LAB GUIDE

Dual-homed internet access – with 1 border GW – Lab1

bgp bestpath as-path multipath-relax

Important! **This guide assumes that the AOS-CX ova has been installed and works in GNS3 or EVE-NG. Please refer to GNS3/EVE-NG initial setup labs if required.**
https://www.eve-ng.net/index.php/documentation/howtos/howto-add-aruba-cx-switch/

At this time, EVE-NG does not support exporting/importing AOS-CX startup-config. The lab user should copy/paste the AOS-CX node configuration from the lab guide as described in the lab guide if required.

TABLE OF CONTENTS

Lab Objective

This lab will enable the reader to gain hands-on experience with BGP attributes and routing engineering for a multi-path case study of Internet access using one border gateway.

Lab Overview

This lab guide explains how to configure BGP on AOS-CX switch with eBGP to Internet Service Provider routers and iBGP to intranet core route-reflector router.

Please read the BGP section of the AOS-CX 10.7 IP Routing Guide (https://www.arubanetworks.com/techdocs/AOS-CX/10.07/HTML/5200-7858/Content/Chp_BGP/bgp-ove28.htm).

During this lab, you'll be able to:

- Configure all the point-to-point circuits
- Configure OSPF
- Configure iBGP for the intranet part with route-reflector function
- Configure eBGP between the border gateway and both ISPs
- Configure eBGP between ISPs
- Configure eBGP filtering to avoid intranet transit
- Configure endpoints for testing
- Analyze the best path based on BGP criteria
- Adjust the BGP configuration for multi-path scenario and test data-path.

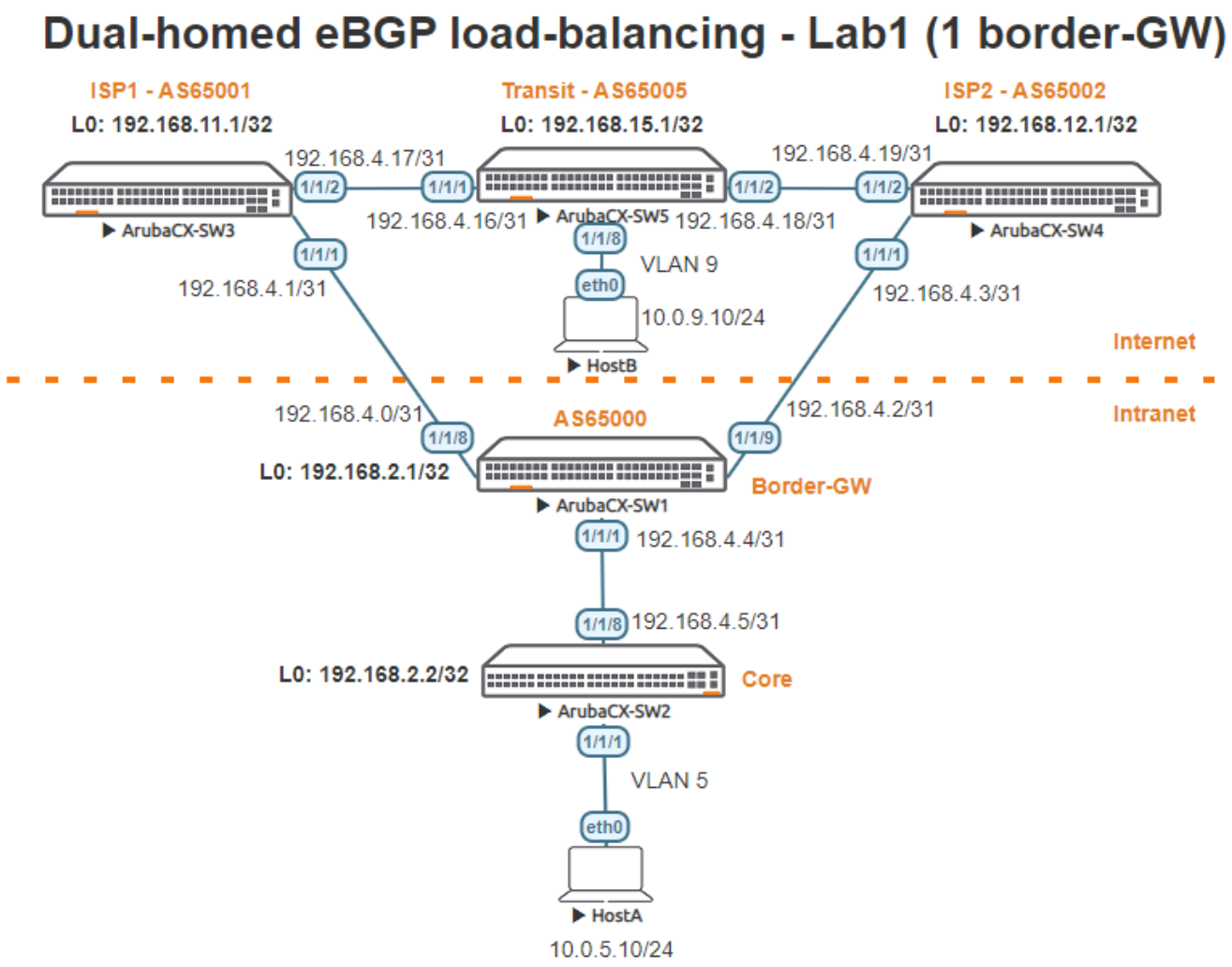
The minimum required AOS-CX Switch Simulator version for this lab is 10.5. It is recommended to use release 10.8 or later.

This lab uses EVE-NG but GNS3 can be used as well.

For educational purpose, this lab uses Private AS Numbers for ISP routers which does not reflect real use-case.

Lab Network Layout

Here is the proposed topology:



Notes:

- The Core SW2 is BGP route-reflector and the border-gateway SW1 a BGP route-reflector-client. This is a typical use-case in production networks.

- For simplicity, HostA is directly attached to the Core, but typically, an access layer would be added in a production network.

- In real networks, a firewall would be inserted between the border-gateway SW1 and the Core layer in order to create a DMZ.

- More than one Core switch is used in production. An additional border-gateway and Core switch is covered in another BGP lab.

