AOS-CX 10.6 Update
November, 2020

aruba
a Hewlett Packard Enterprise company

MP-BGP Graceful-Restart for EVPN AF

Aruba Switching TME

MP-BGP EVPN Graceful-Restart

Agenda

1. Overview
2. Use Cases
3. Details / Caveats
4. Configuration
5. Best Practices
6. Troubleshooting
7. Demo

aruba
a Hewlett Packard
Enterprise company

Overview

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)
AFI	**A**ddress **F**amily **I**dentifier
SAFI	**S**ubsequent **A**ddress **F**amily **I**dentifier
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
GR	**G**raceful **R**estart
FIB	**F**orwarding **I**nformation **B**ase
VXLAN	**V**irtual e**X**tensible **LAN**
VTEP	**V**XLAN **T**unnel **E**nd **P**oint
RIB	**R**outing **I**nformation **B**ase
EoR	**E**nd-**o**f-**R**IB
Restarting Speaker	GR **restarter** or Restarting speaker are used interchangeably
Receiving Speaker	GR **helper** or Receiving speaker are used interchangeably

Overview

Graceful-Restart

- **Graceful-restart** is a mechanism used to freeze the network data-plane while the network control-plane is being restarted, thus avoiding any interruption of network traffic.

- The **restarter** router notifies all the adjacent routers that the routing process has restarted, requesting the neighbors to not remove any FIB entries pointing to the restarter being the Next-Hop for such routes marked as **stale**.

- When the **restarter** router has completed its routing process restart, it sends a completion notification to the neighbors so that the routing protocol peering are restored to their nominal state.

Overview

10.6 Enhancement: GR support for MP-BGP EVPN AF

- **Before 10.6**, GR was supported in OSPFv2/v3 and in MP-BGP IPv4/IPv6 AF, but not for EVPN AF. This has the following consequences for EVPN routes:
  - For the switch rebooting the hpe-routing process:
    - When the process goes down, all the EVPN Routes learnt from all the VTEPs will be cleared.
    - Once BGP session is established again, EVPN routes will be learnt from all the VTEPs.
  - For the BGP neighbors:
    - When the BGP session goes down, all the EVPN Routes learnt from that restarting peer are cleared.
    - Once BGP session is established again, EVPN routes are again learnt from the peer and all the EVPN routes received on this node will be sent to the peer like during any new fresh BGP session establishment.
- **Since 10.6**, GR is also available for MP-BGP EVPN AF, removing any traffic interruption during routing process restart.

aruba
a Hewlett Packard
Enterprise company

Use Cases

Use cases

When should GR be considered?

There are 4 use-cases for Graceful-Restart function in the control-plane and particularly for OSPF/BGP:

1. **VSF Commander switch failover to Standby switch** (due to power-off or reboot of the VSF commander)

   The hpe-routing process is stopped as the VSF **Commander** reboots or is powered-off. The routing infrastructure module on the VSF **Standby** needs to move from passive to active mode in a very short time (<3-5 seconds). Right **after the process restart**, a **BGP OPEN message** is sent to all the peers **before the hold-timer** and interpreted by the peers as a GR event.

2. **Chassis (6400/8400) Management Module failover** (active MM is removed or CLI switchover)

   As the Active MM stops, the Standby MM becomes active and sends a **BGP OPEN message** to all the BGP peers **before the hold-timer** and interpreted by the peers as a GR event, to avoid any routes withdraw.

3. **hpe-routing crash handling**

   In case the hpe-routing process crashes, it is immediately restarted by **systemd** infrastructure, each routing protocol triggering GR mechanism individually. The same principle applies: this GR event prevents any network traffic interruption as the FIB is maintained on the neighbors while the associated protocol (here BGP) reload updates from the peers.

4. **ISSU** (In Service Software Upgrade)

   In the ISSU mechanism, as the data-plane is partially maintained, there is a vey high interest to use Graceful-Restart to mask the routing process restart. This use-case is not supported in 10.6.

Use cases

Clarification or Reminder on upgrade use-case

- **Graceful-Restart** does not help at all for a non-ISSU upgrade.
- GR does not have any impact or effect to the VSX update-software.
- In the VSX upgrade orchestration, the mechanism used to minimize traffic interruption is **Graceful-Shutdown**.

aruba
a Hewlett Packard
Enterprise company

Details

Platform Support

10.6 – Graceful Restart support

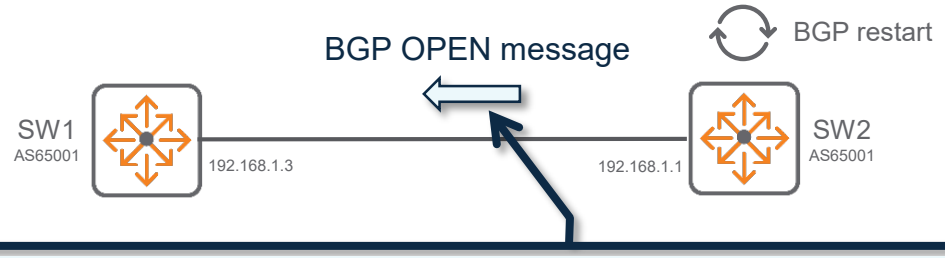
Routing Protocols	6200	6300	6400	8360	8320	8325	8400
RIP	N/A	N/A	N/A	N/A	N/A	N/A	N/A
RIPng		N/A	N/A	N/A	N/A	N/A	N/A
OSPFv2		Y	Y	Y	Y	Y	Y
OSPFv3		Y	Y	Y	Y	Y	Y
MP-BGP IPv4 AF		Y	Y	Y	Y	Y	Y
MP-BGP IPv6 AF		Y	Y	Y	Y	Y	Y
MP-BGP EVPN AF		Y	Y	Y	Y	Y	Y

- Scale:
  BGP GR “Receiving Speaker” (helper) is supported on all possible BGP neighbors (256) at the same time.

BGP GR Details

BGP OPEN message

- BGP GR is **enabled by default.**
- **During BGP session establishment,** GR capability is announced in the BGP OPEN message to all peers per AF.




```
No.     Time           Source                Destination           Protocol Length Info
      1 0.000000       192.168.1.1           192.168.1.3           BGP      131    OPEN Message

Frame 1: 131 bytes on wire (1048 bits), 131 bytes captured (1048 bits) on interface \..., id 0
Ethernet II, Src: ArubaaHe_93:ca:40 (88:3a:30:93:ca:40), Dst: HewlettP_fd:f3:00 (54:80:28:fd:f3:00)
Internet Protocol Version 4, Src: 192.168.1.1, Dst: 192.168.1.3
Transmission Control Protocol, Src Port: 43679, Dst Port: 179, Seq: 1, Ack: 1, Len: 57
Border Gateway Protocol - OPEN Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 57
    Type: OPEN Message (1)
    Version: 4
    My AS: 65001
    Hold Time: 180
    BGP Identifier: 192.168.1.1
    Optional Parameters Length: 28
    Optional Parameters
        Optional Parameter: Capability
            Parameter Type: Capability (2)
            Parameter Length: 26
            Capability: Multiprotocol extensions capability
                Type: Multiprotocol extensions capability (1)
                Length: 4
                AFI: Layer-2 VPN (25)
                Reserved: 00
                SAFI: EVPN (70)
            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: Layer-2 VPN (25)
                SAFI: EVPN (70)
                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
```

