ACI Deep Dive

L3 out troubleshooting

ACI Solution TAC team

Roland Ducomble – Technical Leader EMEAR TAC - CCIE 3745

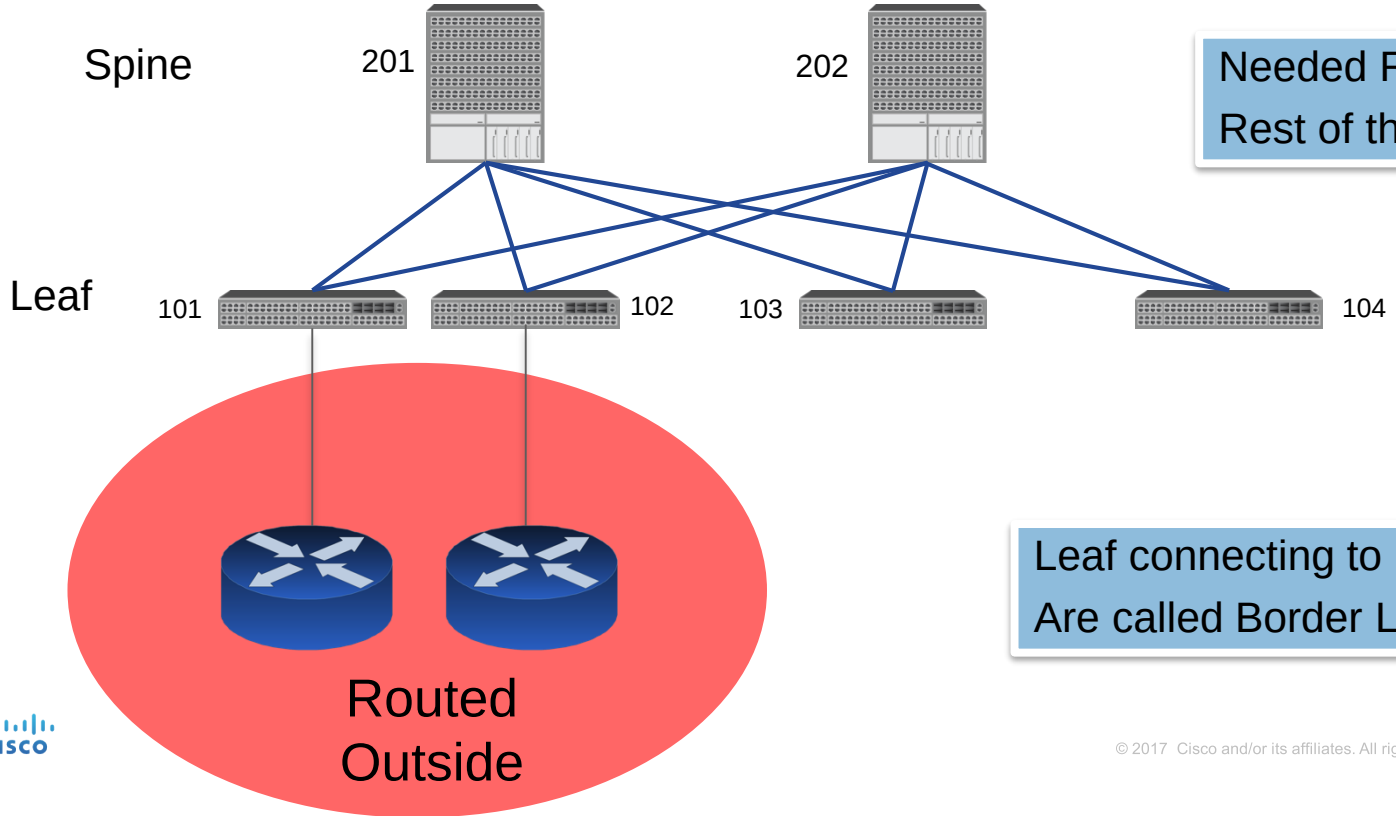
July 2018

V4.6

L3 out – Basic and iBGP inside Fabric

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

External L3 out

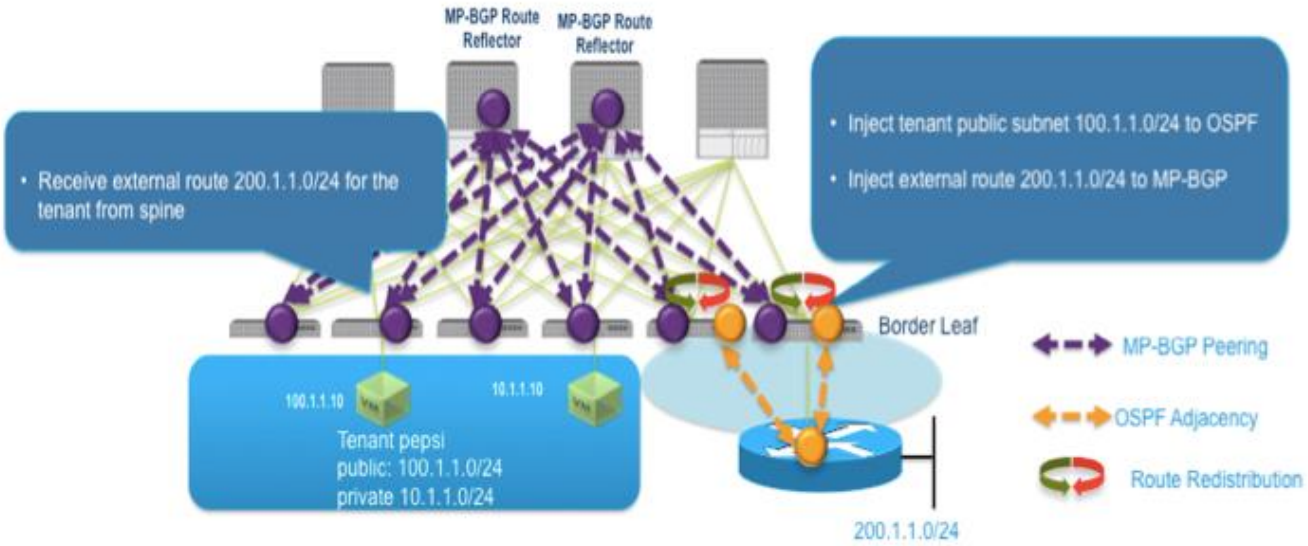


Needed Rx/Tx routes from
Rest of the network

Leaf connecting to router
Are called Border Leaf

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Route Learning Overview

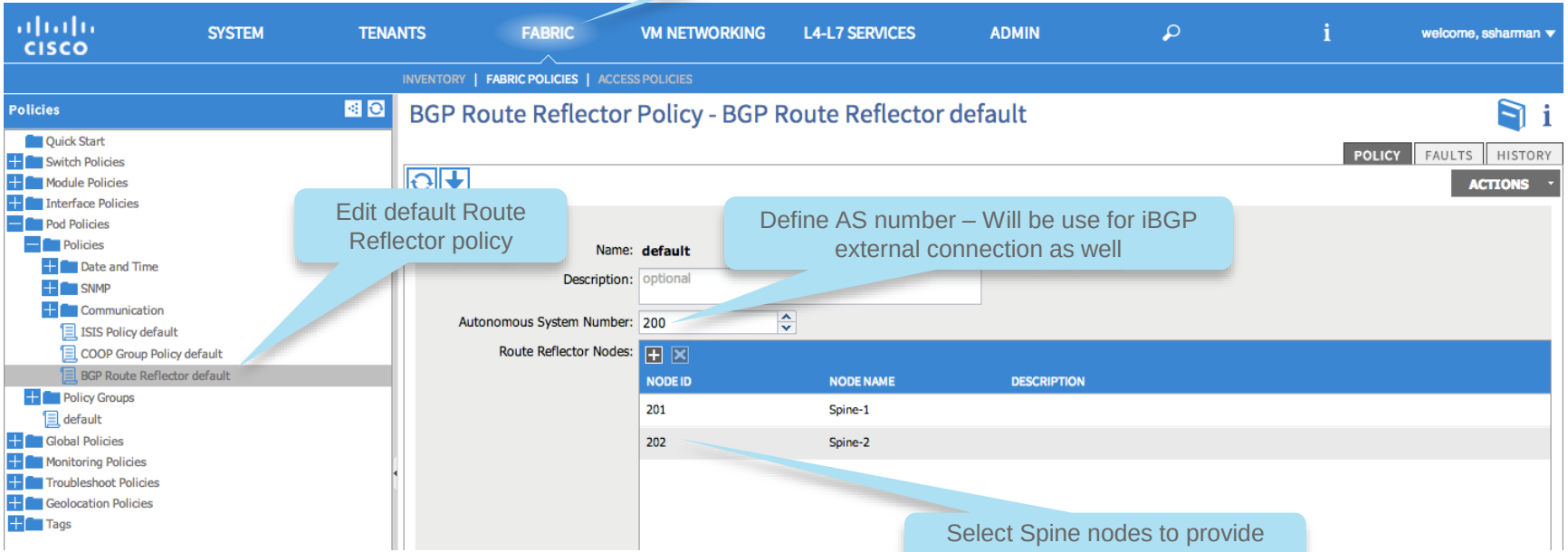


- L3 out peering to external routers (ospf, ibgp or static) on border leaf
- Ext route in border leaf (show ip route vrf tenant:ctx)