Lab Tasks

Task 1 – Lab setup

- In EVE-NG, **import the .zip lab file containing the "unl" file.**
  All the connections between nodes are already set-up. Appropriate numbers of CPUs (2), RAM (4096 MB) and interfaces are already allocated.
- Check the connectivity as proposed above
- Start all the devices (5 AOS-CX switches and 2 hosts)
- Open each switch console and log in with user "admin".
  The switches will ask to enter a new password. This new password can be an empty password for simplicity in this lab.
- Apply (copy/paste) the baseline configuration as proposed below

Baseline configuration proposal (for initial copy/paste):

SW1

```
hostname SW1
!
vlan 1
interface 1/1/1
    no shutdown
    ip address 192.168.4.4/31
interface 1/1/8
    no shutdown
    ip address 192.168.4.0/31
interface 1/1/9
    no shutdown
    ip address 192.168.4.2/31
interface loopback 0
    ip address 192.168.2.1/32
```

SW2

```
hostname SW2
!
vlan 5
interface 1/1/1
    no shutdown
    no routing
    vlan access 5
interface 1/1/8
    no shutdown
    ip address 192.168.4.5/31
interface loopback 0
    ip address 192.168.2.2/32
interface vlan 5
    ip address 10.0.5.1/24
```

SW3

```
hostname SW3
!
interface 1/1/1
    no shutdown
    ip address 192.168.4.1/31
interface 1/1/2
    no shutdown
    ip address 192.168.4.17/31
interface loopback 0
    ip address 192.168.11.1/32
```

SW4

```
hostname SW4
!
interface 1/1/1
    no shutdown
    ip address 192.168.4.3/31
interface 1/1/2
    no shutdown
    ip address 192.168.4.19/31
interface loopback 0
    ip address 192.168.12.1/32
!
```

SW5

```
hostname SW5
!
vlan 9
interface 1/1/1
    no shutdown
    ip address 192.168.4.16/31
interface 1/1/2
    no shutdown
    ip address 192.168.4.18/31
interface 1/1/8
    no shutdown
    no routing
    vlan access 9
interface loopback 0
    ip address 192.168.15.1/32
interface vlan 9
    ip address 10.0.9.1/24
```

HostA / HostB

```
HostA:
VPCS> ip 10.0.5.10/24 10.0.5.1

HostB:
VPCS> ip 10.0.9.10/24 10.0.9.1
```

- Verify the connectivity through LLDP neighbor information as follows:

SW1

```
SW1# show lldp neighbor-info

LLDP Neighbor Information
=========================

Total Neighbor Entries          : 3
Total Neighbor Entries Deleted  : 0
Total Neighbor Entries Dropped  : 0
Total Neighbor Entries Aged-Out : 0

LOCAL-PORT  CHASSIS-ID         PORT-ID                 PORT-DESC                TTL      SYS-NAME
-------------------------------------------------------------------------------------------------
1/1/1       08:00:09:54:97:83  1/1/8                   1/1/8                    120      SW2
1/1/8       08:00:09:5b:7e:2d  1/1/1                   1/1/1                    120      SW3
1/1/9       08:00:09:6d:11:67  1/1/1                   1/1/1                    120      SW4
```

SW2

```
SW2# show lldp neighbor-info

LLDP Neighbor Information
=========================

Total Neighbor Entries          : 1
Total Neighbor Entries Deleted  : 0
Total Neighbor Entries Dropped  : 0
Total Neighbor Entries Aged-Out : 0

LOCAL-PORT  CHASSIS-ID         PORT-ID                 PORT-DESC                TTL      SYS-NAME
-------------------------------------------------------------------------------------------------
1/1/8       08:00:09:d7:5f:0f  1/1/1                   1/1/1                    120      SW1
```

SW3

```
SW3# show lldp neighbor-info

LLDP Neighbor Information
=========================

Total Neighbor Entries          : 2
Total Neighbor Entries Deleted  : 0
Total Neighbor Entries Dropped  : 0
Total Neighbor Entries Aged-Out : 0

LOCAL-PORT  CHASSIS-ID         PORT-ID                 PORT-DESC                TTL      SYS-NAME
-------------------------------------------------------------------------------------------------
1/1/1       08:00:09:d7:5f:0f  1/1/8                   1/1/8                    120      SW1
1/1/2       08:00:09:8b:4a:29  1/1/1                   1/1/1                    120      SW5
```

SW4

```
SW4# show lldp neighbor-info

LLDP Neighbor Information
=========================

Total Neighbor Entries          : 2
Total Neighbor Entries Deleted  : 0
Total Neighbor Entries Dropped  : 0
Total Neighbor Entries Aged-Out : 0

LOCAL-PORT  CHASSIS-ID         PORT-ID                 PORT-DESC                TTL      SYS-NAME
-------------------------------------------------------------------------------------------------
1/1/1       08:00:09:d7:5f:0f  1/1/9                   1/1/9                    120      SW1
1/1/2       08:00:09:8b:4a:29  1/1/2                   1/1/2                    120      SW5
```

SW5

```
SW5# show lldp neighbor-info

LLDP Neighbor Information
=========================

Total Neighbor Entries          : 2
Total Neighbor Entries Deleted  : 0
Total Neighbor Entries Dropped  : 0
Total Neighbor Entries Aged-Out : 0
```

```
LOCAL-PORT  CHASSIS-ID         PORT-ID                PORT-DESC                TTL      SYS-NAME
-------------------------------------------------------------------------------------------------
1/1/1       08:00:09:5b:7e:2d  1/1/2                  1/1/2                    120      SW3
1/1/2       08:00:09:6d:11:67  1/1/2                  1/1/2                    120      SW4
```

- Verify the Hosts connectivity to their default-gateway:

HostA

```
VPCS> show ip

NAME        : VPCS[1]
IP/MASK     : 10.0.5.10/24
GATEWAY     : 10.0.5.1
DNS         :
MAC         : 00:50:79:66:68:07
LPORT       : 20000
RHOST:PORT  : 127.0.0.1:30000
MTU         : 1500

VPCS> ping 10.0.5.1

84 bytes from 10.0.5.1 icmp_seq=1 ttl=64 time=2.182 ms
84 bytes from 10.0.5.1 icmp_seq=2 ttl=64 time=1.228 ms
```

HostB

```
VPCS> show ip

NAME        : VPCS[1]
IP/MASK     : 10.0.9.10/24
GATEWAY     : 10.0.9.1
DNS         :
MAC         : 00:50:79:66:68:06
LPORT       : 20000
RHOST:PORT  : 127.0.0.1:30000
MTU         : 1500

VPCS> ping 10.0.9.1

84 bytes from 10.0.9.1 icmp_seq=1 ttl=64 time=2.671 ms
84 bytes from 10.0.9.1 icmp_seq=2 ttl=64 time=1.697 ms
```

Task 2 – Configure Intranet Routing

There are 2 main steps:

- Configuring OSPF as the IGP (Interior Gateway Protocol) for loopback addresses reachability.
- Configure iBGP between the loopback addresses of the Core route-reflector and the border-gateway.

Note: OSPF and BGP authentication is skipped in this lab for simplicity. It is a best practice to enable such enforcement in a production network.

