0.0	Path ID	: 0
Aggregator ID	:		
Aggregator AS	:		
Atomic Aggregate	:		
RFD Flaps	: 0	RFD Penalty	: 0

```
AS-Path      : 200 100
```

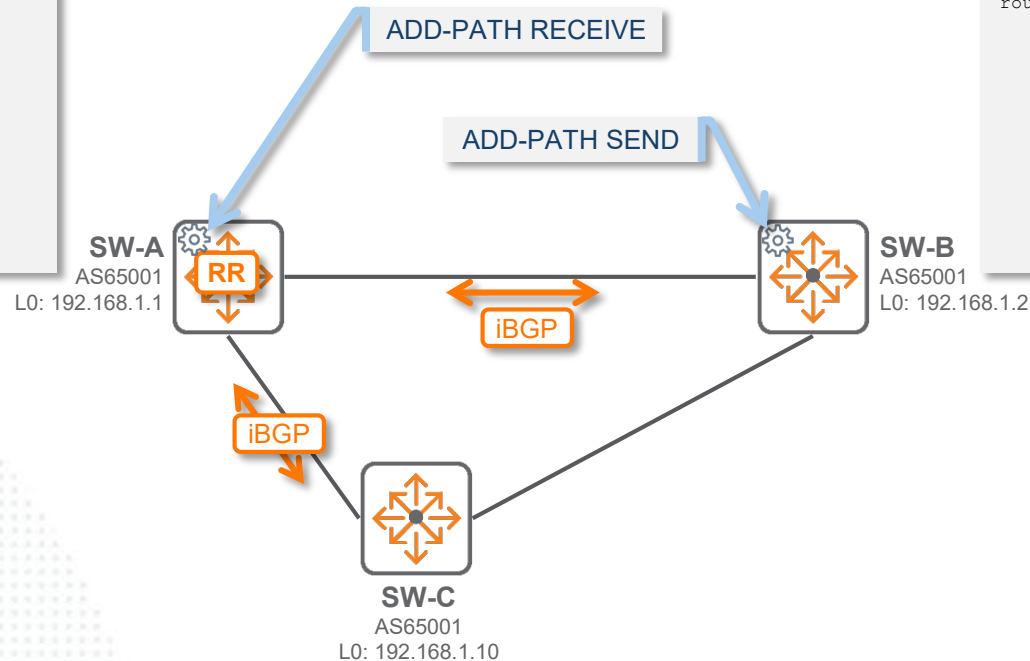
```
Cluster List :
Communities  :
Ext-Communities :
```



ADD-PATH Configuration

ADD-PATH capability + number of paths

```
router bgp 65001
  bgp router-id 192.168.1.1
  bgp cluster-id 1.1.1.1
  neighbor 192.168.1.2 remote-as 65001
  neighbor 192.168.1.2 update-source loopback 0
  neighbor 192.168.1.10 remote-as 65001
  neighbor 192.168.1.10 update-source loopback 0
  neighbor 192.168.3.6 remote-as 100
  address-family ipv4 unicast
    neighbor 192.168.1.2 activate
    neighbor 192.168.1.2 add-paths recv
    neighbor 192.168.1.2 next-hop-self
    neighbor 192.168.1.2 route-reflector-client
    neighbor 192.168.1.10 activate
    neighbor 192.168.1.10 next-hop-self
    neighbor 192.168.1.10 route-reflector-client
    neighbor 192.168.3.6 activate
  exit-address-family
```



```
router bgp 65001
  bgp router-id 192.168.1.2
  neighbor 192.168.1.1 remote-as 65001
  neighbor 192.168.1.1 update-source loopback 0
  neighbor 192.168.3.10 remote-as 200
  address-family ipv4 unicast
    neighbor 192.168.1.1 activate
    neighbor 192.168.1.1 add-paths send
    neighbor 192.168.1.1 add-paths advertise-best 2
    neighbor 192.168.1.1 next-hop-self
    neighbor 192.168.3.10 activate
  exit-address-family
```

```
router bgp 65001
  bgp router-id 192.168.1.10
  neighbor 192.168.1.1 remote-as 65001
  neighbor 192.168.1.1 update-source loopback 0
  address-family ipv4 unicast
    neighbor 192.168.1.1 activate
    redistribute local loopback
  exit-address-family
```