GR Restart Timer:
max time to wait for the restarter to come up

BGP GR Details

BGP AF peering


```
SW1# show bgp l2vpn evpn neighbors 192.168.1.1
Codes: ^ Inherited from peer-group

VRF : default

BGP Neighbor 192.168.1.1 (Internal)
    Description         : Spine and RR peer-group^
    Peer-group          : spine-RR

    Remote Router Id    : 192.168.1.1       Local Router Id    : 192.168.1.3
    Remote AS           : 65001             Local AS           : 65001
    Remote Port         : 43663             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:00m:19s       Alt. Local-AS      : 0
    Local-AS Prepend    : No
    BFD                 : Disabled
    Password            : 30NwaQeCi7BslzS9IQqo4wJvCQbOV/AZGPH8tBhwd+E=^
    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   :
    Weight              : 0                 Fall-over          : Yes^
    Confederation-Peers : No

    Message statistics       Sent     Rcvd
    ------------------       -----    -----
    Open                        3        2
    Notification                0        0
    Updates                    25       30
    Keepalives                 18       18
    Route Refresh               0        0
    Total                      46       50

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

    Address Family : L2VPN EVPN
    ---------------------------

    Rt. Reflect. Client : No                 Send Community    : extended^
    Allow-AS in         : 0                  Advt. Interval    : 30
    Max. Prefix         : 15000              Soft Reconfig In  :
    Nexthop-Self        :                    Default-Originate :
    Cfg. Add-Path       :
    Neg. Add-Path       :

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

GR enabled by default

GR stalepath-time:
max time to hold onto refreshing peer's routes

GR Restart Timer:
max time to wait for the restarter to come up

GR capability advertised / received

BGP GR Details

GR Mechanism

BGP initiates the Graceful-Restart mechanism when an active control-plane switchover occurs and also when acting as a GR-aware device.

- A GR-aware device, also known as **GR helper mode**, is notified that the peer router is transitioning and takes appropriate actions based on configured timers. When a BGP restart happens on the peer router, the **routes** currently held in the forwarding table are **marked** as **stale**. Thus the forwarding state is preserved as the control plane and the forwarding plane operates independently.

- On the **restarter**, BGP starts to establish sessions with all the configured peers.

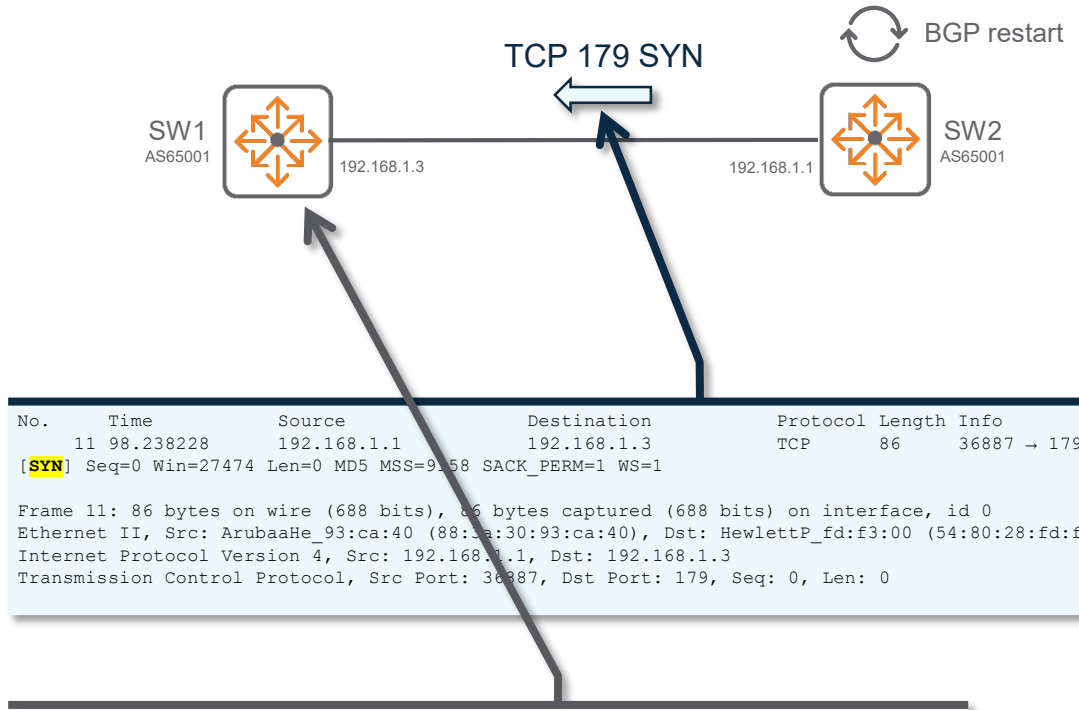
- The **BGP neighbor (helper)** sees a new TCP 179 connection request coming in while the BGP session was in an **established state**. **Such event is the indication** for the non-restarting peer **that the peer has restarted.**

The GR mechanism is triggered **AFTER** the control-plane restart.

BGP GR Details

GR trigger event

- SW1 receives a new TCP 179 connection request coming in while the BGP session was in an established state.
- With such event, SW1 detects that routing protocol on SW2 has restarted.



BGP GR Details

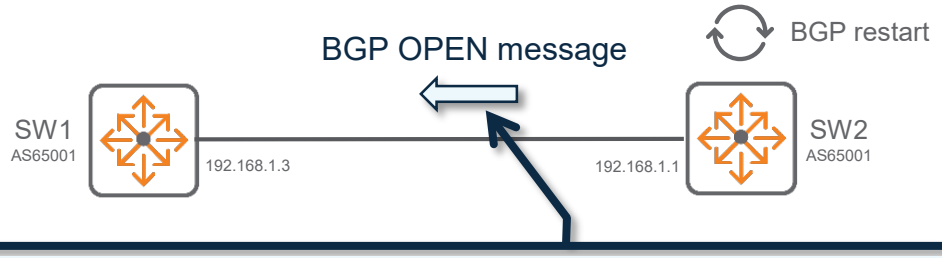
GR Mechanism

- At this point, the **restarter** sends a **BGP OPEN message** with the following GR capability bits:
  - **Restart bit** set to 1
  - **Preserve Forwarding State bit** set to 1.

BGP GR Details

BGP OPEN message with GR

- During BGP GR, the restarter sends a new BGP OPEN message to peers with the following GR capability:
  - Restart bit set to 1
  - Preserve Forwarding State bit set to 1.
  - In the context of MP-BGP EVPN:
    AFI = L2VPN
    SAFI = EVPN




```
No.     Time           Source                Destination           Protocol Length Info
     14 98.243620      192.168.1.1           192.168.1.3           BGP      131    OPEN Message