Step #1: Configure OSPF

SW1(config)#

```
router ospf 1
    passive-interface default
    redistribute local loopback
    area 0.0.0.0
interface 1/1/1
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf network point-to-point
```

SW2(config)#

```
router ospf 1
    passive-interface default
    redistribute local loopback
    area 0.0.0.0
interface 1/1/8
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf network point-to-point
```

Step #2: Verify OSPF routing

SW1

```
SW1# show ip ospf neighbors
VRF : default                          Process : 1
=====================================================

Total Number of Neighbors : 1

Neighbor ID      Priority  State             Nbr Address       Interface
-------------------------------------------------------------------------
192.168.2.2      n/a       FULL              192.168.4.5        1/1/1


SW1# show ip ospf routes
Codes: i - Intra-area route, I - Inter-area route
       E1 - External type-1, E2 - External type-2

OSPF Process ID 1 VRF default, Routing Table
--------------------------------------------

Total Number of Routes : 2

192.168.2.2/32     (E2)
     via 192.168.4.5 interface 1/1/1, cost 25 distance 110
192.168.4.4/31     (i) area: 0.0.0.0
     directly attached to interface 1/1/1, cost 100 distance 110
```

SW2

```
SW2# show ip ospf neighbors
VRF : default                          Process : 1
=====================================================

Total Number of Neighbors : 1

Neighbor ID      Priority  State             Nbr Address       Interface
-------------------------------------------------------------------------
192.168.2.1      n/a       FULL              192.168.4.4        1/1/8


SW2# show ip ospf routes
Codes: i - Intra-area route, I - Inter-area route
       E1 - External type-1, E2 - External type-2

OSPF Process ID 1 VRF default, Routing Table
--------------------------------------------

Total Number of Routes : 2

192.168.2.1/32     (E2)
     via 192.168.4.4 interface 1/1/8, cost 25 distance 110
192.168.4.4/31     (i) area: 0.0.0.0
     directly attached to interface 1/1/8, cost 100 distance 110
```

Step #3: Configure iBGP

It is a best practice to enable logging for BGP neighbor status change, as it can be very useful for the support to troubleshoot past incidents.

iBGP peering are set-up with the Loopback IP addresses that are reachable through OSPF. The BGP TCP session is sourced from the Loopback IP address.

In order to avoid injecting subnets of ISP links into the IGP (OSPF), the border-gateway SW1 will replace the next-hop IP addresses of ISP1 and ISP2 by its own IP address: this is the role of the “next-hop-self” command, towards the Core route-reflector.

It is a best practice to enable “fall-over” for iBGP peering in order to remove all BGP routes coming from the given iBGP neighbor as soon as the iBGP neighbor is unreachable, the loopback reachability being maintained by OSPF. As soon as OSPF discards

the loopback entry from the RIB, BGP removes also the routes learnt from that iBGP neighbor in less than one second.

There are multiple ways of injecting the HostA subnet into BGP:

- Network command (which is proposed in this lab)
- Redistribution of connected in BGP
- Redistribution of OSPF into BGP with user subnet being in OSPF table (through redistribution of connected in OSPF for instance).

SW1(config)#

```
router bgp 65000
    bgp router-id 192.168.2.1
    bgp log-neighbor-changes
    neighbor 192.168.2.2 remote-as 65000
    neighbor 192.168.2.2 description Core-RR
    neighbor 192.168.2.2 fall-over
    neighbor 192.168.2.2 update-source loopback 0
    address-family ipv4 unicast
        neighbor 192.168.2.2 activate
        neighbor 192.168.2.2 next-hop-self
    exit-address-family
```

SW2(config)#

```
router bgp 65000
    bgp router-id 192.168.2.2
    bgp log-neighbor-changes
    neighbor 192.168.2.1 remote-as 65000
    neighbor 192.168.2.1 fall-over
    neighbor 192.168.2.1 update-source loopback 0
    address-family ipv4 unicast
        neighbor 192.168.2.1 activate
        neighbor 192.168.2.1 route-reflector-client
        network 10.0.5.0/24
    exit-address-family
```

Step #4: Verify BGP sessions and BGP routes

Use show bgp, show ip rib, show ip route to check the expected routes in the FIB.

SW1

```
SW1# show bgp ipv4 unicast summary
VRF : default
BGP Summary
-----------
 Local AS               : 65000        BGP Router Identifier  : 192.168.2.1
 Peers                  : 1            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
 192.168.2.2     65000       238     237     03h:23m:16s  Established   Up


SW1# 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 192.168.2.1

    Network            Nexthop         Metric     LocPrf     Weight Path
*>i 10.0.5.0/24        192.168.2.2     0          100        0       i
Total number of entries 1


SW1# show ip rib

Displaying ipv4 routes in RIB

Origin Codes: R - RIP, O - OSPFv2, B - BGP
              C - connected, S - static, L - local
Type Codes:   E - External BGP, I - Internal BGP
              IA - OSPF inter area, ia - OSPF intra area
              E1 - OSPF external type 1, E2 - OSPF external type 2
              EV - BGP EVPN, V - BGP VPN
* indicates selected for forwarding

VRF: default

Prefix              Nexthop          Interface     VRF(egress)       Origin/    Distance/     Age
```

```
                                                                    Type       Metric
-------------------------------------------------------------------------------------------------------
*10.0.5.0/24        192.168.2.2       -              -               B/I        [200/0]      14h:44m:07s
*192.168.2.1/32     -                 loopback0      -               L          [0/0]        -
*192.168.2.2/32     192.168.4.5       1/1/1          -               O/E2       [110/25]     15h:55m:00s
*192.168.4.0/31     -                 1/1/8          -               C          [0/0]        -
*192.168.4.0/32     -                 1/1/8          -               L          [0/0]        -
*192.168.4.2/31     -                 1/1/9          -               C          [0/0]        -
*192.168.4.2/32     -                 1/1/9          -               L          [0/0]        -
 192.168.4.4/31     -                 1/1/1          -               O/ia       [110/100]    15h:55m:05s
*192.168.4.4/31     -                 1/1/1          -               C          [0/0]        -
*192.168.4.4/32     -                 1/1/1          -               L          [0/0]        -

Total Route Count : 10


SW1# show ip route

Displaying ipv4 routes selected for forwarding

Origin Codes: C - connected, S - static, L - local
              R - RIP, B - BGP, O - OSPF
Type Codes:   E - External BGP, I - Internal BGP, V - VPN, EV - EVPN
              IA - OSPF internal area, E1 - OSPF external type 1
              E2 - OSPF external type 2

VRF: default

Prefix              Nexthop          Interface     VRF(egress)       Origin/    Distance/    Age
                                                                     Type       Metric
-------------------------------------------------------------------------------------------------------
10.0.5.0/24         192.168.4.5      1/1/1         -                 B/I        [200/0]      14h:39m:53s
192.168.2.1/32      -                loopback0     -                 L          [0/0]        -
192.168.2.2/32      192.168.4.5      1/1/1         -                 O/E2       [110/25]     15h:50m:45s
192.168.4.0/31      -                1/1/8         -                 C          [0/0]        -
192.168.4.0/32      -                1/1/8         -                 L          [0/0]        -
192.168.4.2/31      -                1/1/9         -                 C          [0/0]        -
192.168.4.2/32      -                1/1/9         -                 L          [0/0]        -
192.168.4.4/31      -                1/1/1         -                 C          [0/0]        -
192.168.4.4/32      -                1/1/1         -                 L          [0/0]        -

Total Route Count : 9
```