ADD-PATH Configuration

ADD-PATH capability

```
SW-A# sh bgp ipv4 unicast neighbors 192.168.1.2
Codes: ^ Inherited from peer-group
```

VRF : default

```
BGP Neighbor 192.168.1.2 (Internal)
Description      :
Peer-group       :
```

Remote Router Id	: 192.168.1.2	Local Router Id	: 192.168.1.1
Remote AS	: 65001	Local AS	: 65001
Remote Port	: 179	Local Port	: 37939
State	: Established	Admin Status	: Up
Conn. Established	: 1	Conn. Dropped	: 0
Passive	: No	Update-Source	: loopback0
Cfg. Hold Time	: 180	Cfg. Keep Alive	: 60
Neg. Hold Time	: 180	Neg. Keep Alive	: 60
Up/Down Time	: 00h:37m:18s	Alt. Local-AS	: 0
Local-AS Prepend	: No		
BFD	: Disabled		
Password	:		
Last Err Sent	: No Error		
Last SubErr Sent	: No Error		
Last Err Rcvd	: No Error		
Last SubErr Rcvd	: No Error		

Graceful-Restart	: Enabled	Gr. Restart Time	: 120
Gr. Stalepath Time	: 300	Remove Private-AS	: No
TTL	: 255	Local Cluster-ID	: 1.1.1.1
Weight	: 0	Fall-over	: No
Confederation-Peers	: No		

Message statistics	Sent	Rcvd
Open	1	1
Notification	0	0
Updates	5	4
Keepalives	42	43
Route Refresh	0	0
Total	48	48

Capability	Advertised	Received
Route Refresh	Yes	Yes
Graceful Restart	Yes	Yes
Add-Path	Yes	Yes
Four Octet ASN	Yes	Yes
Address family IPv4 Unicast	Yes	Yes
Address family IPv6 Unicast	No	No
Address family L2VPN EVPN	No	No

Address Family : IPv4 Unicast

Rt. Reflect. Client	: Yes	Send Community	:
Allow-AS in	: 0	Advt. Interval	: 30
Max. Prefix	: 64000	Soft Reconfig In	:
NextHop-Self	: Yes	Default-Originate	:
Cfg. Add-Path	: Receive		
Neg. Add-Path	: Receive		

Routemap In :
Routemap Out :
ORF type : Prefix-list
ORF capability :

SW-A
AS65001
L0: 192.168.1.1



SW-C
AS65001
L0: 192.168.1.10



SW-B
AS65001
L0: 192.168.1.2

```
SW-B# sh bgp ipv4 unicast neighbors 192.168.1.1
Codes: ^ Inherited from peer-group
```

VRF : default

```
BGP Neighbor 192.168.1.1 (Internal)
Description      :
Peer-group       :
```

Remote Router Id	: 192.168.1.1	Local Router Id	: 192.168.1.2
Remote AS	: 65001	Local AS	: 65001
Remote Port	: 37939	Local Port	: 179
State	: Established	Admin Status	: Up
Conn. Established	: 2	Conn. Dropped	: 1
Passive	: No	Update-Source	: loopback0
Cfg. Hold Time	: 180	Cfg. Keep Alive	: 60
Neg. Hold Time	: 180	Neg. Keep Alive	: 60
Up/Down Time	: 00h:41m:13s	Alt. Local-AS	: 0
Local-AS Prepend	: No		
BFD	: Disabled		
Password	:		
Last Err Sent	: No Error		
Last SubErr Sent	: No Error		
Last Err Rcvd	: No Error		
Last SubErr Rcvd	: No Error		