- For non Border leaf to know about external prefix we need a ibgp with RR (RR are spines)
  - See that as exchange of vpnv4 in an mpls network ☺
- Ext routes seen in non border leaf

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Define Route Reflectors

Go In Fabric – Fabric Policies for BGP RR



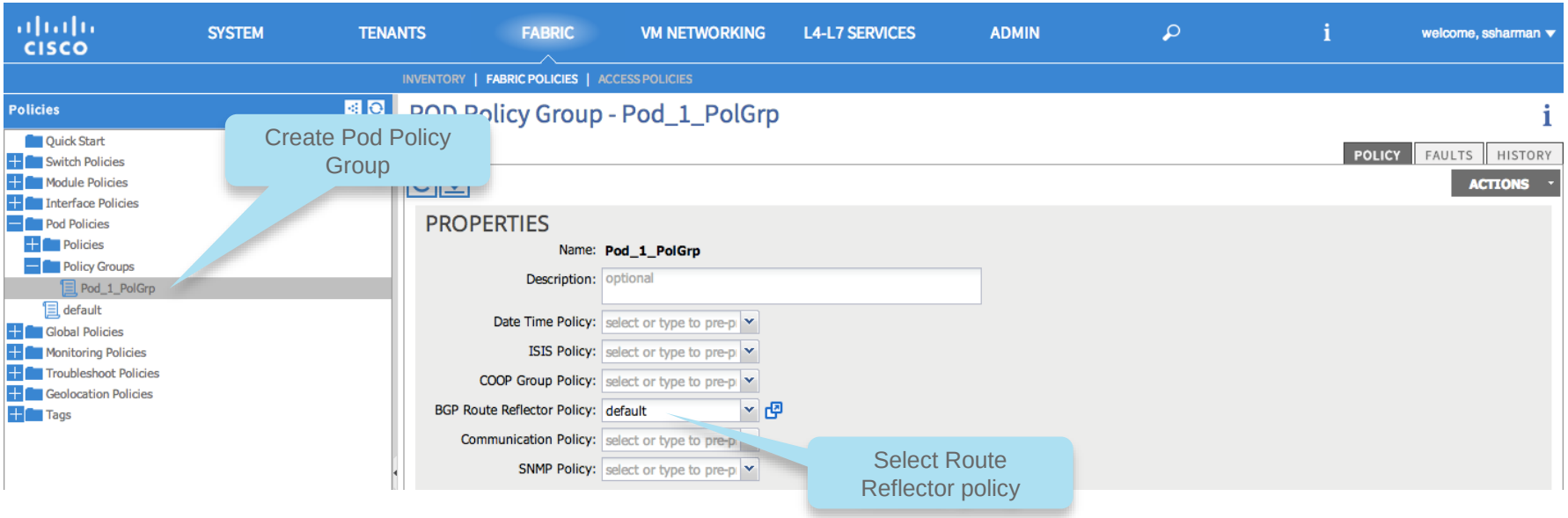
Edit default Route Reflector policy

Define AS number – Will be use for iBGP external connection as well

Select Spine nodes to provide Route Reflector functionality

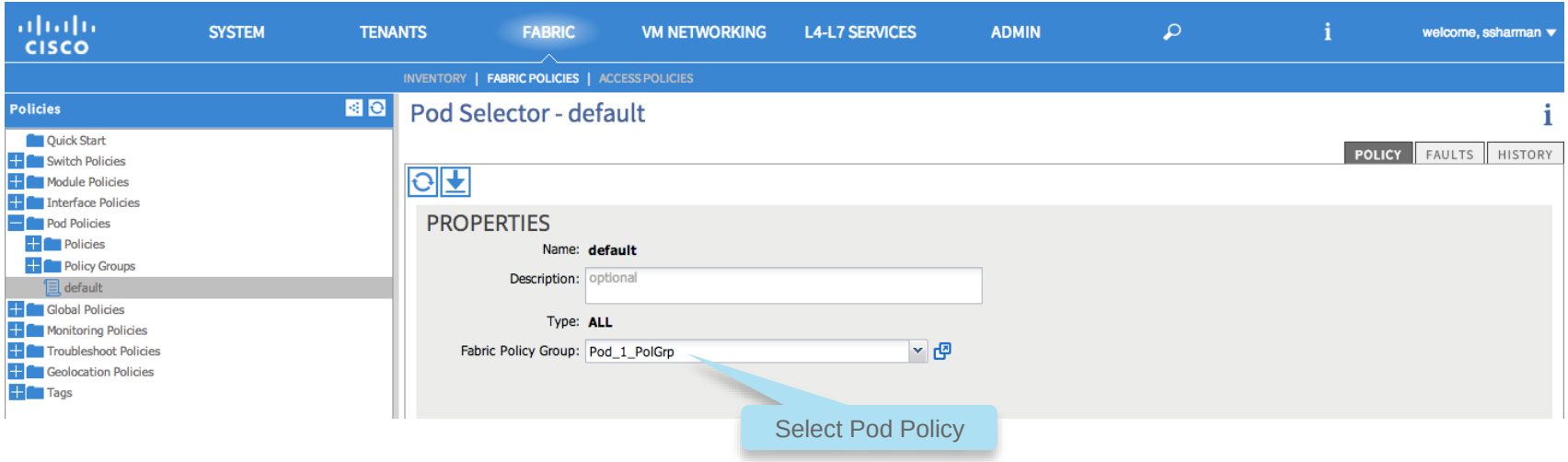
© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Create Policy Group for Pod