SW2

```
SW2#  show bgp ipv4 unicast summary
VRF : default
BGP Summary
-----------
 Local AS               : 65000        BGP Router Identifier  : 192.168.2.2
 Peers                  : 1            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
 192.168.2.1     65000       314     318     03h:23m:44s  Established   Up


SW2# 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 192.168.2.2

    Network            Nexthop         Metric     LocPrf     Weight Path
*>  10.0.5.0/24        0.0.0.0         0          100        0      i
Total number of entries 1


SW2# show ip rib

Displaying ipv4 routes in RIB
```

```
Origin Codes: R - RIP, O - OSPFv2, B - BGP
              C - connected, S - static, L - local
Type Codes:   E - External BGP, I - Internal BGP
              IA - OSPF inter area, ia - OSPF intra area
              E1 - OSPF external type 1, E2 - OSPF external type 2
              EV - BGP EVPN, V - BGP VPN
* indicates selected for forwarding

VRF: default

Prefix              Nexthop          Interface     VRF(egress)       Origin/    Distance/    Age
                                                                     Type       Metric
-------------------------------------------------------------------------------------------------
*10.0.5.0/24        -                vlan5         -                 C          [0/0]        -
*10.0.5.1/32        -                vlan5         -                 L          [0/0]        -
*192.168.2.1/32     192.168.4.4      1/1/8         -                 O/E2       [110/25]     15h:54m:02s
*192.168.2.2/32     -                loopback0     -                 L          [0/0]        -
 192.168.4.4/31     -                1/1/8         -                 O/ia       [110/100]    15h:54m:37s
*192.168.4.4/31     -                1/1/8         -                 C          [0/0]        -
*192.168.4.5/32     -                1/1/8         -                 L          [0/0]        -

Total Route Count : 7


SW2# show ip route

Displaying ipv4 routes selected for forwarding

Origin Codes: C - connected, S - static, L - local
              R - RIP, B - BGP, O - OSPF
Type Codes:   E - External BGP, I - Internal BGP, V - VPN, EV - EVPN
              IA - OSPF internal area, E1 - OSPF external type 1
              E2 - OSPF external type 2

VRF: default

Prefix              Nexthop          Interface     VRF(egress)       Origin/    Distance/    Age
                                                                     Type       Metric
-------------------------------------------------------------------------------------------------
10.0.5.0/24         -                vlan5         -                 C          [0/0]        -
10.0.5.1/32         -                vlan5         -                 L          [0/0]        -
192.168.2.1/32      192.168.4.4      1/1/8         -                 O/E2       [110/25]     15h:53m:43s
192.168.2.2/32      -                loopback0     -                 L          [0/0]        -
192.168.4.4/31      -                1/1/8         -                 C          [0/0]        -
192.168.4.5/32      -                1/1/8         -                 L          [0/0]        -

Total Route Count : 6
```

SW1 received OSPF route of SW2 loopback address, as well as the BGP route of 10.0.5.0/24 subnet.

Task 3 – Configure internet peering

Step #1: Configure BGP and peering between internet nodes and to intranet

Configure BGP to establish peering between ISP1 and Transit-ISP, Transit-ISP and ISP2, ISP1 and Intranet, ISP2 and Intranet.

It is a best practice for eBGP peering to enable fast-external-fallover. This feature allows immediate removal of BGP routes coming from an eBGP neighbor when a link failure is detected, within milliseconds. This alleviates any requirement for BFD (bidirectional forwarding detection) in case of direct point-to-point L3 connectivity between BGP speakers/nodes. This is enabled by default in AOS-CX 10.8 and beyond.

SW5 must advertise the HostB subnet in order to allow connectivity between HostA and HostB.

Here is the additional configuration on each nodes:

SW1	SW3 (ISP1)
`router bgp 65000`	`router bgp 65001`

```
    bgp fast-external-fallover
    neighbor 192.168.4.1 remote-as 65001
    neighbor 192.168.4.1 description ISP1 peering
    neighbor 192.168.4.3 remote-as 65002
    neighbor 192.168.4.3 description ISP2 peering
    address-family ipv4 unicast
        neighbor 192.168.4.1 activate
        neighbor 192.168.4.3 activate
```

```
    bgp router-id 192.168.11.1
    bgp fast-external-fallover
    bgp log-neighbor-changes
    neighbor 192.168.4.0 remote-as 65000
    neighbor 192.168.4.16 remote-as 65005
    address-family ipv4 unicast
        neighbor 192.168.4.0 activate
        neighbor 192.168.4.16 activate
    exit-address-family
```

SW4 (ISP2)

```
router bgp 65002
    bgp router-id 192.168.12.1
    bgp fast-external-fallover
    bgp log-neighbor-changes
    neighbor 192.168.4.2 remote-as 65000
    neighbor 192.168.4.18 remote-as 65005
    address-family ipv4 unicast
        neighbor 192.168.4.2 activate
        neighbor 192.168.4.18 activate
    exit-address-family
```

SW5 (Transit ISP)

```
router bgp 65005
    bgp router-id 192.168.15.1
    bgp fast-external-fallover
    bgp log-neighbor-changes
    neighbor 192.168.4.17 remote-as 65001
    neighbor 192.168.4.19 remote-as 65002
    address-family ipv4 unicast
        neighbor 192.168.4.17 activate
        neighbor 192.168.4.19 activate
        network 10.0.9.0/24
    exit-address-family
```

Step #2: Apply BGP filtering

Note: It is important to enforce that the intranet network is NOT a Transit network for the internet traffic. ISPs apply such Route Filters inbound of customers peering. It is a best practice to also apply outbound filter on the SW1 border gateway to prevent this.

For simplicity, the filtering applied on ISP side is not covered in this lab. The proposed filter is applied outbound on SW1.

A simple rule could be to use AS-path filter with regular expression to allow only the internal prefixes that do not carry any other AS than the internal AS number. Another option could be to filter based on IP prefix-list or based on communities…

- Check the routes advertised by SW1 to ISP1 and ISP2 before applying the filter.

SW1

```
SW1# show bgp ipv4 unicast neighbors 192.168.4.1 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 192.168.2.1

    Network            Nexthop         Metric     LocPrf     Weight Path
*>e 10.0.5.0/24        192.168.4.0     0          0          0       65000 i
Total number of entries 1


SW1# show bgp ipv4 unicast neighbors 192.168.4.3 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 192.168.2.1

    Network            Nexthop         Metric     LocPrf     Weight Path
*>e 10.0.5.0/24        192.168.4.2     0          0          0       65000 i
*>e 10.0.9.0/24        192.168.4.2     0          0          0       65000 65001 65005 i
Total number of entries 2
```

You should see that 10.0.9.0/24 is advertised to one ISP, although 10.0.9.0/24 is not part of the Intranet. This is what the configuration below will prevent.