Graceful-Restart	: Enabled	Gr. Restart Time	: 120
Gr. Stalepath Time	: 300	Remove Private-AS	: No
TTL	: 255	Local Cluster-ID	: 0.0.0.0
Weight	: 0	Fall-over	: No
Confederation-Peers	: No		

Message statistics	Sent	Rcvd
Open	2	2
Notification	0	0
Updates	8	9
Keepalives	48	47
Route Refresh	0	0
Total	58	58

Capability	Advertised	Received
Route Refresh	Yes	Yes
Graceful Restart	Yes	Yes
Add-Path	Yes	Yes
Four Octet ASN	Yes	Yes
Address family IPv4 Unicast	Yes	Yes
Address family IPv6 Unicast	No	No
Address family L2VPN EVPN	No	No

Address Family : IPv4 Unicast

Rt. Reflect. Client	: No	Send Community	:
Allow-AS in	: 0	Advt. Interval	: 30
Max. Prefix	: 64000	Soft Reconfig In	:
NextHop-Self	: Yes	Default-Originate	:
Cfg. Add-Path	: Send		
Neg. Add-Path	: Send		

Routemap In :
Routemap Out :
ORF type : Prefix-list
ORF capability :

ADD-PATH Configuration

Outcome

```
SW-A# sh bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete
```

```
VRF : default
Local Router-ID 192.168.1.1
```

Network	Nexthop	Metric	LocPrf	Weight	Path
*ai 10.1.0.0/16	192.168.1.2	0	100	0	200 100 ?
*>e 10.1.0.0/16	192.168.3.6	0	100	0	100 ?
*>i 192.168.1.10/32	192.168.1.10	0	100	0	?

Total number of entries 3

```
SW-A# sh bgp ipv4 unicast 10.1.0.0/16
```

```
VRF : default
BGP Local AS 65001      BGP Router-id 192.168.1.1
```

Network	: 10.1.0.0/16	Nexthop	: 192.168.1.2
Peer	: 192.168.1.2	Origin	: incomplete
Metric	: 0	Local Pref	: 100
Weight	: 0	Calc. Local Pref	: 100
Best	: No	Valid	: Yes
Type	: internal	Stale	: No
Originator ID	: 0.0.0.0	Path ID	: 327681
Aggregator ID	:		
Aggregator AS	:		
Atomic Aggregate	:		
RFD Flaps	: 0	RFD Penalty	: 0
AS-Path	: 200 100		

Network	: 10.1.0.0/16	Nexthop	: 192.168.3.6
Peer	: 192.168.3.6	Origin	: incomplete
Metric	: 0	Local Pref	: 100
Weight	: 0	Calc. Local Pref	: 100
Best	: Yes	Valid	: Yes
Type	: external	Stale	: No
Originator ID	: 0.0.0.0	Path ID	: 0
Aggregator ID	:		
Aggregator AS	:		
Atomic Aggregate	:		
RFD Flaps	: 0	RFD Penalty	: 0
AS-Path	: 100		