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Apply Policy Group to Pod



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Verify BGP vpnv4 session are up

```
bdsol-aci32-spine2# show bgp vpnv4 unicast summary vrf overlay-1
BGP summary information for VRF overlay-1, address family VPNv4 Unicast
BGP router identifier 10.10.32.212, local AS number 132
BGP table version is 418, VPNv4 Unicast config peers 10, capable peers 8
144 network entries and 161 paths using 28632 bytes of memory
BGP attribute entries [108/15552], BGP AS path entries [1/6]
BGP community entries [0/0], BGP clusterlist entries [4/24]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.0.80.94      4   132   12597   12795      418    0    0     1w1d 24
10.0.88.65      4   132   12577   12785      418    0    0     1w1d 0
10.0.88.90      4   132   12619   13056      418    0    0     1w1d 21
10.0.88.91      4   132   12613   12904      418    0    0     1w1d 27
10.0.88.95      4   132   12615   12873      418    0    0     1w1d 27
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

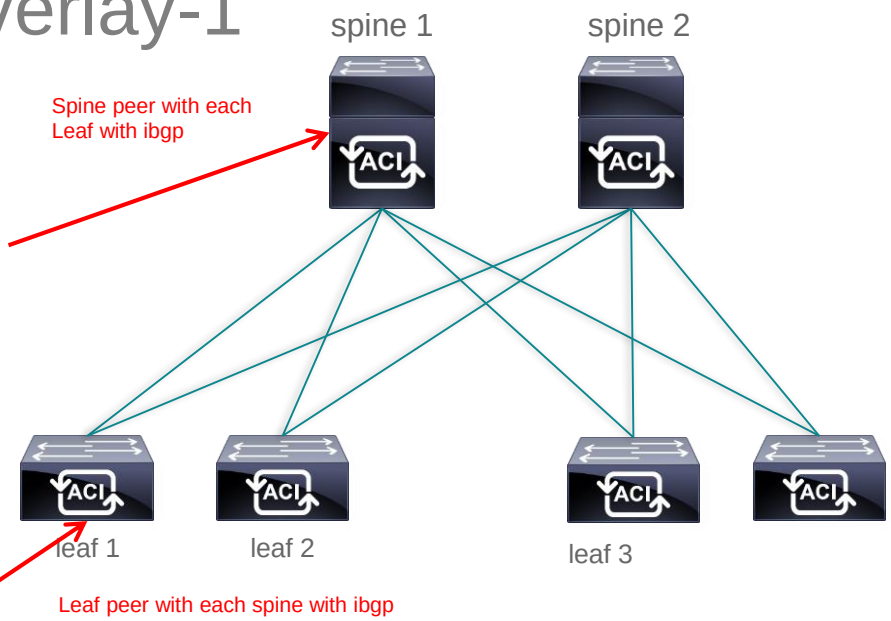
L3 out – Example 1
eBGP between ACI and ext Router
Configuration and troubleshooting

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Prereq - Verifying iBGP in overlay-1

```
pod2-spine1# show bgp session vrf overlay-1
Total peers 4, established peers 4
ASN 101
VRF overlay-1, local ASN 101
peers 4, established peers 4, local router-id 10.0.168.92
State: I-Idle, A-Active, O-Open, E-Established, C-Closing, S-Shutdown

Neighbor        ASN    Flaps LastUpDn|LastRead|LastWrit St Port(L/R)
Notif(S/R)
10.0.168.91      101 0     06w48d  |never   |never    E  179/56931  0/0
10.0.168.93      101 0     06w48d  |never   |never    E  179/44502  0/0
10.0.168.90      101 0     06w48d  |never   |never    E  179/42567  0/0
10.0.168.95      101 1     02w14d  |never   |never    E  179/35542  1/0
```

```
pod2-leaf1# show  bgp session vrf overlay-1
Total peers 2, established peers 2
ASN 101
VRF overlay-1, local ASN 101
peers 2, established peers 2, local router-id 10.0.168.95
State: I-Idle, A-Active, O-Open, E-Established, C-Closing, S-Shutdown

Neighbor        ASN    Flaps LastUpDn|LastRead|LastWrit St Port(L/R)
Notif(S/R)
10.0.168.94      101 0     02w14d  |never   |never    E  36966/179  0/0
10.0.168.92      101 0     02w14d  |never   |never    E  35542/179  0/0
pod2-leaf1#
```



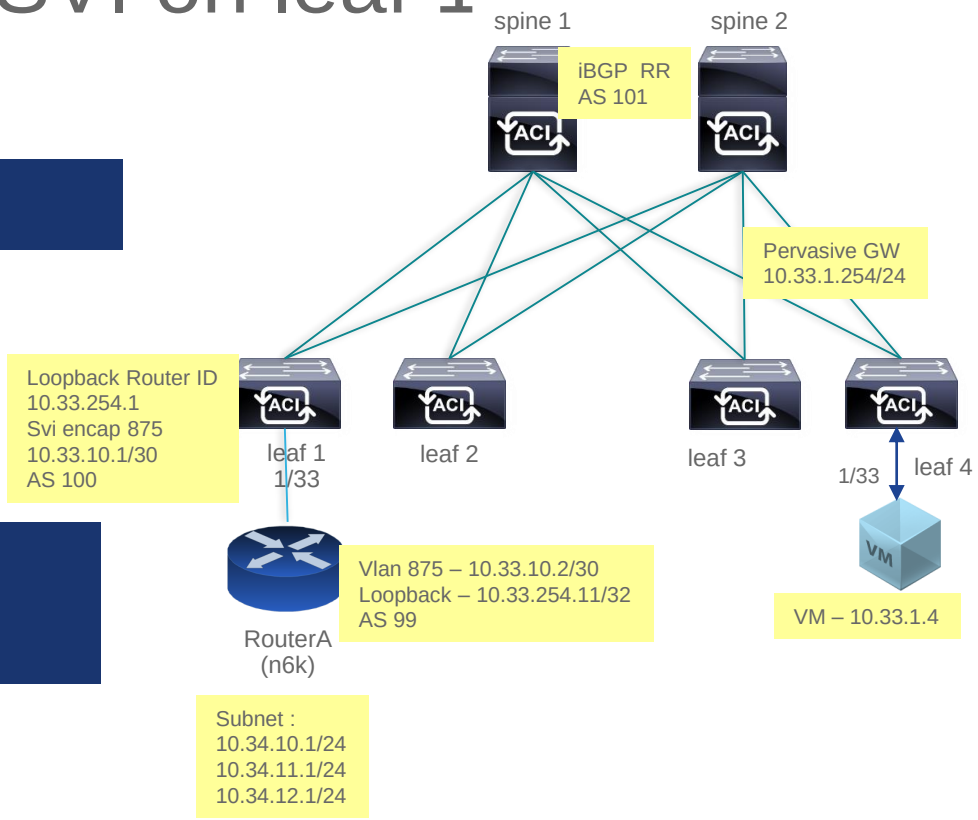