Frame 14: 131 bytes on wire (1048 bits), 131 bytes captured (1048 bits) on interface \..., id 0
Ethernet II, Src: ArubaaHe_93:ca:40 (88:3a:30:93:ca:40), Dst: HewlettP_fd:f3:00 (54:80:28:fd:f3:00)
Internet Protocol Version 4, Src: 192.168.1.1, Dst: 192.168.1.3
Transmission Control Protocol, Src Port: 36887, Dst Port: 179, Seq: 1, Ack: 58, Len: 57
Border Gateway Protocol - OPEN Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 57
    Type: OPEN Message (1)
    Version: 4
    My AS: 65001
    Hold Time: 180
    BGP Identifier: 192.168.1.1
    Optional Parameters Length: 28
    Optional Parameters
        Optional Parameter: Capability
            Parameter Type: Capability (2)
            Parameter Length: 26
            Capability: Multiprotocol extensions capability
                Type: Multiprotocol extensions capability (1)
                Length: 4
                AFI: Layer-2 VPN (25)
                Reserved: 00
                SAFI: EVPN (70)
            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: 0x8078, Restart state
                    1... .... .... .... = Restart state: Yes
                    .0.. .... .... .... = Graceful notification: No
                    .... 0000 0111 1000 = Time: 120
                AFI: Layer-2 VPN (25)
                SAFI: EVPN (70)
                Flag: 0x80, Preserve forwarding state
                    1... .... = Preserve forwarding state: Yes
            Capability: Support for 4-octet AS number capability
                Type: Support for 4-octet AS number capability (65)
                Length: 4
                AS Number: 65001
```

BGP GR Details

Effect on the BGP neighbors

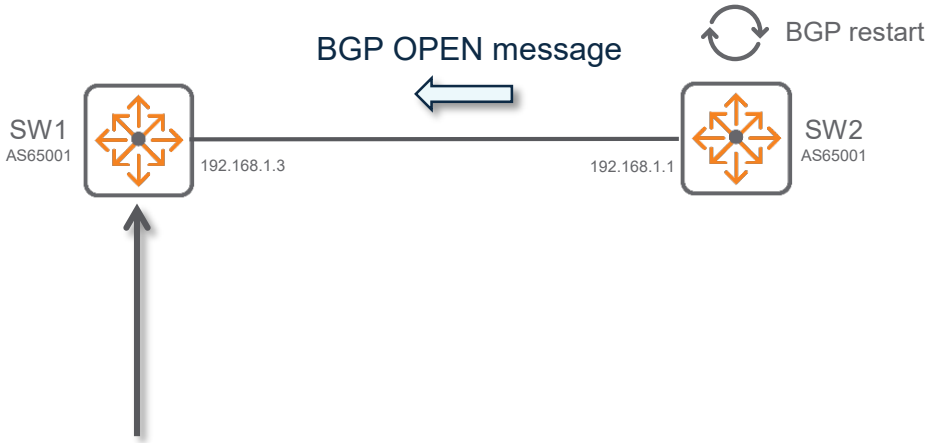
- clean up the old BGP session.
- Previous TCP 179 connection is closed.
- A new TCP socket and a new BGP session are established.


```
**************************************
Iteration : 7 Command : sh bgp l2 e s
**************************************
VRF : default
BGP Summary
-----------
 Local AS               : 65001        BGP Router Identifier  : 192.168.1.3
 Peers                  : 2            Log Neighbor Changes   : Yes
 Cfg. Hold Time         : 180          Cfg. Keep Alive        : 60
 Confederation Id       : 0

 Neighbor        Remote-AS MsgRcvd MsgSent   Up/Down Time State        AdminStatus
 192.168.1.1     65001       9950    9374    00h:19m:26s  Established   Up
 192.168.1.2     65001       9791    9246    05d:00h:12m  Established   Up


**************************************
Iteration : 8 Command : sh bgp l2 e s
**************************************
VRF : default
BGP Summary
-----------
 Local AS               : 65001        BGP Router Identifier  : 192.168.1.3
 Peers                  : 2            Log Neighbor Changes   : Yes
 Cfg. Hold Time         : 180          Cfg. Keep Alive        : 60
 Confederation Id       : 0

 Neighbor        Remote-AS MsgRcvd MsgSent   Up/Down Time State        AdminStatus
 192.168.1.1     65001       9952    9388    00h:00m:00s  Established   Up
 192.168.1.2     65001       9791    9246    05d:00h:12m  Established   Up


**************************************
Iteration : 9 Command : sh bgp l2 e s
**************************************
VRF : default
BGP Summary
-----------
 Local AS               : 65001        BGP Router Identifier  : 192.168.1.3
 Peers                  : 2            Log Neighbor Changes   : Yes
 Cfg. Hold Time         : 180          Cfg. Keep Alive        : 60
 Confederation Id       : 0

 Neighbor        Remote-AS MsgRcvd MsgSent   Up/Down Time State        AdminStatus
 192.168.1.1     65001       9952    9388    00h:00m:02s  Established   Up
 192.168.1.2     65001       9791    9246    05d:00h:12m  Established   Up
```

BGP GR Details

Effect on BGP neighbors

- clean up the old BGP sessions.



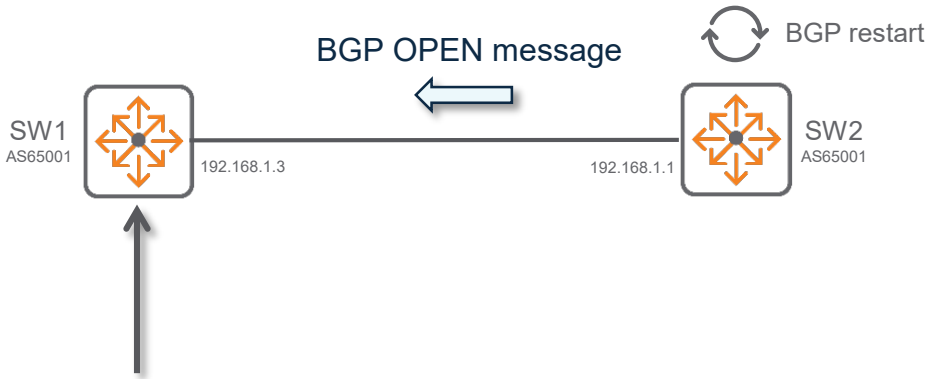

```
SW1# show events -r
--------------------------------------------------
Event logs from current boot
--------------------------------------------------
2020-10-14T16:04:27.582720+02:00 8325-1 hpe-routing[8810]: Event|2401|LOG_INFO|AMM|1/1|AdjChg: Nbr  rtr ID 192.168.1.1 on IP addr 192.168.3.1( area ID 0.0.0.0): Exchange -> Full
2020-10-14T16:04:27.581470+02:00 8325-1 hpe-routing[8810]: Event|2401|LOG_INFO|AMM|1/1|AdjChg: Nbr  rtr ID 192.168.1.1 on IP addr 192.168.3.1( area ID 0.0.0.0): Exstart -> Exchange
2020-10-14T16:04:27.580725+02:00 8325-1 hpe-routing[8810]: Event|2401|LOG_INFO|AMM|1/1|AdjChg: Nbr  rtr ID 192.168.1.1 on IP addr 192.168.3.1( area ID 0.0.0.0): Two-way -> Exstart
2020-10-14T16:04:27.580467+02:00 8325-1 hpe-routing[8810]: Event|2401|LOG_INFO|AMM|1/1|AdjChg: Nbr  rtr ID 192.168.1.1 on IP addr 192.168.3.1( area ID 0.0.0.0): Init -> Two-way
2020-10-14T16:04:25.189446+02:00 8325-1 hpe-routing[8810]: Event|2401|LOG_INFO|AMM|1/1|AdjChg: Nbr  rtr ID 192.168.1.1 on IP addr 192.168.3.1( area ID 0.0.0.0): Full -> Init
2020-10-14T16:04:15.187427+02:00 8325-1 hpe-routing[8810]: Event|2901|LOG_INFO|AMM|1/1|192.168.1.1: Peer up. vrf-name: default
2020-10-14T16:04:14.238694+02:00 8325-1 hpe-routing[8810]: Event|2902|LOG_INFO|AMM|1/1|192.168.1.1: Peer down. error-code: Unrecognized error code, error-sub-code: Unrecognized error subcode. vrf-
name: default
```

The error-code is not very explicit.
Improvement is being investigated.

BGP GR Details

Effect on BGP neighbors

- Mark all the routes that were received from the restarting peer as **stale** in the BGP table.
- No impact on the FIB as the routes are maintained.



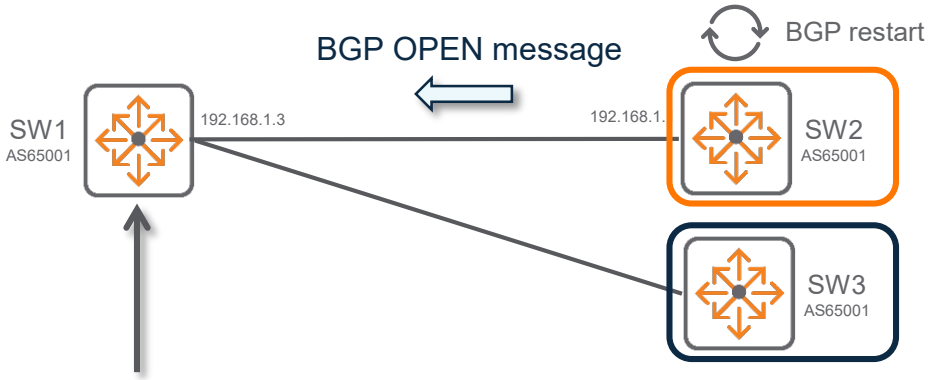
Stale