Here is a proposal for the lab based on AS-path filter (regexp used allows only empty AS-path) then adding prefix-list

SW1(config)#

```
ip aspath-list intranet-AS seq 10 permit ^$

route-map isp1-out permit seq 10
```

```
     match aspath-list intranet-AS
route-map isp2-out permit seq 10
     match aspath-list intranet-AS
!
router bgp 65000
    address-family ipv4 unicast
        neighbor 192.168.4.1 route-map isp1-out out
        neighbor 192.168.4.3 route-map isp2-out out
    exit-address-family
```

- Check refreshed advertisement:

SW1

```
SW1# clear bgp *
SW1# show bgp ipv4 unicast summary
VRF : default
BGP Summary
-----------
 Local AS               : 65000        BGP Router Identifier  : 192.168.2.1
 Peers                  : 3            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
 192.168.2.2     65000       4       4       00h:00m:08s  Established   Up
 192.168.4.1     65001       5       4       00h:00m:08s  Established   Up
 192.168.4.3     65002       4       4       00h:00m:08s  Established   Up


SW1# show bgp ipv4 unicast neighbors 192.168.4.1 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 192.168.2.1

    Network            Nexthop         Metric     LocPrf     Weight Path
*>e 10.0.5.0/24        192.168.4.0     0          0          0       65000 i
Total number of entries 1


SW1# show bgp ipv4 unicast neighbors 192.168.4.3 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 192.168.2.1

    Network            Nexthop         Metric     LocPrf     Weight Path
*>e 10.0.5.0/24        192.168.4.2     0          0          0       65000 i
Total number of entries 1
```

You should no longer see 10.0.9.0/24 advertised from SW1 to ISPs.

- Complete filtering with prefix-list:

SW1(config)#

```
ip aspath-list intranet-AS seq 10 permit ^$
!
ip prefix-list intranet-IP seq 10 permit 10.0.5.0/24
!
route-map isp1-out permit seq 10
     match ip address prefix-list intranet-IP
     match aspath-list intranet-AS
!
route-map isp2-out permit seq 10
     match ip address prefix-list intranet-IP
     match aspath-list intranet-AS
```

- Check again refreshed advertisement:

SW1

```
SW1# clear bgp *
SW1# show bgp ipv4 unicast summary
VRF : default
BGP Summary
-----------
 Local AS               : 65000        BGP Router Identifier  : 192.168.2.1
 Peers                  : 3            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
 192.168.2.2     65000       4       4       00h:00m:10s  Established   Up
 192.168.4.1     65001       4       4       00h:00m:10s  Established   Up
 192.168.4.3     65002       5       4       00h:00m:10s  Established   Up


SW1# show bgp ipv4 unicast neighbors 192.168.4.1 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 192.168.2.1

    Network            Nexthop         Metric     LocPrf     Weight Path
*>e 10.0.5.0/24        192.168.4.0     0          0          0       65000 i
Total number of entries 1


SW1# show bgp ipv4 unicast neighbors 192.168.4.3 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 192.168.2.1

    Network            Nexthop         Metric     LocPrf     Weight Path
*>e 10.0.5.0/24        192.168.4.2     0          0          0       65000 i
Total number of entries 1
```

You should see no difference as 10.0.9.0/24 was already filtered out from the previous AS-path filter. This additional prefix-list filter is another protection allowing only specific subnets (a subset of intranet IP space).

Step #3: Verify routing

- All BGP sessions on SW1, SW3, SW4, SW5 should be in “Established” state:

SW1

```
SW1# show bgp ipv4 unicast summary
VRF : default
BGP Summary
-----------
 Local AS               : 65000        BGP Router Identifier  : 192.168.2.1
 Peers                  : 3            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
 192.168.2.2     65000       1078    1080    15h:34m:02s  Established   Up
 192.168.4.1     65001       1060    1085    00h:40m:06s  Established   Up
 192.168.4.3     65002       1067    1092    00h:39m:58s  Established   Up
```

SW3 (ISP1)

```
SW3#  show bgp ipv4 unicast summary
VRF : default
BGP Summary
-----------
 Local AS               : 65001        BGP Router Identifier  : 192.168.11.1
```

```
 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
 192.168.4.0     65000       1207    1211    00h:40m:39s  Established   Up
 192.168.4.16    65005       1193    1188    17h:05m:34s  Established   Up
```

SW4 (ISP2)

```
SW4# show bgp ipv4 unicast summary
VRF : default
BGP Summary
-----------
 Local AS              : 65002        BGP Router Identifier  : 192.168.12.1
 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
 192.168.4.2     65000       1201    1204    00h:41m:38s  Established   Up
 192.168.4.18    65005       1191    1186    17h:06m:36s  Established   Up
```

SW5 (Transit)

```
SW5# show bgp ipv4 unicast summary
VRF : default
BGP Summary
-----------
 Local AS              : 65005        BGP Router Identifier  : 192.168.15.1
 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
 192.168.4.17    65001       1186    1194    17h:08m:31s  Established   Up
 192.168.4.19    65002       1187    1193    17h:08m:26s  Established   Up
```

- On SW5, toggle BGP session between ISP1 and Transit in order to make routes coming from ISP2 the most stable eBGP routes on Transit node. **This is for educational purpose on the BGP decision algorithm (most stable route)**.

SW5 (Transit)

```
SW5(config)# router bgp 65005
SW5(config-bgp)# neighbor 192.168.4.17 shut
SW5(config-bgp)# no neighbor 192.168.4.17 shut

SW5# show bgp ipv4 unicast summary
VRF : default
BGP Summary
-----------
 Local AS              : 65005        BGP Router Identifier  : 192.168.15.1
 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
 192.168.4.17    65001       1228    1237    00h:11m:05s  Established   Up
 192.168.4.19    65002       1225    1232    00h:18m:24s  Established   Up
```

Peering with ISP2 should be older than peering with ISP1.

- BGP routes on ISP1, ISP2 and Transit

SW3 (ISP1)

```
SW3# 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 192.168.11.1

    Network            Nexthop         Metric     LocPrf     Weight Path
*>e 10.0.5.0/24        192.168.4.0     0          100        0       65000 i
* e 10.0.5.0/24        192.168.4.16    0          100        0       65005 65002 65000 i
```

```
*>e 10.0.9.0/24         192.168.4.16     0          100        0       65005 i
Total number of entries 3
```

SW5 (Transit)

```
SW5# 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 192.168.15.1

    Network            Nexthop          Metric     LocPrf     Weight Path
* e 10.0.5.0/24         192.168.4.17     0          100        0       65001 65000 i
*>e 10.0.5.0/24         192.168.4.19     0          100        0       65002 65000 i
*>  10.0.9.0/24         0.0.0.0          0          100        0       i
Total number of entries 3
```

SW4 (ISP2)

```
SW4# 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 192.168.12.1

    Network            Nexthop          Metric     LocPrf     Weight Path
*>e 10.0.5.0/24         192.168.4.2      0          100        0       65000 i
*>e 10.0.9.0/24         192.168.4.18     0          100        0       65005 i
Total number of entries 2
```