```
pod2-leaf1# acidiag fnvread
      ID              Name     Serial Number         IP Address     Role        State    LastUpdMsgId
--------------------------------------------------------------------------------------------------------
     101        pod2-leaf1      SAL1820SMHV     10.0.168.95/32    leaf       active    0
     102        pod2-leaf2      SAL1816QVBC     10.0.168.93/32    leaf       active    0
     103        pod2-leaf3      SAL1818RUHM     10.0.168.91/32    leaf       active    0
     104        pod2-leaf4      SAL1818RP59     10.0.168.90/32    leaf       active    0
     201       pod2-spine1      SAL1811NN5S     10.0.168.92/32   spine       active    0
     202       pod2-spine2      SAL1811NN5X     10.0.168.94/32   spine       active    0
```

Lab setup – L3 out on SVI on leaf 1 using eBGP

All in tenant L3, context L3
VM in EPG1 part of BD1 subnet 10.33.1.254/24 on leaf4

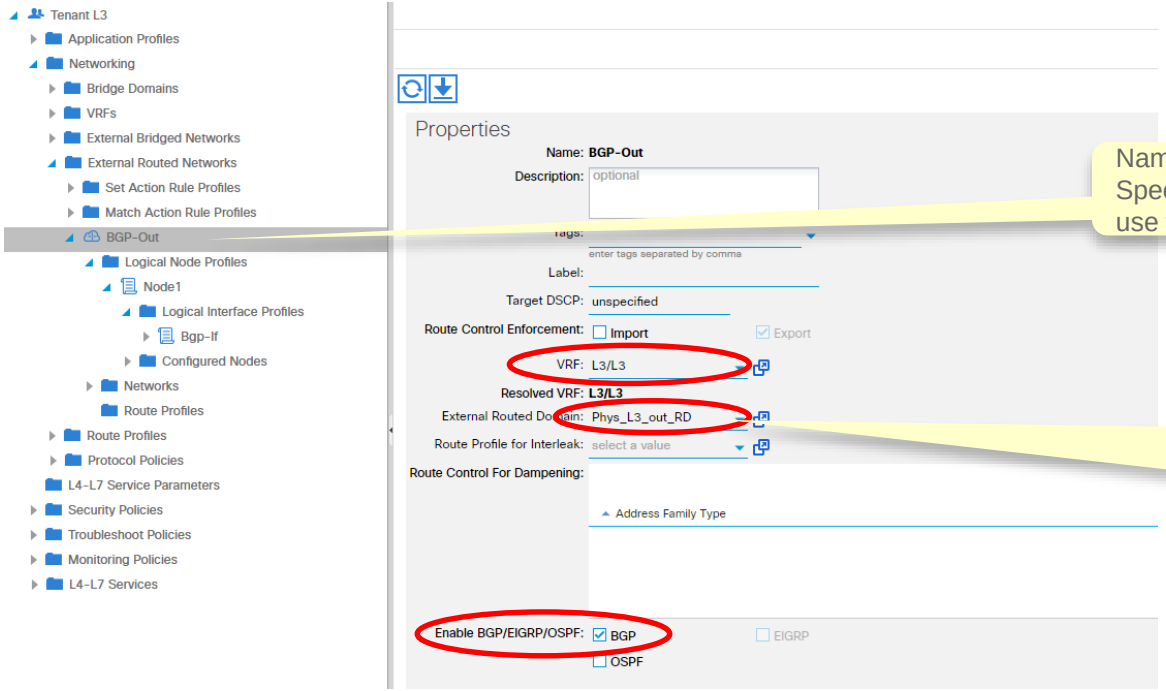
Goal setup eBGP between leaf1 and external router
AS in ACI will be 100, AS in n6k will be 99
Get route exchange in both direction
Ping from 10.33.1.4 to 10.34.10.1

Software used is 1.2(2g)



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 1a – L3 out + Node + if + BGP peer + L3 EPG

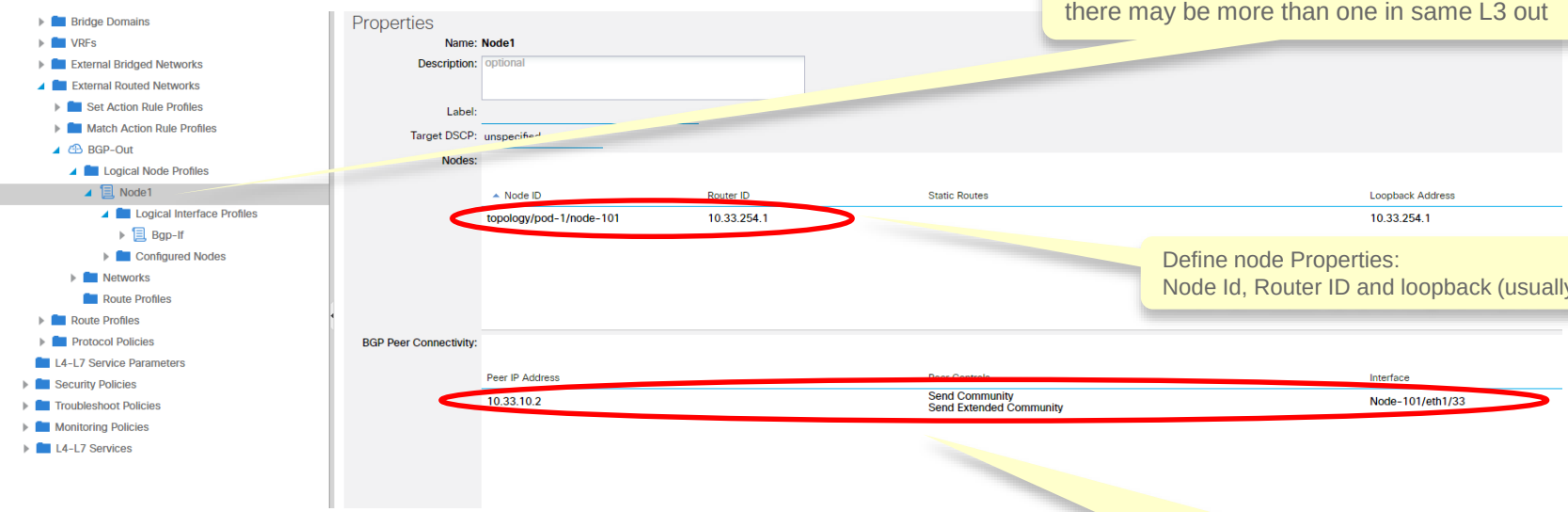


Name of the L3 out extension
Specify the ctx and the ext Routed domain to use for it

The L3 must be bound to a Ext Routed domain (define in Fabric/access-policy).
That domain must be in AEP attached to the port used for L3
There may need a vlan pool (if svi) but might not need pool if routed IF

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 1b – L3 out + Node + if + BGP peer



List of Node (Border leaf), here only one but there may be more than one in same L3 out

Define node Properties:
Node Id, Router ID and loopback (usually the same)

Define BGP peer , Address and interface

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 1b –BGP peer property

Peer Connectivity Profile - BGP Peer Connectivity Profile 10.33.10.2

Properties

Address: **10.33.10.2**

Description: optional

BGP Controls:
- [ ] Allow Self AS
- [ ] Disable Peer AS Check
- [ ] Next-hop Self
- [x] Send Community
- [x] Send Extended Community

CHECK ALL | UNCHECK ALL

Password:

Confirm Password:

Allowed Self AS Count: 3

Peer Controls:
- [ ] Bidirectional Forwarding Detection
- [ ] Disable Connected Check

EBGP Multihop TTL: 1

Weight for routes from this neighbor: 0

Private AS Control:
- [ ] Remove all private AS
- [ ] Remove private AS
- [ ] Replace private AS with local AS

BGP Peer Prefix Policy: select a value

Remote Autonomous System Number: 99

Local-AS Number Config: no options

Local-AS Number: 100

This value must not match the MP-BGP RR policy

You can select a series of usual BGP peer property

Remote AS is 99

We want to present outself as AS 100 to external world, using local-as feature (without local-as we would use the iBGP AS 101 configured in ninitial fabric setup

Step 1c – L3 out + Node + if + BGP peer

- Networking
  - Bridge Domains
  - VRFs
  - External Bridged Networks
  - External Routed Networks
    - Set Action Rule Profiles
    - Match Action Rule Profiles
    - BGP-Out
      - Logical Node Profiles
        - Node1
          - Logical Interface Profiles
            - Bgp-If
              - BGP Peer Connectivity Profile 1...
          - Configured Nodes
        - Networks
        - Route Profiles
  - Route Profiles
  - Protocol Policies
  - L4-L7 Service Parameters
- Security Policies
- Troubleshoot Policies
- Monitoring Policies
- L4-L7 Services

Properties

Name: Bgp-If

Description: optional

Label:

ND policy: select a value

Egress Data Plane Policing Policy: select a value

Ingress Data Plane Policing Policy: select a value

Routed Interfaces:

Path	IP Address	MAC Address	MTU (Bytes)
No items have been found. Select Actions to create a new item.			

SVI:

Path	IP Address	Side A IP	Side B IP	MAC Address	MTU (Bytes)	Encap
Node-101/eth1/33	10.33.10.1/30			00:22:BD:F8:19:FF	inherit	vlan-875

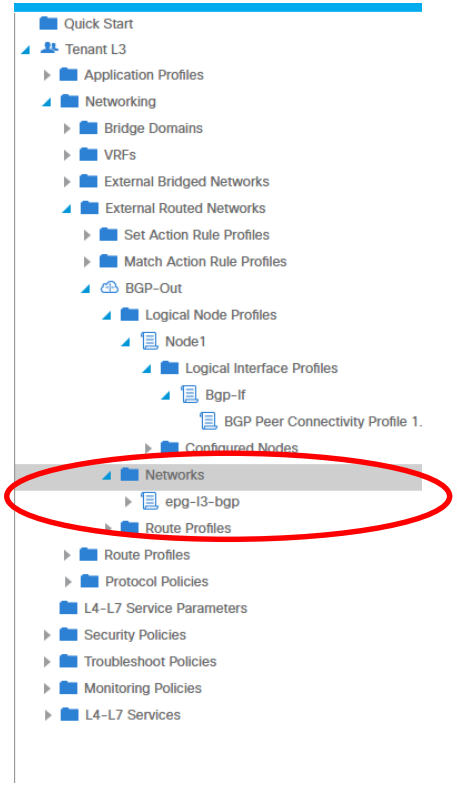
Routed Sub-Interfaces:

Path	IP Address	MAC Addr	MTU (Bytes)	Encap

Define the Interface on L3 out property.
Here SVI on 1/33 with encap vlan 875

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 1d – Creating L3 out EPG



L3 out EPG is under Networks of the L3 out.
Here I just create it with a name and no property in it.
THIS IS MANDATORY to get anything deploy the leaf (ip,..).
Like regular Epg, this is adding an interface in the epg that push the info the leaf , here all interface in the l3 out (1/33 on leaf1) are implicitely part of the epg

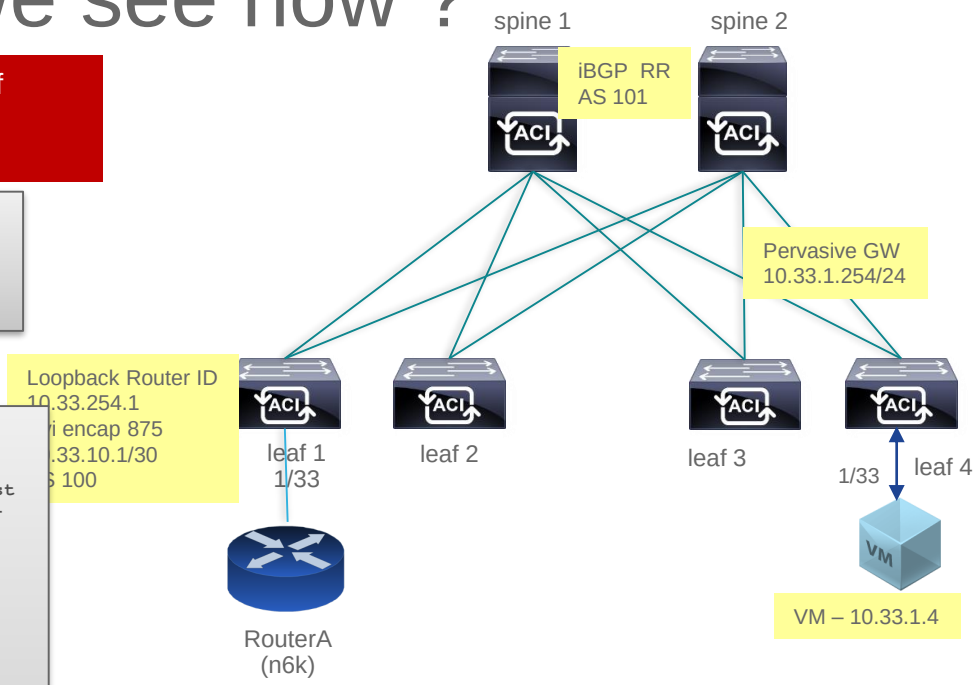
```
pod2-apic1# moquery -d  uni/tn-L3/out-BGP-Out/instP-epg-l3-bgp
Total Objects shown: 1