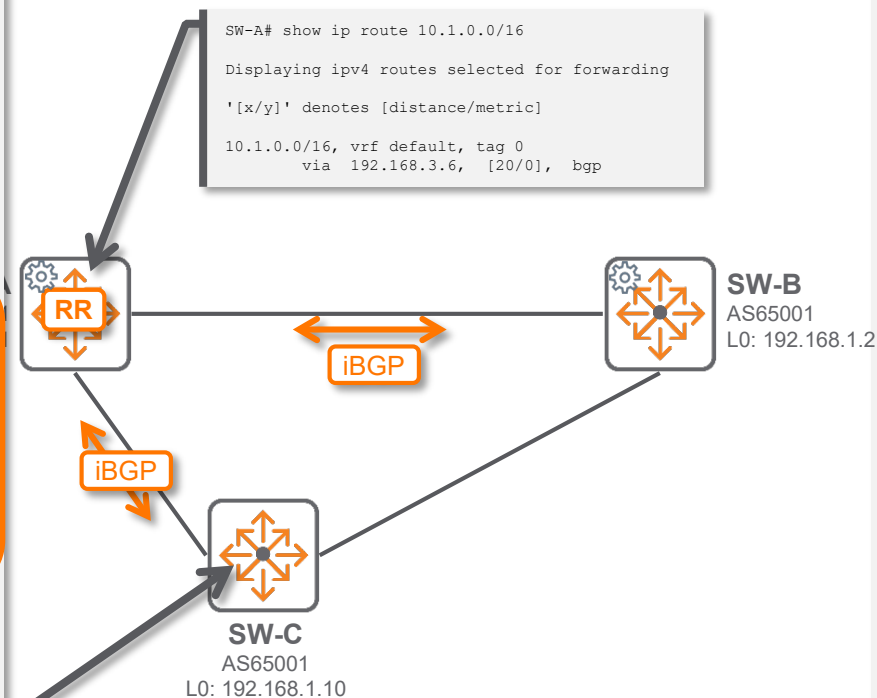
```
Cluster List :
Communities :
  Ext-Communities :
```

```
SW-A# show ip route 10.1.0.0/16
```

Displaying ipv4 routes selected for forwarding

'[x/y]' denotes [distance/metric]

```
10.1.0.0/16, vrf default, tag 0
via 192.168.3.6, [200/0], bgp
```



```
SW-C# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete
```

```
VRF : default
Local Router-ID 192.168.1.10
```

Network	Nexthop	Metric	LocPrf	Weight	Path
*>i 10.1.0.0/16	192.168.1.1	0	100	0	100 ?
*> 192.168.1.10/32	0.0.0.0	0	100	0	?

Total number of entries 2

```
SW-B# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete
```

```
VRF : default
Local Router-ID 192.168.1.2
```

Network	Nexthop	Metric	LocPrf	Weight	Path
*>i 10.1.0.0/16	192.168.1.1	0	100	0	100 ?
* e 10.1.0.0/16	192.168.3.10	0	100	0	200 100 ?
*>i 192.168.1.10/32	192.168.1.10	0	100	0	?

Total number of entries 3

```
SW-B# sh bgp ipv4 unicast 10.1.0.0/16
```

```
VRF : default
BGP Local AS 65001      BGP Router-id 192.168.1.2
```

Network	: 10.1.0.0/16	Nexthop	: 192.168.1.1
Peer	: 192.168.1.1	Origin	: incomplete
Metric	: 0	Local Pref	: 100
Weight	: 0	Calc. Local Pref	: 100
Best	: Yes	Valid	: Yes
Type	: internal	Stale	: No
Originator ID	: 0.0.0.0	Path ID	: 0
Aggregator ID	:		
Aggregator AS	:		
Atomic Aggregate	:		
RFD Flaps	: 0	RFD Penalty	: 0
AS-Path	: 100		

Network	: 10.1.0.0/16	Nexthop	: 192.168.3.10
Peer	: 192.168.3.10	Origin	: incomplete
Metric	: 0	Local Pref	: 100
Weight	: 0	Calc. Local Pref	: 100
Best	: No	Valid	: Yes
Type	: external	Stale	: No
Originator ID	: 0.0.0.0	Path ID	: 0
Aggregator ID	:		
Aggregator AS	:		
Atomic Aggregate	:		
RFD Flaps	: 0	RFD Penalty	: 0
AS-Path	: 200 100		

```
Cluster List :
Communities :
  Ext-Communities :
```

Best Practices

aruba

a Hewlett Packard
Enterprise company

ADD-PATH Best Practices

Recommendation

- For capability, use **both**: Configure **advertise and receive** additional paths for the neighbor.
- The number of candidate additional paths selected for each destination should be kept small.
- The number of “best paths” to advertise to the neighbor should be kept as default:
a value of 2 is sufficient to populate the best path and the next best backup path for each destination.