```
SW1# show bgp l2vpn evpn neighbor 192.168.1.1 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

EVPN Route-Type 2 prefix: [2]:[ESI]:[EthTag]:[MAC]:[OrigIP]
EVPN Route-Type 3 prefix: [3]:[EthTag]:[OrigIP]
EVPN Route-Type 5 prefix: [5]:[ESI]:[EthTag]:[IPAddrLen]:[IPAddr]
VRF : default
Local Router-ID 192.168.1.3

     Network                                                Nexthop         Metric     LocPrf     Weight    Path
-------------------------------------------------------------------------------------------------------------------
Route Distinguisher: 192.168.11.5:10       (L2VNI 10010)
S>i [2]:[0]:[0]:[12:00:00:00:01:00]:[10.1.10.1]           192.168.11.5    0          100        0       ?
S>i [2]:[0]:[0]:[12:00:00:00:01:00]:[fe80:10:1:10::1]     192.168.11.5    0          100        0       ?
S>i [3]:[0]:[192.168.11.5]                                192.168.11.5    0          100        0       ?

Route Distinguisher: 192.168.11.5:12       (L2VNI 10012)
S>i [2]:[0]:[0]:[00:50:56:8e:68:5e]:[10.1.12.11]          192.168.11.5    0          100        0       ?
S>i [2]:[0]:[0]:[00:50:56:8e:68:5e]:[]                    192.168.11.5    0          100        0       ?
S>i [2]:[0]:[0]:[00:50:56:8e:68:5e]:[fd00:10:1:12:2489:3e32:bafe:ebbd]
                                                          192.168.11.5    0          100        0       ?
S>i [2]:[0]:[0]:[00:50:56:8e:68:5e]:[fe80::80a0:5d8b:f992:cb1d]
                                                          192.168.11.5    0          100        0       ?
S>i [2]:[0]:[0]:[12:00:00:00:01:00]:[10.1.12.1]           192.168.11.5    0          100        0       ?
S>i [2]:[0]:[0]:[12:00:00:00:01:00]:[fe80:10:1:12::1]     192.168.11.5    0          100        0       ?
S>i [3]:[0]:[192.168.11.5]                                192.168.11.5    0          100        0       ?

Route Distinguisher: 192.168.11.5:20       (L2VNI 10020)
S>i [2]:[0]:[0]:[00:50:56:8e:6b:d8]:[10.2.20.11]          192.168.11.5    0          100        0       ?
S>i [2]:[0]:[0]:[00:50:56:8e:6b:d8]:[]                    192.168.11.5    0          100        0       ?
S>i [2]:[0]:[0]:[12:00:00:00:01:00]:[10.2.20.1]           192.168.11.5    0          100        0       ?
S>i [3]:[0]:[192.168.11.5]                                192.168.11.5    0          100        0       ?

Route Distinguisher: 192.168.1.4:1         (L3VNI 100001)
S>i [5]:[0]:[0]:[0]:[0.0.0.0]                             192.168.11.3    0          100        0       ?
S>i [5]:[0]:[0]:[24]:[10.1.10.0]                          192.168.11.3    0          100        0       ?
S>i [5]:[0]:[0]:[24]:[10.1.11.0]                          192.168.11.3    0          100        0       ?
S>i [5]:[0]:[0]:[64]:[fd00:10:1:10::]                     192.168.11.3    0          100        0       ?

Route Distinguisher: 192.168.1.5:1         (L3VNI 100001)
<.. Omitted ..>
```

BGP GR Details

Effect on BGP neighbors

- Routes from other peers are not impacted and still valid.




```
SW1# show bgp l2vpn evpn
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

EVPN Route-Type 2 prefix: [2]:[ESI]:[EthTag]:[MAC]:[OrigIP]
EVPN Route-Type 3 prefix: [3]:[EthTag]:[OrigIP]
EVPN Route-Type 5 prefix: [5]:[ESI]:[EthTag]:[IPAddrLen]:[IPAddr]
VRF : default
Local Router-ID 192.168.1.3

     Network                                                Nexthop         Metric     LocPrf     Weight    Path
-------------------------------------------------------------------------------------------------------------------
Route Distinguisher: 192.168.11.3:10       (L2VNI 10010)
*>   [2]:[0]:[0]:[00:50:56:8e:cf:69]:[10.1.10.12]           192.168.11.3    0          100        0       ?
*>   [2]:[0]:[0]:[00:50:56:8e:cf:69]:[]                     192.168.11.3    0          100        0       ?
*>   [2]:[0]:[0]:[00:50:56:8e:cf:69]:[fe80::f4af:f785:51cb:403c]
                                                            192.168.11.3    0          100        0       ?
*>   [2]:[0]:[0]:[12:00:00:00:01:00]:[10.1.10.1]            192.168.11.3    0          100        0       ?
*>   [2]:[0]:[0]:[12:00:00:00:01:00]:[fe80:10:1:10::1]      192.168.11.3    0          100        0       ?
*>   [3]:[0]:[192.168.11.3]                                 192.168.11.3    0          100        0       ?

Route Distinguisher: 192.168.11.5:10       (L2VNI 10010)
S>i  [2]:[0]:[0]:[12:00:00:00:01:00]:[10.1.10.1]            192.168.11.5    0          100        0       ?
*  i  [2]:[0]:[0]:[12:00:00:00:01:00]:[10.1.10.1]            192.168.11.5    0          100        0       ?
S>i  [2]:[0]:[0]:[12:00:00:00:01:00]:[fe80:10:1:10::1]      192.168.11.5    0          100        0       ?
* i  [2]:[0]:[0]:[12:00:00:00:01:00]:[fe80:10:1:10::1]      192.168.11.5    0          100        0       ?
S>i  [3]:[0]:[192.168.11.5]                                 192.168.11.5    0          100        0       ?
* i  [3]:[0]:[192.168.11.5]                                 192.168.11.5    0          100        0       ?

Route Distinguisher: 192.168.11.3:11       (L2VNI 10011)
*>   [2]:[0]:[0]:[00:50:56:8e:4d:9c]:[10.1.11.10]           192.168.11.3    0          100        0       ?
*>   [2]:[0]:[0]:[00:50:56:8e:4d:9c]:[]                     192.168.11.3    0          100        0       ?
*>   [2]:[0]:[0]:[12:00:00:00:01:00]:[10.1.11.1]            192.168.11.3    0          100        0       ?
*>   [3]:[0]:[192.168.11.3]                                 192.168.11.3    0          100        0       ?