# l3ext.InstP                                          !! Class of the object
name          : epg-l3-bgp
childAction   :
configIssues  :
configSt      : applied
descr         :
dn            : uni/tn-L3/out-BGP-Out/instP-epg-l3-bgp    !! Dn of the l3 epg
lcOwn         : local
matchT        : AtleastOne
modTs         : 2016-04-07T22:33:09.733+02:00
monPolDn      : uni/tn-common/monepg-default
pcTag         : 32775                                      !! Epg id in the vrf
prio          : unspecified
rn            : instP-epg-l3-bgp
scope         : 2785280                                     !! Vrf number (vnid)
status        :
targetDscp    : unspecified
triggerSt     : triggerable
txId          : 10376293541467381302
uid           : 15374
```

Step 1 – What should we see now ?

Interface vlan and ip address should have been pushed to the leaf
BGP should be started and trying to initiate connection to Peer.
Our local prefix should be in BGP table

```
pod2-leaf1# show ip interface brief vrf L3:L3
IP Interface Status for VRF "L3:L3"(15)
Interface            Address            Interface Status
vlan65               10.33.10.1/30      protocol-up/link-up/admin-up
lo7                  10.33.254.1/32     protocol-up/link-up/admin-up
```

```
pod2-leaf1# show ip bgp vrf L3:L3
BGP routing table information for VRF L3:L3, address family IPv4 Unicast
BGP table version is 8, local router ID is 10.33.254.1
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-
injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

   Network          Next Hop            Metric     LocPrf     Weight Path
*>r10.33.10.0/30    0.0.0.0                  0        100      32768 ?
*>r10.33.254.1/32   0.0.0.0                  0        100      32768 ?

pod2-leaf1# show ip bgp vrf L3:L3 summary
BGP summary information for VRF L3:L3, address family IPv4 Unicast
BGP router identifier 10.33.254.1, local AS number 101
BGP table version is 8, IPv4 Unicast config peers 1, capable peers 0
2 network entries and 2 paths using 272 bytes of memory
BGP attribute entries [2/272], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [6/24]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.33.10.2      4    99       0       0        0    0    0 00:01:47 Active
```



Note :
We do not have yet the prefix for the BD VM 10.33.1.0/24
BGP session is not established as remote peer is not configured yet

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 2 – configure Interface and bgp on remote route

```
interface loopback0
  vrf member L3:L3
  ip address 10.33.254.11/32

interface Vlan875
  no shutdown
  vrf member L3:L3
  ip address 10.33.10.2/30

interface Vlan2300
  no shutdown
  vrf member L3:L3
  ip address 10.34.10.1/24

interface Vlan2301
  no shutdown
  vrf member L3:L3
  ip address 10.34.11.1/24

interface Vlan2302
  no shutdown
  vrf member L3:L3
  ip address 10.34.12.1/24
```