aruba

a Hewlett Packard
Enterprise company

Troubleshooting

ADD-PATH Troubleshooting

- Ensure the Capability is exchanged (Send/Receive or Both) for the intended address-family on both sender and receiver.
- Example: For the IPv4 AF neighbor capability of 'send ' and 'receive' are sent.

```
BGP Neighbor 12.0.0.2 (Internal)
Description      :
```

```
<--omitted the snippet here -->
```

```
Address Family : IPv4 Unicast
```

```
-----
Rt. Reflect. Client : No           Send Community   :
Allow-AS in        : 0             Advt. Interval    : 30
Max. Prefix         : 64000         Soft Reconfig In  :
Nexthop-Self        :              Default-Originate  :
```

```
Cfg. Add-Path      : Send and Receive
```

```
Neg. Add-Path      : Send and Receive
```

```
Routemap In        :
Routemap Out        :
ORF type            : Prefix-list
ORF capability       :
```

ADD-PATH Troubleshooting

- Sender has #n of paths expected to be advertised and all of them must have a valid NH for reachability.
- The number of advertised paths can never be more than the configured “add-path advertise-best value”.
- Only paths with associated valid reachable NH is eligible for being candidate backup path for advertisement.
- Example: ensure all the back-up paths marked with * (i.e. valid if NH is reachable)

```
RR# show bgp ipv4 unicast
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
              i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 1.1.1.1
```

	Network	Nexthop	Metric	LocPrf	Weight	Path
*>	12.0.0.0/24	0.0.0.0	0	100	0	i
*>	13.0.0.0/24	0.0.0.0	0	100	0	i
*>	14.0.0.0/24	0.0.0.0	0	100	0	i
*>i	50.0.0.0/24	12.0.0.2	0	100	0	200 i
*i	50.0.0.0/24	13.0.0.3	0	100	0	200 i
*>i	90.0.0.1/32	12.0.0.2	0	100	0	?

Total number of entries 6

ADD-PATH Troubleshooting

- Checking the **advertised** routes from the sender:
 - show bgp ipv4 unicast neighbors <nbr addr> advertised-routes

```
show bgp ipv4 unicast neighbors 14.0.0.2 advertised-routes
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
               i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete

VRF : default
Local Router-ID 1.1.1.1
```

	Network	Nexthop	Metric	LocPrf	Weight	Path
*>i	50.0.0.0/24	12.0.0.2	0	100	0	200 i
*>i	50.0.0.0/24	13.0.0.3	0	100	0	200 i
*>i	90.0.0.1/32	12.0.0.2	0	100	0	?

- Checking the **received** routes on the receiver:
 - show bgp ipv4 unicast neighbors <nbr addr> routes
 - Validity of the back-up paths(reachability of NH.)

```
DUT# show bgp ipv4 unicast neighbors 14.0.0.1 routes
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
               i internal, e external S Stale, R Removed, a additional-paths
Origin codes: i - IGP, e - EGP, ? - incomplete
```

```
VRF : default
Local Router-ID 5.5.5.5
```

	Network	Nexthop	Metric	LocPrf	Weight	Path
*>i	50.0.0.0/24	12.0.0.2	0	100	0	200 i
*a i	50.0.0.0/24	13.0.0.3	0	100	0	200 i
*>i	90.0.0.1/32	12.0.0.2	0	100	0	?

Total number of entries 6



ADD-PATH Troubleshooting

- These are the following commands available in 'diagnostic' mode.
 - **“diag bgp dump mib”** : This would take lesser time and minimum required for a dev engineer to troubleshoot further.
 - **“diag bgp dump all”** : This may be more exhaustive and take more time relative to scale. Good if you can collect this.