Route Distinguisher: 192.168.11.5:12       (L2VNI 10012)
S>i  [2]:[0]:[0]:[00:50:56:8e:68:5e]:[10.1.12.11]           192.168.11.5    0          100        0       ?
* i  [2]:[0]:[0]:[00:50:56:8e:68:5e]:[10.1.12.11]           192.168.11.5    0          100        0       ?
S>i  [2]:[0]:[0]:[00:50:56:8e:68:5e]:[]                     192.168.11.5    0          100        0       ?
* i  [2]:[0]:[0]:[00:50:56:8e:68:5e]:[]                     192.168.11.5    0          100        0       ?
<.. Omitted ..>
```

Stale

Valid

BGP GR Details

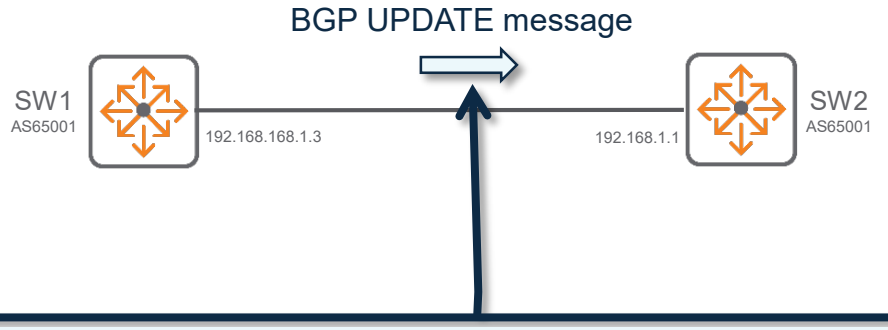
Neighbor: Routing table update and EoR marker

- Sends to the restarter an initial routing table update, followed by an **End-of-RIB (EoR) Marker**

```
SW1# diag utilities tcpdump verbosity level4 destination-ip 192.168.1.1
tcpdump: listening on any, link-type LINUX_SLL (Linux cooked), capture size 262144 bytes
    1  17:53:00.722820 IP (tos 0xc0, ttl 64, id 32860, offset 0, flags [DF], proto TCP (6), length 60)
    8325-1.bgp > 192.168.1.1.40919: Flags [.], cksum 0x7446 (correct), seq 1479401499, ack 2474042295, win 391, options [nop,nop,md5 shared secret not supplied with -M, can't check -
0088403cbfee8ab23c0d6c1cfad1d427], length 0
    2  17:53:39.252732 IP (tos 0xc0, ttl 64, id 32861, offset 0, flags [DF], proto TCP (6), length 60)
    8325-1.bgp > 192.168.1.1.40919: Flags [F.], cksum 0x7d71 (correct), seq 0, ack 2, win 391, options [nop,nop,md5 shared secret not supplied with -M, can't check -
80c3e5d206565379e8c66fabba58262b], length 0
    3  17:53:40.192431 IP (tos 0xc0, ttl 64, id 0, offset 0, flags [DF], proto TCP (6), length 72)
    8325-1.bgp > 192.168.1.1.42323: Flags [S.], cksum 0x2820 (correct), seq 1786372186, ack 3367608433, win 27474, options [nop,nop,md5 shared secret not supplied with -M, can't check -
23c399c0488edf20110c291e22f639a7,mss 9158,nop,nop,sackOK,nop,wscale 7], length 0
    4  17:53:40.193170 IP (tos 0xc0, ttl 64, id 44948, offset 0, flags [DF], proto TCP (6), length 117)
    8325-1.bgp > 192.168.1.1.42323: Flags [P.], cksum 0x75d8 (correct), seq 1:58, ack 1, win 215, options [nop,nop,md5 shared secret not supplied with -M, can't check -
d629a53ff1f09bf9d3cfd9d9c06d9edd], length 57: BGP
        Open Message (1), length: 57
          Version 4, my AS 65001, Holdtime 180s, ID 8325-1
          Optional parameters, length: 28
            Option Capabilities Advertisement (2), length: 26
              Multiprotocol Extensions (1), length: 4
                AFI VPLS (25), SAFI Unknown (70)
                0x0000:  0019 0046
              Route Refresh (2), length: 0
              Route Refresh (Cisco) (128), length: 0
              Enhanced Route Refresh (70), length: 0
                no decoder for Capability 70
              Graceful Restart (64), length: 6
                Restart Flags: [none], Restart Time 120s
                  AFI VPLS (25), SAFI Unknown (70), Forwarding state preserved: no
                0x0000:  0078 0019 4600
              32-Bit AS Number (65), length: 4
                 4 Byte AS 65001
                0x0000:  0000 fde9

<- output omitted see next hidden slide for details->


        Update Message (2), length: 30
          Multi-Protocol Unreach NLRI (15), length: 3, Flags [OE]:
            AFI: VPLS (25), SAFI: Unknown SAFI (70)
              End-of-Rib Marker (empty NLRI)
            0x0000:  0019 46
```

BGP GR Details

Neighbor: Routing table update and EoR marker



- For the **EVPN address family**, an UPDATE message with no reachable Network Layer Reachability Information (NLRI) and empty withdrawn NLRI is specified as the **End-of-RIB marker** that can be used by a BGP speaker to indicate to its peer the **completion of the initial routing update** after the session is established.

- For the **IPv4/IPv6 unicast address family**, the End-of-RIB marker is an UPDATE message with the minimum length. For any **other address family**, it is an UPDATE message that contains only the **MP_UNREACH_NLRI** attribute with **no withdrawn routes** for that <AFI, SAFI>.

```
No.     Time           Source                Destination           Protocol Length Info
      1 0.000000000    192.168.1.3           192.168.1.1           BGP      1996   UPDATE Message, UPDATE
Message, UPDATE Message, UPDATE Message, UPDATE Message, UPDATE Message, UPDATE Message, UPDATE Message,
UPDATE Message, UPDATE Message, UPDATE Message