```
vrf context L3:L3
router bgp 99
  vrf L3:L3
    router-id 10.33.254.11
    address-family ipv4 unicast
      network 10.34.10.0/24
      network 10.34.11.0/24
    neighbor 10.33.10.1 remote-as 100
      address-family ipv4 unicast
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 2 – What should we see now?

```
pod2-leaf1#  iping -V L3:L3 10.33.10.2
PING 10.33.10.2 (10.33.10.2) from 10.33.10.1: 56 data bytes
64 bytes from 10.33.10.2: icmp_seq=0 ttl=255 time=0.735 ms
64 bytes from 10.33.10.2: icmp_seq=1 ttl=255 time=0.646 ms
64 bytes from 10.33.10.2: icmp_seq=2 ttl=255 time=0.78 ms
```

Ping across L3 out should work

```
pod2-leaf1# show ip bgp vrf L3:L3  summary
BGP summary information for VRF L3:L3, address family IPv4 Unicast
BGP router identifier 10.33.254.1, local AS number 101
BGP table version is 11, IPv4 Unicast config peers 1, capable peers 1
4 network entries and 4 paths using 544 bytes of memory
BGP attribute entries [3/408], BGP AS path entries [1/10]
BGP community entries [0/0], BGP clusterlist entries [6/24]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.33.10.2      4    99       5       4       11    0    0 00:00:26 2
pod2-leaf1#
pod2-leaf1#
pod2-leaf1# show ip bgp vrf L3:L3
BGP routing table information for VRF L3:L3, address family IPv4 Unicast
BGP table version is 11, local router ID is 10.33.254.1
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

   Network          Next Hop            Metric     LocPrf     Weight Path
*>r10.33.10.0/30    0.0.0.0                  0        100      32768 ?
*>r10.33.254.1/32   0.0.0.0                  0        100      32768 ?
*>e10.34.10.0/24    10.33.10.2                                     0 100 99 i
*>e10.34.11.0/24    10.33.10.2                                     0 100 99 i
```

BGP should be up and we should receive prefix from NXOS router

Note most BGP command must be done under vsh

Note: we do not have yet the BD prefix 10.33.1.0/24 in our BGP table → so it won't be send to remote router

Step 2 – What should we see now? (cont)

```
pod2-leaf4# show bgp vpnv4 unicast   vrf L3:L3
BGP routing table information for VRF overlay-1, address family VPNv4 Unicast
BGP table version is 812, local router ID is 10.0.56.95
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

   Network            Next Hop            Metric     LocPrf     Weight Path
Route Distinguisher: 10.0.56.95:10    (VRF L3:L3)
*>i10.33.10.0/30      10.0.168.95              0        100          0 ?
*>i10.33.254.1/32     10.0.168.95              0        100          0 ?
*>i10.34.10.0/24      10.0.168.95                       100          0 100 99 i
*>i10.34.11.0/24      10.0.168.95                       100          0 100 99 i
```

Route from external router are injected in vpnv4 bgp and propagated through ibgp to leaf4

(now I am on remote non Border leaf)

```
pod2-leaf4# show ip route vrf L3:L3
IP Route Table for VRF "L3:L3"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

..
10.34.10.0/24, ubest/mbest: 1/0
    *via 10.0.168.95%overlay-1, [200/0], 00:15:03, bgp-101, internal, tag 100
10.34.11.0/24, ubest/mbest: 1/0
    *via 10.0.168.95%overlay-1, [200/0], 00:15:03, bgp-101, internal, tag 100
```

Routing table on leaf4 is populated with remote subnet

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 2 – What if BGP does not come up ?

Do we have TCP session established for BGP ?

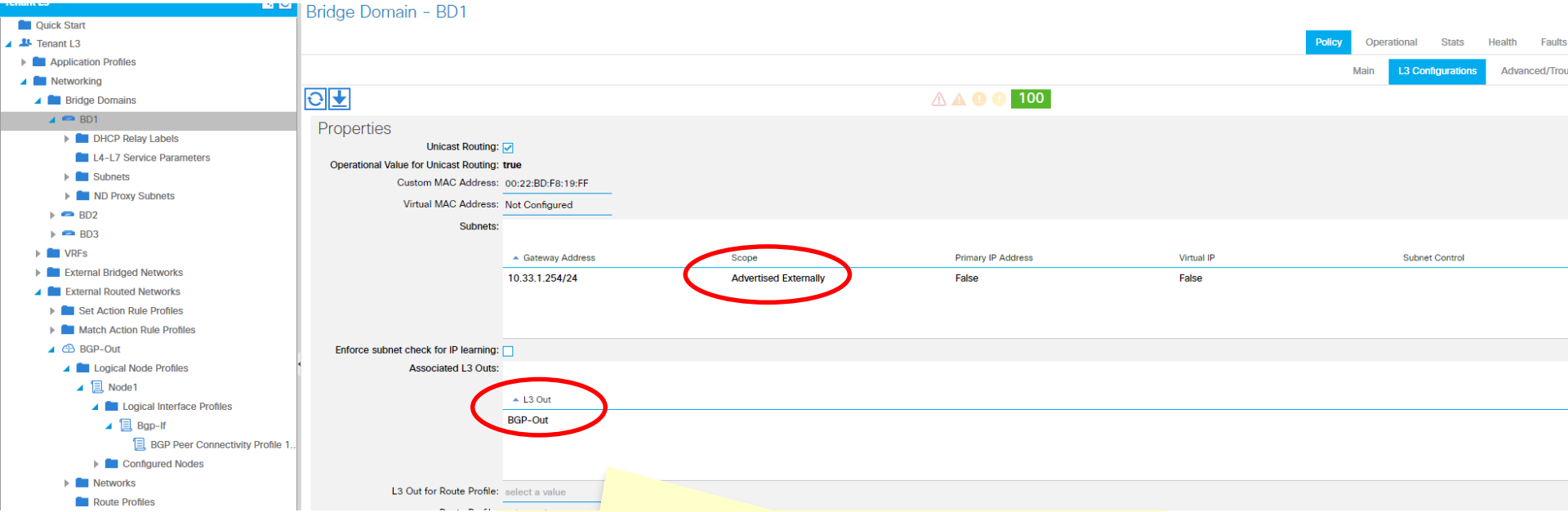
```
pod2-leaf1# netstat -tupn | egrep 179
(No  info could be read for "-p": geteuid()=15374 but you should be root.)
tcp        0      0 10.33.10.1:179          10.33.10.2:62415        ESTABLISHED -
tcp        0      0 10.0.168.95:44599       10.0.168.94:179         ESTABLISHED -
tcp        0      0 10.0.168.95:57261       10.0.168.92:179         ESTABLISHED -
```