- BGP RIB and IP FIB on intranet nodes: SW1 and SW2

SW1

```
SW1# 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 192.168.2.1

    Network            Nexthop          Metric     LocPrf     Weight Path
*>i 10.0.5.0/24         192.168.2.2      0          100        0       i
* e 10.0.9.0/24         192.168.4.1      0          100        0       65001 65005 i
*>e 10.0.9.0/24         192.168.4.3      0          100        0       65002 65005 i
Total number of entries 3


SW1# show ip route

Displaying ipv4 routes selected for forwarding

Origin Codes: C - connected, S - static, L - local
              R - RIP, B - BGP, O - OSPF
Type Codes:   E - External BGP, I - Internal BGP, V - VPN, EV - EVPN
              IA - OSPF internal area, E1 - OSPF external type 1
              E2 - OSPF external type 2

VRF: default

Prefix              Nexthop          Interface     VRF(egress)       Origin/    Distance/    Age
                                                                     Type       Metric
-------------------------------------------------------------------------------------------------
10.0.5.0/24         192.168.4.5      1/1/1         -                 B/I        [200/0]      00h:33m:12s
10.0.9.0/24         192.168.4.3      1/1/9         -                 B/E        [20/0]       00h:26m:12s
192.168.2.1/32      -                loopback0     -                 L          [0/0]        -
192.168.2.2/32      192.168.4.5      1/1/1         -                 O/E2       [110/25]     17h:33m:39s
192.168.4.0/31      -                1/1/8         -                 C          [0/0]        -
192.168.4.0/32      -                1/1/8         -                 L          [0/0]        -
192.168.4.2/31      -                1/1/9         -                 C          [0/0]        -
192.168.4.2/32      -                1/1/9         -                 L          [0/0]        -
```

```
192.168.4.4/31      -               1/1/1           -               C           [0/0]       -
192.168.4.4/32      -               1/1/1           -               L           [0/0]       -

Total Route Count : 10
```

SW2

```
SW2# 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 192.168.2.2

    Network            Nexthop         Metric     LocPrf     Weight Path
*>  10.0.5.0/24        0.0.0.0         0          100        0      i
*>i 10.0.9.0/24        192.168.2.1     0          100        0      65002 65005 i
Total number of entries 2


SW2# show ip route

Displaying ipv4 routes selected for forwarding

Origin Codes: C - connected, S - static, L - local
              R - RIP, B - BGP, O - OSPF
Type Codes:   E - External BGP, I - Internal BGP, V - VPN, EV - EVPN
              IA - OSPF internal area, E1 - OSPF external type 1
              E2 - OSPF external type 2

VRF: default

Prefix              Nexthop          Interface     VRF(egress)       Origin/    Distance/    Age
                                                                     Type       Metric
-------------------------------------------------------------------------------------------------
10.0.5.0/24         -                vlan5         -                 C          [0/0]        -
10.0.5.1/32         -                vlan5         -                 L          [0/0]        -
10.0.9.0/24         192.168.4.4      1/1/8         -                 B/I        [200/0]      00h:28m:38s
192.168.2.1/32      192.168.4.4      1/1/8         -                 O/E2       [110/25]     17h:36m:06s
192.168.2.2/32      -                loopback0     -                 L          [0/0]        -
192.168.4.4/31      -                1/1/8         -                 C          [0/0]        -
192.168.4.5/32      -                1/1/8         -                 L          [0/0]        -

Total Route Count : 7
```

SW1 sends to SW2 the best route with its own next-hop IP address, instead of the ISP1 and ISP2 peering IP addresses, thanks to the next-hop-self command.

Step #3: Check connectivity between HostA and HostB

Ping HostB from HostA

HostA

```
VPCS> ping 10.0.9.10

84 bytes from 10.0.9.10 icmp_seq=1 ttl=60 time=15.683 ms
84 bytes from 10.0.9.10 icmp_seq=2 ttl=60 time=5.194 ms
84 bytes from 10.0.9.10 icmp_seq=3 ttl=60 time=5.217 ms
84 bytes from 10.0.9.10 icmp_seq=4 ttl=60 time=12.034 ms
84 bytes from 10.0.9.10 icmp_seq=5 ttl=60 time=7.935 ms
```

Ping HostA from HostB

HostB

```
VPCS> ping 10.0.5.10

84 bytes from 10.0.5.10 icmp_seq=1 ttl=60 time=12.031 ms
84 bytes from 10.0.5.10 icmp_seq=2 ttl=60 time=5.268 ms
84 bytes from 10.0.5.10 icmp_seq=3 ttl=60 time=5.292 ms
84 bytes from 10.0.5.10 icmp_seq=4 ttl=60 time=5.939 ms
84 bytes from 10.0.5.10 icmp_seq=5 ttl=60 time=5.439 ms
```

Traceroute HostB from HostA

HostA

```
VPCS> trace 10.0.9.10
trace to 10.0.9.10, 8 hops max, press Ctrl+C to stop
 1   10.0.5.1   1.409 ms  1.239 ms  1.108 ms
 2   192.168.4.4   3.909 ms  1.989 ms  2.065 ms
 3   192.168.4.3   4.117 ms  3.524 ms  3.381 ms
 4   192.168.4.18   7.726 ms  5.145 ms  4.158 ms
 5   *10.0.9.10   14.312 ms (ICMP type:3, code:3, Destination port unreachable)
```

Traceroute HostA from HostB

HostB

```
VPCS> trace 10.0.5.10
trace to 10.0.5.10, 8 hops max, press Ctrl+C to stop
 1   10.0.9.1   2.050 ms  0.932 ms  0.948 ms
 2   192.168.4.19   2.463 ms  2.117 ms  1.702 ms
 3   192.168.4.2   3.494 ms  2.662 ms  2.773 ms
 4   192.168.4.5   11.474 ms  5.203 ms  4.068 ms
 5   *10.0.5.10   5.071 ms (ICMP type:3, code:3, Destination port unreachable)
```

Task 4 – Analysis of selected BGP routes

Explain why, on SW1, ISP2 is preferred over ISP1 for HostB destination. To help you, you'll find a reminder of the **BGP decision algorithm** in the appendix.