Frame 1: 1996 bytes on wire (15968 bits), 1996 bytes captured (15968 bits) on interface MirrorRxNet, id 0
Ethernet II, Src: HewlettP_fd:f3:00 (54:80:28:fd:f3:00), Dst: ArubaaHe_93:ca:40 (88:3a:30:93:ca:40)
Internet Protocol Version 4, Src: 192.168.1.3, Dst: 192.168.1.1
Transmission Control Protocol, Src Port: 179, Dst Port: 41633, Seq: 1, Ack: 1, Len: 1922
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - UPDATE Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 30
    Type: UPDATE Message (2)
    Withdrawn Routes Length: 0
    Total Path Attribute Length: 7
    Path attributes
        Path Attribute - MP_UNREACH_NLRI
            Flags: 0x90, Optional, Extended-Length, Non-transitive, Complete
                1... .... = Optional: Set
                .0.. .... = Transitive: Not set
                ..0. .... = Partial: Not set
                ...1 .... = Extended-Length: Set
                .... 0000 = Unused: 0x0
            Type Code: MP_UNREACH_NLRI (15)
            Length: 3
            Address family identifier (AFI): Layer-2 VPN (25)
            Subsequent address family identifier (SAFI): EVPN (70)
            Withdrawn routes (0 bytes)
```

BGP GR Details

Neighbor: stale management

- Under normal conditions, **Stale** flag of routes will be removed if the corresponding routes update is received from the Restarting Speaker.
- The neighbor will purge all stale routes after the **Restart-Time** (120s by default) expires if the restarting peer does not re-establish the BGP session before the expiration of this restart-timer.
- The neighbor will also purge all stale routes after the **Stalepath-Time** (300s by default) expires if the restarting peer does not send routes updates before the expiration of this stalepath-timer.
- Both GR timers are started upon reception of the BGP OPEN message with GR bits set.

BGP GR Details

GR Process completion

- On the **restarter**:
  - Delay best-path calculation until after receiving End-of-RIB marker from all peers.
  - Generate updates for its peers and send its complete BGP routing table followed by the End-of-RIB marker.

- On the **BGP neighbors**:
  - Receive the routing updates from the restarting switch.
  - Remove stale marking for any refreshed routes.
  - Purge any remaining stale routes after End-of-RIB marker is received from the restarting peer or when the stalepath timer expires.

VXLAN Tunnels

No EVPN impact

- If the **Restarting Speaker** is a **VTEP**, the associated VXLAN tunnel pointers are kept unchanged in the ASIC for data-plane continuity:
  - Although protocols are restarted, the OVSDB Route table is kept unchanged for overlay and underlay.
  - While Host (Route-type 2) EVPN routes are stale, the associated MAC-address table is maintained.

aruba
a Hewlett Packard
Enterprise company

Configuration

BGP GR configuration

- BGP GR is **enabled by default and can not be disabled (specifically for BGP).**
  GR can however be globally disabled, that would be for all protocols (not only for BGP).
- MP-BGP EVPN GR timers are configured in the default VRF global BGP configuration part.
- Two BGP GR timers can be configured:
  1. Restart-time: defining the maximum time to wait for the restarter to come back
  2. Stalepath-time: defining the maximum time to hold onto refreshing the peer's routes
- Change to timers will reset the BGP sessions in the given VRF (here the default).

```
SW1(config-bgp)# bgp graceful-restart
  restart-time     Set the max time to wait for graceful restart peer to come up
  stalepath-time   Set the max time to hold onto restarting peer's stale paths

SW1(config-bgp)# bgp graceful-restart restart-time
  <1-3600>  Delay value (Default: 120 seconds)
SW1(config-bgp)# bgp graceful-restart restart-time 120
  <cr>
SW1(config-bgp)# bgp graceful-restart restart-time 120
All current BGP sessions in VRF default will be restarted
Do you want to continue (y/n)? N

SW1(config-bgp)# bgp graceful-restart stalepath-time
  <1-3600>  Delay value (Default: 300 seconds)
SW1(config-bgp)# bgp graceful-restart stalepath-time 300
All current BGP sessions in VRF default will be restarted
Do you want to continue (y/n)? n
SW1(config-bgp)#
```

aruba
a Hewlett Packard
Enterprise company

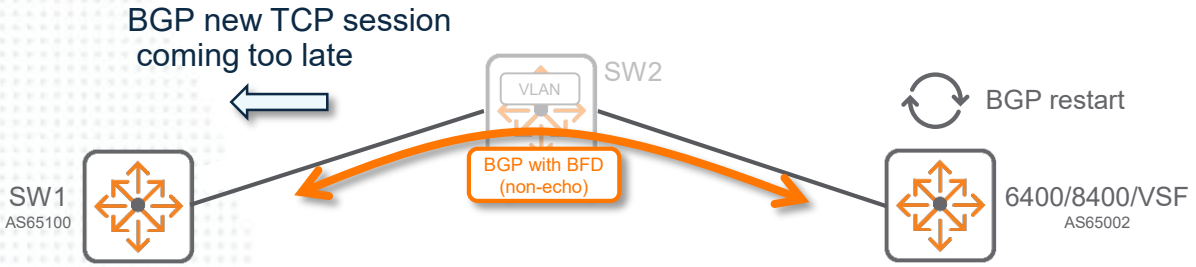
Best Practices

BGP GR Best Practices

- **BGP Graceful-Restart is enabled by default** (and can not be disabled specifically for BGP).
  This provides the benefit of seamless restart of hpe-routing process, and is useful for this reason even if the switch is not a chassis with dual MM or a VSF stack.

- **GR and BFD combination**
  - HW BFD and BGP GR combination works well for high-availability solution.
  - SW BFD and BGP GR combination might require special attention and testing to guarantee high-availability.
    BFD might detect neighborhood failure before the GR mechanism is kicked in.

    Note: BFD is useless for direct point-to-point connection supporting the BGP peering.



aruba
a Hewlett Packard
Enterprise company

Troubleshooting

EVPN GR Troubleshooting

BGP Debug

- BGP sessions must be in **Established** state after GR.

```
leaf1# show bgp l2vpn evpn sum
VRF : default
BGP Summary
-----------
 Local AS               : 1           BGP Router Identifier  : 192.168.1.24
 Peers                  : 1           Log Neighbor Changes   : No
 Cfg. Hold Time         : 180         Cfg. Keep Alive        : 60
 Confederation Id       : 0

 Neighbor        Remote-AS MsgRcvd MsgSent   Up/Down Time State        AdminStatus
 192.168.1.22    1           23      23      00h:00m:30s  Established   Up

leaf1#
```

EVPN GR Troubleshooting

BGP Debug

```
2020-10-09:16:21:55.523702|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Received Capability Multiprotocol Extension, AFI/SAFI 25/70 in OPEN message.VRF Name = default.
2020-10-09:16:21:55.523754|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Received Capability Route Refresh in OPEN message.VRF Name = default.
2020-10-09:16:21:55.523801|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Received Capability Cisco Route Refresh in OPEN message.VRF Name = default.
2020-10-09:16:21:55.523847|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Received Capability Enhanced Route Refresh in OPEN message.VRF Name = default.
2020-10-09:16:21:55.523908|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Received Capability Graceful Restart in OPEN message.VRF Name = default.
2020-10-09:16:21:55.523957|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Received Capability Four-octet AS in OPEN message.VRF Name = default.
2020-10-09:16:21:55.524152|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|A valid OPEN message has been received from a neighbor.
2020-10-09:16:21:55.524165|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|NM entity index          = 268763136
2020-10-09:16:21:55.524176|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Configured local address = 192.168.1.3
2020-10-09:16:21:55.524188|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Configured local port    = 0
2020-10-09:16:21:55.524199|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Selected local address   = 192.168.1.3
2020-10-09:16:21:55.524210|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Selected local port      = 179
2020-10-09:16:21:55.524221|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Remote address          = 192.168.1.1
2020-10-09:16:21:55.524233|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Remote port             = 0
2020-10-09:16:21:55.524244|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Scope ID                = 0
2020-10-09:16:21:55.524255|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Incoming connection?    = True
2020-10-09:16:21:55.524266|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Remote AS number        = 65001
2020-10-09:16:21:55.524277|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Remote BGP ID           = 0.0.0.0
2020-10-09:16:21:55.524288|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Received hold time      = 180
2020-10-09:16:21:55.524299|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Number of Capabilities  = 6
2020-10-09:16:21:55.524309|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Restart capable?        = True
2020-10-09:16:21:55.524320|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Capabilities offset     = 31
2020-10-09:16:21:55.524331|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Open Message packet     =
2020-10-09:16:21:55.524344|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|   FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF 00390104 FDE900B4 C0A80101 1C021A01
2020-10-09:16:21:55.524355|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|   04001900 46020080 00460040 06807800 19468041 040000FD E9VRF Name = default.
2020-10-09:16:21:55.528223|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|A connection has entered Established state.
2020-10-09:16:21:55.528238|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Local address:  192.168.1.3
2020-10-09:16:21:55.528252|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Local port:     0
2020-10-09:16:21:55.528263|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Remote address: 192.168.1.1
2020-10-09:16:21:55.528274|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Remote port:    0
2020-10-09:16:21:55.528284|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Scope ID:       0
2020-10-09:16:21:55.528295|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Neg hold time:  180
2020-10-09:16:21:55.528306|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Passive:?       False
2020-10-09:16:21:55.528319|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Entity index:   268763136VRF Name = default.
2020-10-09:16:21:55.528454|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|BGP 268763136 established a session with peer 192.168.1.1.
2020-10-09:16:21:55.528465|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Configured local address and port:  192.168.1.3:0
2020-10-09:16:21:55.528476|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Connection local address and port:  192.168.1.3:179
2020-10-09:16:21:55.528487|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Configured remote address and port: 192.168.1.1:0
2020-10-09:16:21:55.528498|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Connection remote port:            45171
2020-10-09:16:21:55.528510|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Scope ID:                        0
2020-10-09:16:21:55.528521|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Local AS number:                  65001
2020-10-09:16:21:55.528533|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|Remote AS number:                 65001VRF Name = default.
2020-10-09:16:21:55.528681|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|A BGP peer session has restarted.
2020-10-09:16:21:55.528693|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|RIB Manager entity index: 268763136
2020-10-09:16:21:55.528704|hpe-routing|LOG_INFO|AMM|-|BGP|BGP|BGP peer ID:             3232235777VRF Name = default.
2020-10-09:16:21:55.530523|hpe-routing|LOG_INFO|AMM|-|BGP|BGP_EVENT|VRF default: Peer 192.168.1.1 state changed to Established.
2020-10-09:16:21:55.563003|hpe-routing|LOG_ERR|AMM|-|OSPFV2|OSPFv2_EVENT|Could not send ospfNbrRestartHelperStatusChange Trap
```

EVPN GR Troubleshooting

- Underlay Route must be unchanged. Deletion of underlay routes leads to deletion of tunnels and subsequently IP routes/MAC gets deleted. This can further cause traffic disruption.