Do we receive / transmit BGP packets ?

```
pod2-leaf1# tcpdump -i kpm_inb -f port 179
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on kpm_inb, link-type EN10MB (Ethernet), capture size 65535 bytes
22:38:59.188850 IP 10.33.10.2.62415 > 10.33.10.1.bgp: Flags [P.], seq 2891078262:2891078281,
ack 1176057492, win 17376, options [nop,nop,TS val 172217473 ecr 506104871], length 19: BGP,
length: 19
22:38:59.188921 IP 10.33.10.1.bgp > 10.33.10.2.62415: Flags [.], ack 19, win 145, options
[nop,nop,TS val 506113132 ecr 172217473], length 0
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 3 – Advertise our BD to external router



In the BD config (L3 configurations part) you need to make sure the BD subnet is marked as Advertised Externally
And that the L3 out we create in step 1 is bound to the BD

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

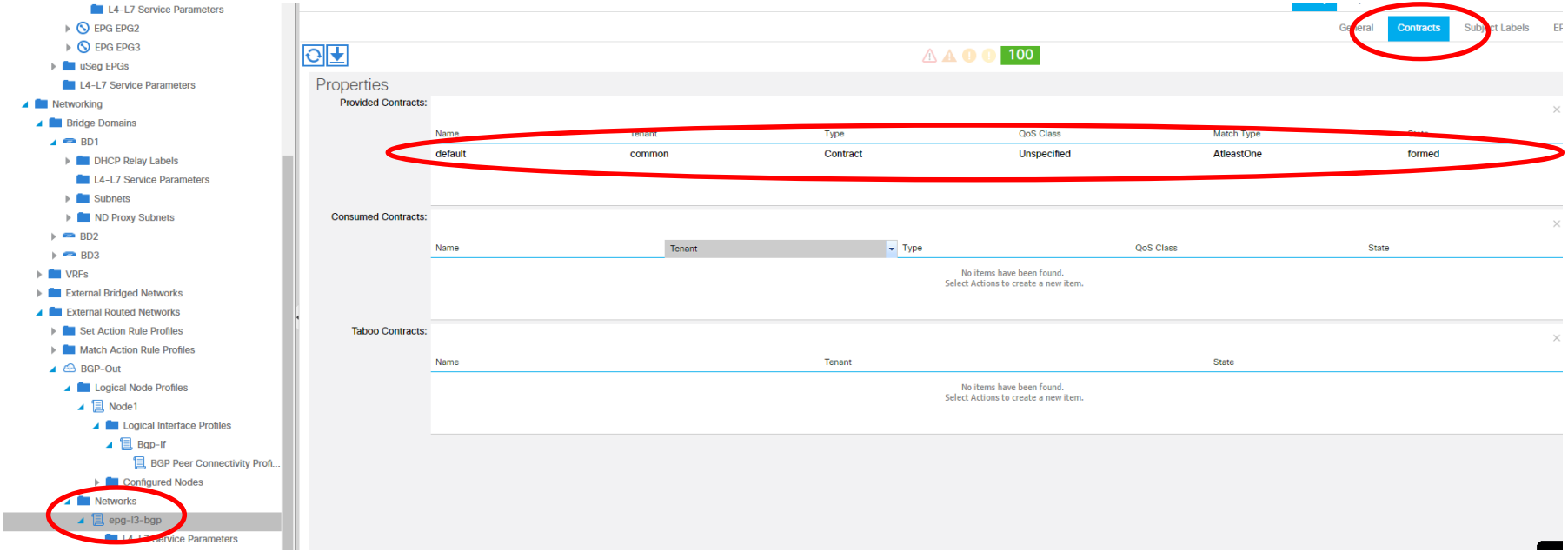
Step 3 – What should we see now?

```
pod2-leaf1# show ip route vrf L3:L3 10.33.1.0
IP Route Table for VRF "L3:L3"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

Route not found
```

No real changes – BD subnet 10.33.1.0/24 is still not send to leaf1 through
And of course not pushed externally..

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 4 – Adding contract between L3 out EPG and VM EPG



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 4 – What should we see now?

```
pod2-leaf1# show ip route vrf L3:L3 10.33.1.0
IP Route Table for VRF "L3:L3"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

10.33.1.0/24, ubest/mbest: 1/0, attached, direct, pervasive
    *via 10.0.232.65%overlay-1, [1/0], 00:03:58, static
```

```
pod2-leaf1# show ip bgp vrf L3:L3
BGP routing table information for VRF L3:L3, address family IPv4 Unicast
BGP table version is 31, local router ID is 10.33.254.1
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

   Network          Next Hop            Metric     LocPrf     Weight Path
*>r10.33.1.0/24     0.0.0.0                  0        100      32768 ?
*>r10.33.2.0/24     0.0.0.0                  0        100      32768 ?
*>r10.33.3.0/24     0.0.0.0                  0        100      32768 ?
*>r10.33.10.0/30    0.0.0.0                  0        100      32768 ?
*>r10.33.254.1/32   0.0.0.0                  0        100      32768 ?
*>e10.34.10.0/24    10.33.10.2                                     0 100 99 i
*>e10.34.11.0/24    10.33.10.2                                     0 100 99 i
```

Now route for BD subnet is on leaf 1
And it appears in our BGP table

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 4 – What should we see now?

```
bdsol-n6004-01# show ip bgp vrf L3:L3
BGP routing table information for VRF L3:L3, address family IPv4 Unicast
BGP table version is 14, local router ID is 10.33.254.11
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath

   Network            Next Hop            Metric     LocPrf     Weight Path
*>e10.33.1.0/24       10.33.10.1               0                     0 100 101 ?
*>e10.33.2.0/24       10.33.10.1               0                     0 100 101 ?
*>e10.33.3.0/24       10.33.10.1               0                     0 100 101 ?
*>110.34.10.0/24      0.0.0.0                             100      32768 i
*>110.34.11.0/24      0.0.0.0                             100      32768 i
\bdsol-n6004-01# show ip route vrf L3:L3 10.33.1.0
IP Route Table for VRF "L3:L3"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

10.33.1.0/24, ubest/mbest: 1/0
    *via 10.33.10.1, [20/0], 00:14:02, bgp-99, external, tag 100
```

External router does now have bgp prefix and the route is learned.
Control is now fine end to end !!!

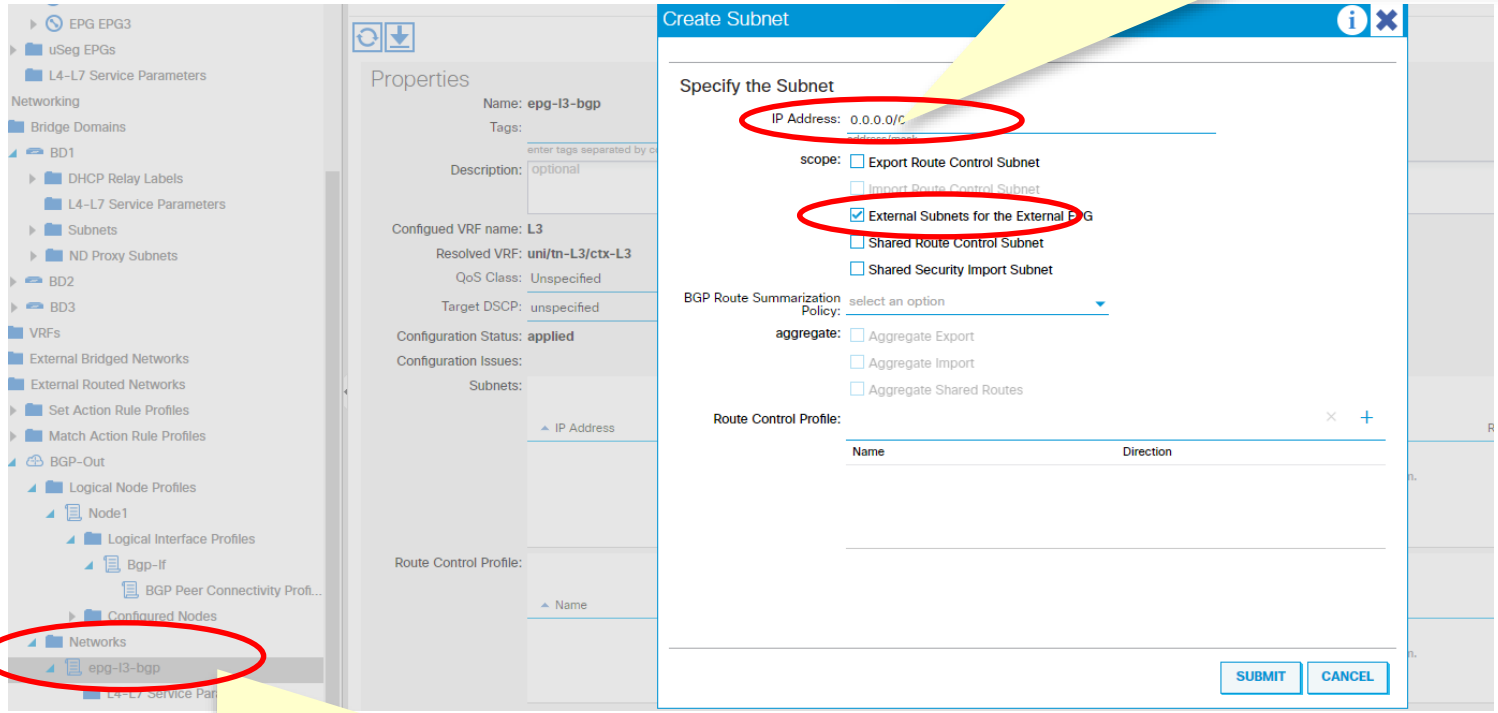
```
bdsol-n6004-01# ping 10.33.1.4 source 10.34.10.1 vrf L3:L3
PING 10.33.1.4 (10.33.1.4) from 10.34.10.1: 56 data bytes
Request 0 timed out
Request 1 timed out
```

Ping still fails !!!

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 5 – Mark external subnet for External EPG

You can either add 0.0.0.0/0 (most common scenario) – This will allow any subnet received from external router to communicate with internal sunet
Or you specify a subset like 10.34.0.0/16 or 10.34.1.0/24 for example



Under External Networks (ext l3 out epg) you must add the External Subnet (aka from external router) as – ‘external Subnets for External EPG’

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 5 – What should we see now?