- Example:

```
DUT# diagnostics
```

```
DUT# diag bgp dump all
```

```
*****show bgp all-vrf all summary*****
```

```
<omitted the o/p here.>
```

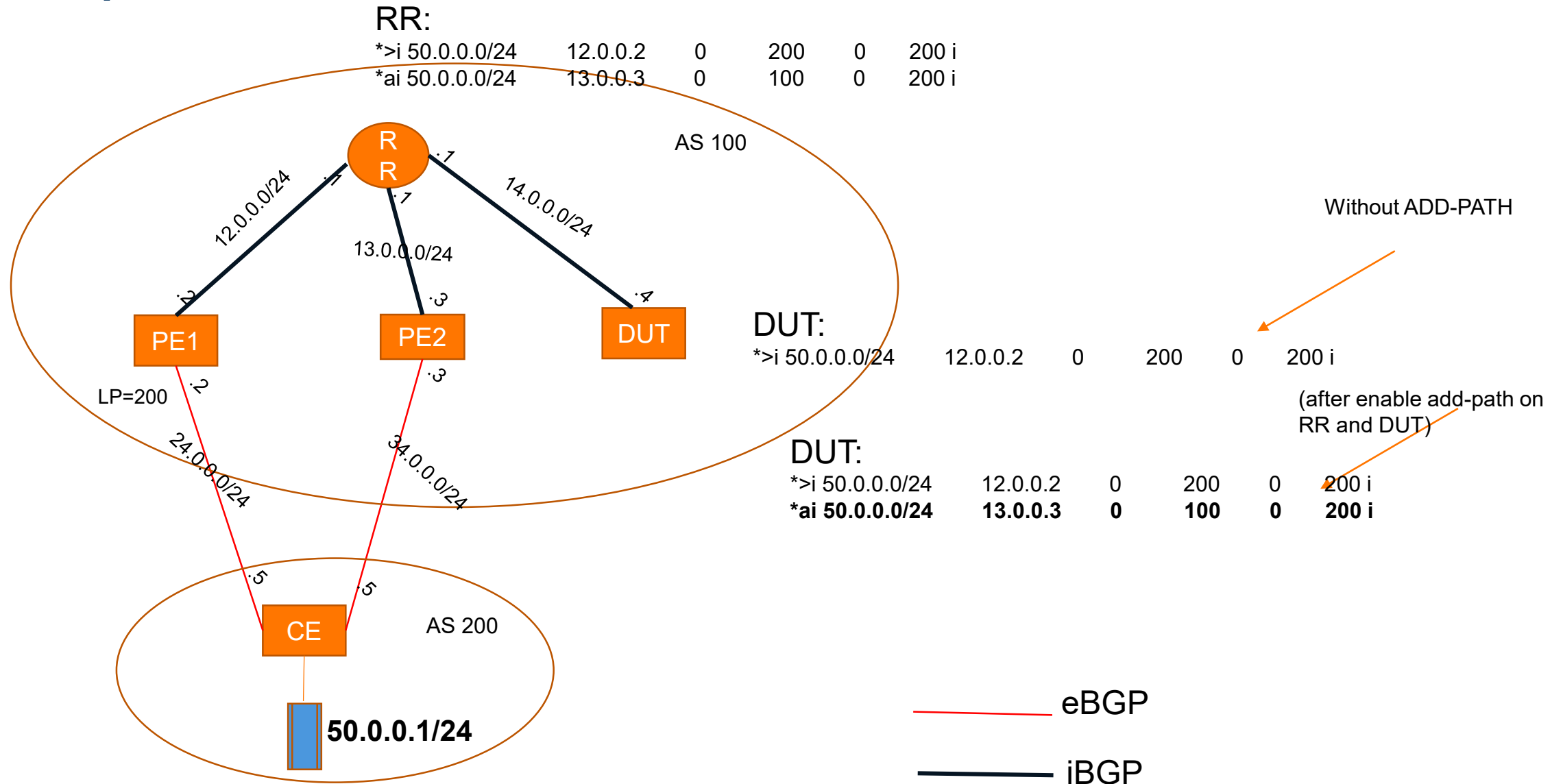
- Note: These are in addition to the supportability files.