- Compare the all the BGP attributes of 10.0.9.0/24:

SW1

```
SW1# show bgp ipv4 unicast 10.0.9.0/24

VRF : default
BGP Local AS 65000       BGP Router-id 192.168.2.1

   Network           : 10.0.9.0/24             Nexthop           : 192.168.4.1
   Peer              : 192.168.4.1             Origin            : IGP
   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           :  65001 65005

   Cluster List      :
   Communities       :
     Ext-Communities :
   Network           : 10.0.9.0/24             Nexthop           : 192.168.4.3
   Peer              : 192.168.4.3             Origin            : IGP
   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           :  65002 65005

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

All BGP criteria look the same: weight, local-pref, AS-path length, metric...The last criteria is the peer IP address, the lowest IP address being preferred. It is not the case in the previous command output; as the "best" router selected is from eBGP 192.168.4.3 which is higher than 192.168.4.1. So it means that another criteria is used before lowest peer IP address. This criteria is not immediately visible in the current show bgp output. It is about the eBGP route stability, i.e. the oldest route is preferred. Please remember that you toggled the BGP peering between SW5 and SW3 to make the eBGP route for 10.0.9.0/24 the most stable through ISP2.

- You can check route age on SW1 against the overall peering age:

SW1

```
SW1# sho bgp ipv4 unicast summary
VRF : default
BGP Summary
-----------
 Local AS               : 65000        BGP Router Identifier  : 192.168.2.1
 Peers                  : 3            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
 192.168.2.2     65000       28      30      00h:21m:44s  Established   Up
 192.168.4.1     65001       32      28      00h:21m:44s  Established   Up
 192.168.4.3     65002       30      28      00h:21m:44s  Established   Up


SW1# show ip rou 10.0.9.0/24

VRF: default

   Prefix           : 10.0.9.0/24                VRF(egress)        : -
   Nexthop          : 192.168.4.3                Interface          : 1/1/9
   Origin           : bgp                        Type               : bgp_ext
   Distance         : 20                         Metric             : 0
   Age              : 00h:09m:10s                Tag                : 0
   Encap Type       : -                          Encap Details      : -
```

The route coming from ISP2 is oldest than the one coming from ISP1.

- Similarly, you can clear on SW1 the BGP session with ISP2, and you will now see the best path through ISP1 (new most stable):

SW1

```
SW1# show bgp ipv4 unicast summary
VRF : default
BGP Summary
-----------
 Local AS               : 65000        BGP Router Identifier  : 192.168.2.1
 Peers                  : 3            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
 192.168.2.2     65000       28      30      00h:21m:44s  Established   Up
 192.168.4.1     65001       32      28      00h:21m:44s  Established   Up
 192.168.4.3     65002       30      28      00h:21m:44s  Established   Up

SW1# show ip route 10.0.9.0/24

VRF: default

   Prefix           : 10.0.9.0/24                VRF(egress)        : -
   Nexthop          : 192.168.4.3                Interface          : 1/1/9
   Origin           : bgp                        Type               : bgp_ext
   Distance         : 20                         Metric             : 0
   Age              : 00h:09m:10s                Tag                : 0
   Encap Type       : -                          Encap Details      : -
```

```
SW1# clear bgp 192.168.4.3
SW1# show ip route 10.0.9.0/24

VRF: default

   Prefix           : 10.0.9.0/24              VRF(egress)        : -
   Nexthop          : 192.168.4.1              Interface          : 1/1/8
   Origin           : bgp                      Type               : bgp_ext
   Distance         : 20                       Metric             : 0
   Age              : 00h:00m:04s              Tag                : 0
   Encap Type       : -                        Encap Details      : -

SW1# show bgp ipv4 unicast summary
VRF : default
BGP Summary
-----------
 Local AS               : 65000        BGP Router Identifier  : 192.168.2.1
 Peers                  : 3            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
 192.168.2.2     65000       32      33      00h:24m:20s  Established   Up
 192.168.4.1     65001       35      31      00h:24m:20s  Established   Up
 192.168.4.3     65002       4       4       00h:00m:11s  Established   Up

SW1# show ip route 10.0.9.0/24

VRF: default

   Prefix           : 10.0.9.0/24              VRF(egress)        : -
   Nexthop          : 192.168.4.1              Interface          : 1/1/8
   Origin           : bgp                      Type               : bgp_ext
   Distance         : 20                       Metric             : 0
   Age              : 00h:00m:14s              Tag                : 0
   Encap Type       : -                        Encap Details      : -
```

This criteria of oldest eBGP route is important to keep in mind while troubleshooting eBGP routing.

As you can see only one best route is preferred although the AS-path length is the same. In some situations like load-balancing request, it could be useful to have these routes considered as equal cost.

Task 5 – set-up load-balancing for outbound traffic to ISP1 and to ISP2

As the AS-path length is the same, the next BGP criteria is considered to select the best route. The intent is to stop such processing when AS-path lengths from 2 routes are identical. This is the purpose of “bgp bestpath as-path multipath-relax” command.

- Configure multipath-relax on SW1:

SW1

```
SW1(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

SW1(config-bgp)# bgp bestpath as-path multipath-relax
All current BGP sessions in VRF default will be restarted
Do you want to continue (y/n)? y
```

This command will reset the BGP sessions, so it is recommended to plan a maintenance window to configure it.

- Check the BGP routing table

SW1

```
SW1# show bgp ipv4 unicast 10.0.9.0/24

VRF : default
BGP Local AS 65000       BGP Router-id 192.168.2.1

   Network           : 10.0.9.0/24              Nexthop           : 192.168.4.1
   Peer              : 192.168.4.1              Origin            : IGP
   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           :  65001 65005

   Cluster List      :
   Communities       :
     Ext-Communities :
   Network           : 10.0.9.0/24              Nexthop           : 192.168.4.3
   Peer              : 192.168.4.3              Origin            : IGP
   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           :  65002 65005

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

The route coming from IPS2 is still the best route. However, you will now see “=” sign in front of the route received from ISP1, showing a multi-path.

SW1

```
SW1# 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 192.168.2.1

    Network          Nexthop         Metric     LocPrf     Weight Path
*>i 10.0.5.0/24      192.168.2.2     0          100        0       i
*=e 10.0.9.0/24      192.168.4.1     0          100        0       65001 65005 i
*>e 10.0.9.0/24      192.168.4.3     0          100        0       65002 65005 i
Total number of entries 3
```

Consequently, the FIB is populated with both Next-hops for the same destination, due to the equal cost paths:

SW1

```
SW1# show ip route 10.0.9.0/24

VRF: default

   Prefix            : 10.0.9.0/24              VRF(egress)       : -
   Nexthop           : 192.168.4.1              Interface         : 1/1/8
   Origin            : bgp                      Type              : bgp_ext
   Distance          : 20                       Metric            : 0
   Age               : 00h:07m:53s              Tag               : 0
   Encap Type        : -                        Encap Details     : -
```

```
   Prefix           : 10.0.9.0/24              VRF(egress)        : -
   Nexthop          : 192.168.4.3              Interface          : 1/1/9
   Origin           : bgp                      Type               : bgp_ext
   Distance         : 20                       Metric             : 0
   Age              : 00h:07m:53s              Tag                : 0
   Encap Type       : -                        Encap Details      : -
```

- Check the data-path on HostA, by doing multiple iterations of trace command.