```
leaf1# show ip route
snip
20.20.20.0/24, vrf default
        via  10.10.10.2,  [110/20],  ospf
192.168.1.0/24, vrf default
        via  loopback1,  [0/0],  connected
192.168.1.21/32, vrf default
        via  10.10.10.2,  [110/20],  ospf
192.168.1.22/32, vrf default
        via  10.10.10.2,  [110/10],  ospf
192.168.1.24/32, vrf default
        via  loopback1,  [0/0],  local

leaf1#
```

EVPN GR Troubleshooting

After GR, check that the EVPN routes are received via “show bgp l2vpn evpn” command.

```
leaf1# show bgp l2vpn evpn
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

EVPN Route-Type 2 prefix: [2]:[ESI]:[EthTag]:[MAC]:[OrigIP]
EVPN Route-Type 3 prefix: [3]:[EthTag]:[OrigIP]
EVPN Route-Type 5 prefix: [5]:[ESI]:[EthTag]:[IPAddrLen]:[IPAddr]
VRF : default
Local Router-ID 192.168.1.24

     Network                                          Nexthop         Metric     LocPrf     Weight   Path
-------------------------------------------------------------------------------------------------------
Route Distinguisher: 192.168.1.21:10      (L2VNI 10)
*>i [2]:[0]:[0]:[00:00:01:00:01:11]:[100.10.0.1]          192.168.1.21    0          100        0       ?
*>i [2]:[0]:[0]:[00:50:56:96:62:4f]:[100.10.0.12]         192.168.1.21    0          100        0       ?
*>  [2]:[0]:[0]:[00:50:56:96:62:4f]:[]                   192.168.1.21    0          100        0       ?
*>i [3]:[0]:[192.168.1.21]                                192.168.1.21    0          100        0       ?

Route Distinguisher: 192.168.1.24:10      (L2VNI 10)
*>  [2]:[0]:[0]:[00:00:01:00:01:11]:[100.10.0.1]          192.168.1.24    0          100        0       ?
*>  [2]:[0]:[0]:[00:50:56:96:d7:dd]:[100.10.0.13]         192.168.1.24    0          100        0       ?
*>  [2]:[0]:[0]:[00:50:56:96:d7:dd]:[]                   192.168.1.24    0          100        0       ?
*>  [3]:[0]:[192.168.1.24]                                192.168.1.24    0          100        0       ?

Route Distinguisher: 192.168.1.21:20      (L2VNI 20)
*>i [2]:[0]:[0]:[00:00:01:00:01:11]:[100.30.0.1]          192.168.1.21    0          100        0       ?
*>i [2]:[0]:[0]:[00:50:56:96:94:3f]:[100.30.0.11]         192.168.1.21    0          100        0       ?
*>i [2]:[0]:[0]:[00:50:56:96:94:3f]:[]                    192.168.1.21    0          100        0       ?
*>i [3]:[0]:[192.168.1.21]                                192.168.1.21    0          100        0       ?

Route Distinguisher: 192.168.1.24:20      (L2VNI 20)
```

EVPN GR Troubleshooting

ovs-appctl -t hpe-routing fastlog show evpn_dump | grep stale

- Tunnels/ EVPN MACs and Routes [Host/Prefix] gets deleted as part of stale management if the corresponding route is not received.

```
 leaf1:/home/admin# ovs-appctl -t hpe-routing fastlog show evpn_dump | grep stale
(23 Oct 20 11:37:20.181204251): INFO: ms_idl_bd_stale_all_mac_ip_bindings:80: Stale entry marking completed for all OVSDB tables
(23 Oct 20 11:37:20.181206777): INFO: evpn_gr_start_stale_timer:90: EVPN GR Timer Started for 300 seconds
(23 Oct 20 11:42:20.181228658): INFO: evpn_gr_timer_handle_stale_timer_expiry:139: EVPN GR Stale timer expired, clear stale entries
leaf1:/home/admin#
leaf1:/home/admin#
```

As of now, there are no stale entries which are deleted, but this is a place to look out for

- Tunnels/EVPN MACs and Routes [Host/Prefix] must not get deleted on remote VTEP as well. For example, GR on VTEP1 must not in any way have a change on VTEP2.
- If the Tunnels/ EVPN MACs and Routes [Host/Prefix] are getting deleted as part of stale management, increasing the stalepath timer depending on the scale may help.

EVPN GR Troubleshooting

ovsdb-client dump Route