Demo



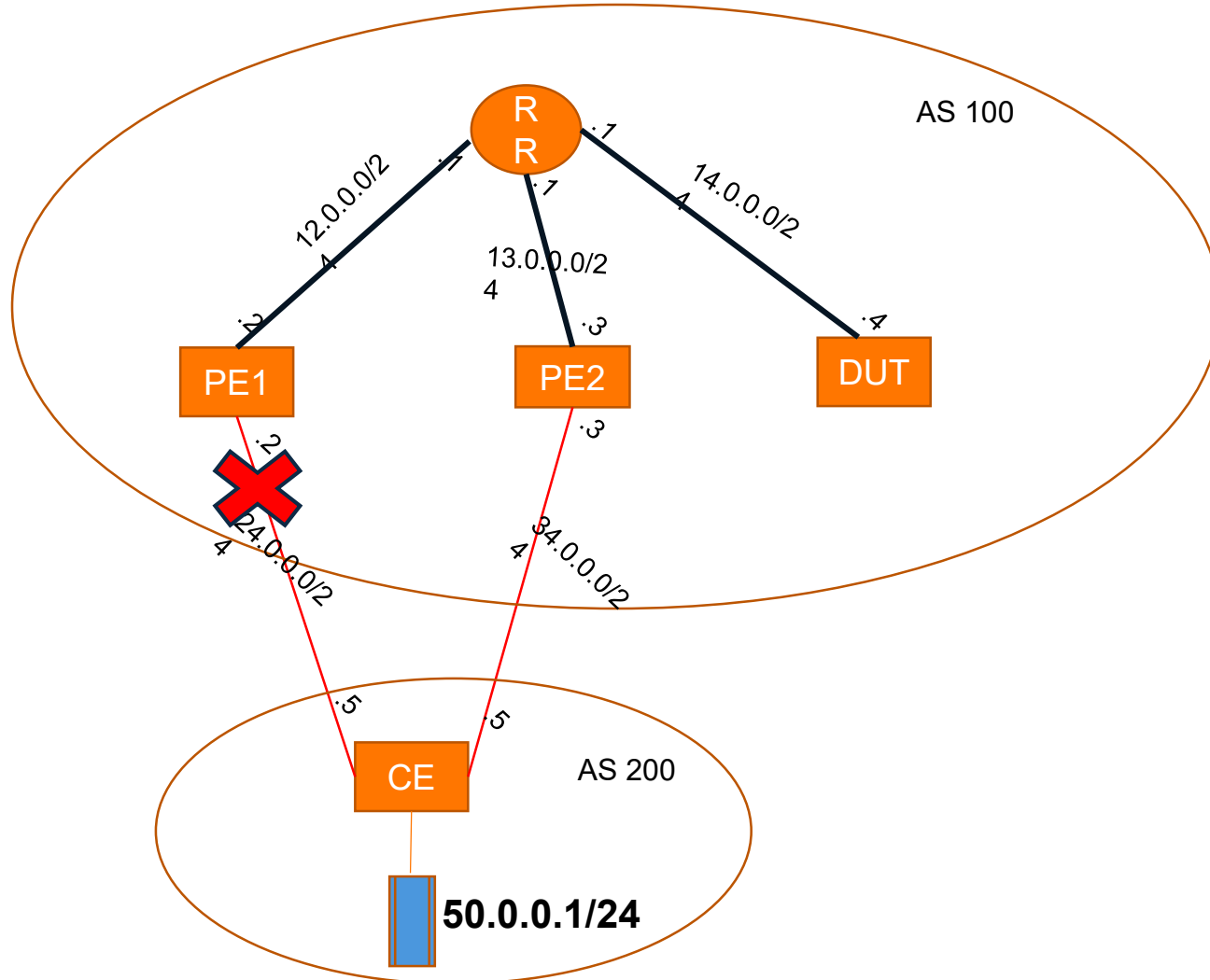
ADD-PATH demo topology

Set-up



ADD-PATH demo event

Event



Event : Shut the link between PE1 and CE
Result checked: The DUT is choosing the alternative path.

Thank you

vincent.giles@hpe.com