HostA

```
VPCS> trace 10.0.9.10
trace to 10.0.9.10, 8 hops max, press Ctrl+C to stop
 1   10.0.5.1   1.115 ms  0.804 ms  0.814 ms
 2   192.168.4.4   2.050 ms  1.666 ms  1.850 ms
 3   192.168.4.3   7.038 ms  2.630 ms  2.580 ms
 4   192.168.4.18   3.701 ms  3.813 ms  6.123 ms
 5   *10.0.9.10   4.149 ms (ICMP type:3, code:3, Destination port unreachable)

VPCS> trace 10.0.9.10
trace to 10.0.9.10, 8 hops max, press Ctrl+C to stop
 1   10.0.5.1   1.302 ms  0.913 ms  0.784 ms
 2   192.168.4.4   2.164 ms  1.595 ms  1.570 ms
 3   192.168.4.1   5.251 ms  2.963 ms  2.744 ms
 4   192.168.4.18   3.737 ms  2.981 ms  3.431 ms
 5   *10.0.9.10   3.684 ms (ICMP type:3, code:3, Destination port unreachable)

VPCS> trace 10.0.9.10
trace to 10.0.9.10, 8 hops max, press Ctrl+C to stop
 1   10.0.5.1   0.760 ms  0.710 ms  0.765 ms
 2   192.168.4.4   2.145 ms  1.739 ms  1.711 ms
 3   192.168.4.1   2.770 ms  2.712 ms  2.486 ms
 4   192.168.4.18   3.608 ms  3.601 ms  3.084 ms
 5   *10.0.9.10   13.041 ms (ICMP type:3, code:3, Destination port unreachable)

VPCS> trace 10.0.9.10
trace to 10.0.9.10, 8 hops max, press Ctrl+C to stop
 1   10.0.5.1   1.249 ms  1.003 ms  0.661 ms
 2   192.168.4.4   1.942 ms  1.570 ms  1.621 ms
 3   192.168.4.3   2.604 ms  2.524 ms  2.099 ms
 4   192.168.4.18   3.702 ms  3.247 ms  3.198 ms
 5   *10.0.9.10   3.566 ms (ICMP type:3, code:3, Destination port unreachable)
```

As you can verify, the outbound data-path uses now alternatively ISP1 and ISP2 exit points, which is the objective of this lab.

For inbound data-path load-balancing, other mechanisms must be used (like AS path prepending), which is out of the scope of this current lab.

This is the end of this lab.

Appendix

BGP Decision tree algorithm

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 the 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
10	oldest eBGP route		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

Reference Configurations

If you face issues during your lab, you can verify your configuration with the configuration extract listed in this section.

SW1

```
hostname SW1
!
interface 1/1/1
    no shutdown
    ip address 192.168.4.4/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf network point-to-point
interface 1/1/8
    no shutdown
    ip address 192.168.4.0/31
interface 1/1/9
    no shutdown
    ip address 192.168.4.2/31
```

```
interface loopback 0
    ip address 192.168.2.1/32
ip prefix-list intranet-IP seq 10 permit 10.0.5.0/24
!
ip aspath-list intranet-AS seq 10 permit ^$
!
route-map isp1-out permit seq 10
     match ip address prefix-list intranet-IP
     match aspath-list intranet-AS
route-map isp2-out permit seq 10
     match ip address prefix-list intranet-IP
     match aspath-list intranet-AS
!
router ospf 1
    passive-interface default
    redistribute local loopback
    area 0.0.0.0
router bgp 65000
    bgp router-id 192.168.2.1
    bgp fast-external-fallover
    bgp log-neighbor-changes
    bgp bestpath as-path multipath-relax
    neighbor 192.168.2.2 remote-as 65000
    neighbor 192.168.2.2 description Core-RR
    neighbor 192.168.2.2 fall-over
    neighbor 192.168.2.2 update-source loopback 0
    neighbor 192.168.4.1 remote-as 65001
    neighbor 192.168.4.1 description ISP1 peering
    neighbor 192.168.4.3 remote-as 65002
    neighbor 192.168.4.3 description ISP2 peering
    address-family ipv4 unicast
        neighbor 192.168.2.2 activate
        neighbor 192.168.2.2 next-hop-self
        neighbor 192.168.4.1 activate
        neighbor 192.168.4.1 route-map isp1-out out
        neighbor 192.168.4.3 activate
        neighbor 192.168.4.3 route-map isp2-out out
    exit-address-family
```

SW2

```
hostname SW2
!
vlan 1,5
!
interface 1/1/1
    no shutdown
    no routing
    vlan access 5
interface 1/1/8
    no shutdown
    ip address 192.168.4.5/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf network point-to-point
interface loopback 0
    ip address 192.168.2.2/32
interface vlan 5
    ip address 10.0.5.1/24
!
router ospf 1
    passive-interface default
    redistribute local loopback
    area 0.0.0.0
router bgp 65000
    bgp router-id 192.168.2.2
    bgp log-neighbor-changes
    neighbor 192.168.2.1 remote-as 65000
    neighbor 192.168.2.1 fall-over
    neighbor 192.168.2.1 update-source loopback 0
    address-family ipv4 unicast
        neighbor 192.168.2.1 activate
```

```
        neighbor 192.168.2.1 route-reflector-client
        network 10.0.5.0/24
    exit-address-family
```

SW3

```
hostname SW3
!
interface 1/1/1
    no shutdown
    ip address 192.168.4.1/31
interface 1/1/2
    no shutdown
    ip address 192.168.4.17/31
interface loopback 0
    ip address 192.168.11.1/32
!
router bgp 65001
    bgp router-id 192.168.11.1
    bgp fast-external-fallover
    bgp log-neighbor-changes
    neighbor 192.168.4.0 remote-as 65000
    neighbor 192.168.4.16 remote-as 65005
    address-family ipv4 unicast
        neighbor 192.168.4.0 activate
        neighbor 192.168.4.16 activate
    exit-address-family
```

SW4

```
hostname SW4
!
interface 1/1/1
    no shutdown
    ip address 192.168.4.3/31
interface 1/1/2
    no shutdown
    ip address 192.168.4.19/31
interface loopback 0
    ip address 192.168.12.1/32
!
router bgp 65002
    bgp router-id 192.168.12.1
    bgp fast-external-fallover
    bgp log-neighbor-changes
    neighbor 192.168.4.2 remote-as 65000
    neighbor 192.168.4.18 remote-as 65005
    address-family ipv4 unicast
        neighbor 192.168.4.2 activate
        neighbor 192.168.4.18 activate
    exit-address-family
```

SW5

```
hostname SW5
!
vlan 1,9
interface 1/1/1
    no shutdown
    ip address 192.168.4.16/31
interface 1/1/2
    no shutdown
    ip address 192.168.4.18/31
interface 1/1/8
    no shutdown
    no routing
    vlan access 9
interface loopback 0
    ip address 192.168.15.1/32
interface vlan 9
    ip address 10.0.9.1/24
```

```
!
router bgp 65005
    bgp router-id 192.168.15.1
    bgp fast-external-fallover
    bgp log-neighbor-changes
    neighbor 192.168.4.17 remote-as 65001
    neighbor 192.168.4.19 remote-as 65002
    address-family ipv4 unicast
        neighbor 192.168.4.17 activate
        neighbor 192.168.4.19 activate
        network 10.0.9.0/24
    exit-address-family
```

a Hewlett Packard Enterprise company

www.arubanetworks.com

3333 Scott Blvd. Santa Clara, CA 95054
1.844.472.2782 | T: 1.408.227.4500 | FAX: 1.408.227.4550 | info@arubanetworks.com