```
8325-1:/home/admin# ovsdb-client dump Route
Route table
_uuid                                address_family                     coalescence_id                   distance                         dp_state                                ecmp_group
from                                 l3_destination                     metric                           nexthops                         prefix                                  protocol_private
selected                             source_vrfs                        sub_address_family               tag                              type                                    vrf
------------------------------------ ---------------------------------- -------------------------------- -------------------------------- --------------------------------------- --------------------
---------------- ------------------------------------ ---------------------------------- -------------------------------- -------------------------------- --------------------------------------- -
-------------------------------- ------------------------------------ ---------------------------------- -------------------------------- -------------------------------- --------------------------------------- --------------------
---------------- ------------------------------------
d6a8b3ac-aaee-4cd6-858f-7ef6173463d4 "ipv4"        181027733     0       []               []                                  local    []                                      0      [dd8d6e97-655b-
4506-9cad-07b1730a59ad]                          "10.2.20.1/32"                   []               true    []                                      []                     0   local
95329f87-3a6b-49e5-add3-ec4cf7d2a631
110e773b-0d14-4d31-b473-1d4eedf19e40 "ipv4"        226381135     200     single           []                                  bgp      abd89474-7e2a-496d-9ebd-0e3eb8847c42 0      [8610f0a6-6fd0-
402f-b31a-e424c609a9bb]                          "10.1.155.2/31"                  []               true    [e5770645-53ab-4897-bd41-3e67248520f8] []                     0   forward
5f7b72c9-91d3-402f-b963-f5c046f83d0d
fc33d242-16be-4099-9f48-bd61a363ef0e "ipv4"        369364506     200     host_route_single []                                 bgp      8e437c2d-15a5-4139-bdb1-6615eb86d37b 0      [0a9a1904-e504-
42ba-8e2b-713f0f70f464]                          "10.2.20.11/32"                  []               true    []                                      []                     0   forward
95329f87-3a6b-49e5-add3-ec4cf7d2a631
7578a020-cd50-4a5e-adff-25f87d971a96 "ipv4"        401616467     200     single           []                                  bgp      abd89474-7e2a-496d-9ebd-0e3eb8847c42 0      [8610f0a6-6fd0-
402f-b31a-e424c609a9bb]                          "10.1.155.2/31"                  []               true    []                                      []                     0   forward
e5770645-53ab-4897-bd41-3e67248520f8
5efeaf5e-7840-4d31-95fb-29752407ddef "ipv4"        474035719     110     single           []                                  ospf     e77239e8-dc8b-4c54-9efd-d63536241377 14     [513b36f2-2bc9-
4e8b-aed9-1f70aeb703f3]                          "192.168.3.14/31"                []               true    []                                      []                     0   forward
a4755730-09e2-4a87-81c4-8ee140dbef81
bd6f871a-4ad9-4767-a8dd-69ed1c1574b9 "ipv4"        496905985     200     single           []                                  bgp      abd89474-7e2a-496d-9ebd-0e3eb8847c42 0      [8610f0a6-6fd0-
402f-b31a-e424c609a9bb]                          "10.1.12.0/24"                   []               true    []                                      []                     0   forward
e5770645-53ab-4897-bd41-3e67248520f8
9dfd36cd-8201-4556-9ada-5f99c28c1d73 "ipv4"        581322276     1       host_route_single []                                 static   a10e88da-0f9c-41e3-9674-1f64e3febc0a 0      [b52009df-726d-
4e7d-860a-e7c89a9c7d8a]                          "192.168.3.251/32"               []               true    []                                      []                     0   forward
a4755730-09e2-4a87-81c4-8ee140dbef81
8aa2783d-2164-4db4-9bc9-5b5e4db90f3c "ipv4"        596083736     0       connected        []                                  connected []                                     0      [dd8d6e97-655b-
4506-9cad-07b1730a59ad]                          "10.2.20.0/24"                   []               true    []                                      []                     0   forward
95329f87-3a6b-49e5-add3-ec4cf7d2a631
77f92ff6-3b59-4e94-9597-e77e51cef78c "ipv4"        770894840     0       []               []                                  local    []                                      0      [8863b0e5-7965-
4aa3-979b-33319e1fdb41]                          "192.168.3.9/32"                 []               true    []                                      []                     0   local
a4755730-09e2-4a87-81c4-8ee140dbef81
07e4bb35-9114-45f5-8b17-d5cfc332ff9e "ipv4"        880605441     0       []               []                                  local    []                                      0      [df0e6d4d-fb68-
4fda-b988-b48a0b4b9bda]                          "192.168.3.1/32"                 []               true    []                                      []                     0   local
a4755730-09e2-4a87-81c4-8ee140dbef81
40453f65-c92f-4373-bfcb-c6418d9311aa "ipv4"        978947083     200     port             []                                  bgp      []                                      0      [bd5358b8-0a6c-
4440-96f0-99af75e82b56]                          "10.1.11.0/24"                   []               true    [e5770645-53ab-4897-bd41-3e67248520f8] []                     0   forward
5f7b72c9-91d3-402f-b963-f5c046f83d0d
fe432ab9-fab8-41b4-afb9-cb929e7476f0 "ipv4"        1074118932    0       []               []                                  local    []                                      0      [a66d03cb-c899-
4af9-af5d-7d2e8477face]                          "192.168.150.0/32"               []               true    []                                      []                     0   local
5f7b72c9-91d3-402f-b963-f5c046f83d0d
e225c0f1-b454-4056-994c-6ca562c7528f "ipv4"        1079837387    200     host_route_single []                                 bgp      8e437c2d-15a5-4139-bdb1-6615eb86d37b 0      [0a9a1904-e504-
42ba-8e2b-713f0f70f464]                          "10.2.20.11/32"                  []               true    [95329f87-3a6b-49e5-add3-ec4cf7d2a631] []                     0   forward
5f7b72c9-91d3-402f-b963-f5c046f83d0d
20cf4076-9e40-42dd-9b83-1d138334d443 "ipv4"        1276762726    0       connected        []                                  connected []                                     0      [bd5358b8-0a6c-
```

BGP GR Troubleshooting

- ovsdb-client dump Tunnel_Endpoint

```
leaf1:/home/admin# ovsdb-client dump Tunnel_Endpoint
Tunnel_Endpoint table
_uuid                                configuration_state                destination                          forwarding_state
hw_id                                hw_vlan                            interface                            macs_invalid
mcast_repl_hw_id                     mcast_repl_l2_port                 network_id                           origin
state                                statistics                         vrf
------------------------------------ ---------------------------------- ---------------------------------- ---------------------------------- -
---------------------------------- ---------------------------------- ---------------------------------- ---------------------------------- --
---------------------------------- ---------------------------------- ---------------------------------- ---------------------------------- ---
---------------------------------- ---------------------------------- ----------------------------------
c637c407-0a91-4dbc-bee6-c2e7a14a831b valid               "192.168.1.21" operational      1     []      f8bee46f-90dd-4799-a081-9184dbfdea03 []
17687           d515d8b8-59b3-458b-8856-3ba851947607 [525770f7-47c7-4e1c-8dec-692914e1e75c, becd9d3f-83a7-4627-8da2-de9305b6b060, eb63015f-47e4-
4e15-bfaa-d3c85d4da7ae] evpn   operational {}         2014494f-a018-4d3c-9122-a2e367f809dc
leaf1:/home/admin#
```

- ovsdb-client dump MAC

```
leaf1:/home/admin# ovsdb-client dump MAC | grep evpn
1654f787-8b6a-43dc-a8ea-65b3edc2c459 []                                      []           []     []           []                       evpn    []
"00:50:56:96:62:4f" true         31a67304-cd06-4c0c-9a95-559608235479 true     c637c407-0a91-4dbc-bee6-c2e7a14a831b 31a0582b-b8b9-4b86-81e2-
40c0ba337fb2
85638578-f9d2-40ac-901f-bd74878d52a0 []                                      []           []     []           []                       evpn    []
"00:50:56:96:94:3f" true         31a67304-cd06-4c0c-9a95-559608235479 true     c637c407-0a91-4dbc-bee6-c2e7a14a831b c6bf3ccf-d889-4e4f-8219-
35a3fa831a78
leaf1:/home/admin#
leaf1:/home/admin#
leaf1:/home/admin#
```

BGP GR Troubleshooting

Live debugging

- For higher scales, during GR on both restarter and helper command “**debug db rx table <table_name>**” can be used to monitor incremental changes. “MAC”, “Tunnel_Endpoint” and “Route” tables can be monitored for any deletions/re-additions. The changes can be seen using “show debug buffer”.
- For lower scales, the tables can also be monitored using “**ovsdb-client monitor <table_name>**”.

Demo

aruba
a Hewlett Packard
Enterprise company

Thank you

vincent.giles@hpe.com