```
bdsol-n6004-01# ping 10.33.1.4 source 10.34.10.1 vrf L3:L3
PING 10.33.1.4 (10.33.1.4) from 10.34.10.1: 56 data bytes
64 bytes from 10.33.1.4: icmp_seq=0 ttl=252 time=1.024 ms
64 bytes from 10.33.1.4: icmp_seq=1 ttl=252 time=0.75 ms
64 bytes from 10.33.1.4: icmp_seq=2 ttl=252 time=0.888 ms
64 bytes from 10.33.1.4: icmp_seq=3 ttl=252 time=0.816 ms
64 bytes from 10.33.1.4: icmp_seq=4 ttl=252 time=0.748 ms
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

How to trace what BGP does ?

A lot of bgp event-traces available

```
pod2-leaf1# show bgp event-history detail | egrep "10.34.10.0"

2016 Apr  7 21:48:39.116315 bgp 101 [9488]: TID 9518: (14) RIB: dest: 10.34.10.0/24 rnh: 10.33.10.2/32 [4] path#1: set to rid=0xa21fe0b,
flags=0x7, changed=1
2016 Apr  7 21:48:39.116311 bgp 101 [9488]: TID 9518: (14) RIB: dest: 10.34.10.0/24 [4] Begin select adv_path adv_all=0, cal_nth=0,
install_to_rib=0, flags=0x88d0000,0x100000
2016 Apr  7 21:48:39.116308 bgp 101 [9488]: TID 9518: (15) RIB: dest: 10.34.10.0/24 [0] Selected new bestpath flags=0x8880000 rid=0xa21fe0b
nh=10.33.10.2
2016 Apr  7 21:48:39.116303 bgp 101 [9488]: TID 9518: (15) RIB: dest: 10.34.10.0/24 [0] Updating path to dest in URIB, nhs added 1, deleted
0
2016 Apr  7 21:48:39.116299 bgp 101 [9488]: TID 9518: (15) RIB: dest: 10.34.10.0/24 rnh: 10.33.10.2/32 [0] Adding path (NH iod 0) in URIB
(table-id 0xe, flags 0x10) extcomm-len=0, preference=20
2016 Apr  7 21:48:39.116297 bgp 101 [9488]: TID 9518: (15) RIB: dest: 10.34.10.0/24 [0] Sending route to URIB
2016 Apr  7 21:48:39.116285 bgp 101 [9488]: TID 9518: (15) RIB: dest: 10.34.10.0/24 rnh: 10.33.10.2/32 [0] path#1: set to rid=0xa21fe0b,
flags=0x6, changed=1
2016 Apr  7 21:48:39.116272 bgp 101 [9488]: TID 9518: (15) RIB: dest: 10.34.10.0/24 [0] Begin select bestpath adv_all=0, cal_nth=0,
install_to_rib=0, flags=0x8080000
2016 Apr  7 21:48:39.115903 bgp 101 [9488]: TID 9518: (15) EVT: dest: 10.34.10.0/24 [0] Removing prefix from new-list
2016 Apr  7 21:48:39.111063 bgp 101 [9488]: TID 9516:bgp_brib_add:2308: (L3:L3) BRIB: [IPv4 Unicast] (10.34.10.0/24 (10.33.10.2)): returning
from bgp_brib_add, new_path: 1, change: 1, undelete: 0, history: 0, for
ce: 1, (pflags=0x2020), reeval=0
2016 Apr  7 21:48:39.111048 bgp 101 [9488]: TID 9516: (15) RIB: dest: 10.34.10.0/24 [0] Triggering bestpath selection, flags=0x8080000
repopulate_rib: 0, is_bestpath_marker: 0
2016 Apr  7 21:48:39.110955 bgp 101 [9488]: TID 9516:bgp_brib_add:2058: (L3:L3) BRIB: [IPv4 Unicast] (10.34.10.0/24 (10.33.10.2)):
bgp_brib_add: handling nexthop
2016 Apr  7 21:48:39.110948 bgp 101 [9488]: TID 9516:bgp_brib_add:2024: (L3:L3) BRIB: [IPv4 Unicast] Created new path to 10.34.10.0/24 via
0.0.0.0 (pflags=0x0)
2016 Apr  7 21:48:39.110927 bgp 101 [9488]: TID 9516:bgp_brib_add:1846: (L3:L3) BRIB: [IPv4 Unicast] Installing prefix 10.34.10.0/24
(10.33.10.2) via 10.33.10.2  into BRIB
2016 Apr  7 21:48:39.110898 bgp 101 [9488]: TID 9516:bgp_af_eval_export_policy:402: (L3:L3) BRIB: [IPv4 Unicast] Prefix 10.34.10.0/24 from
peer 10.33.10.2, set export attr RT:101:2785280
2016 Apr  7 21:48:39.110880 bgp 101 [9488]: TID 9516:bgp_af_process_nlri:3874: (L3:L3) UPD: [IPv4 Unicast] Received prefix 10.34.10.0/24
from peer 10.33.10.2, origin 0, next hop 10.33.10.2, localpref 0, med 0
```

Bgp trace in file

```
bdsol-aci32-leaf1# ls -al *bgp*
-rw-rw-rw- 1 root root 90647413 Dec  6 08:36 bgp_trace.txt
lrwxrwxrwx 1 root root       44 Nov 29 08:00 ipv6mgr-api-bgp-283.log -> /var/sysmgr/mem_logs/ipv6mgr-api-
bgp-283.log
bdsol-aci32-leaf1# pwd
/var/sysmgr/tmp_logs
bdsol-aci32-leaf1#
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

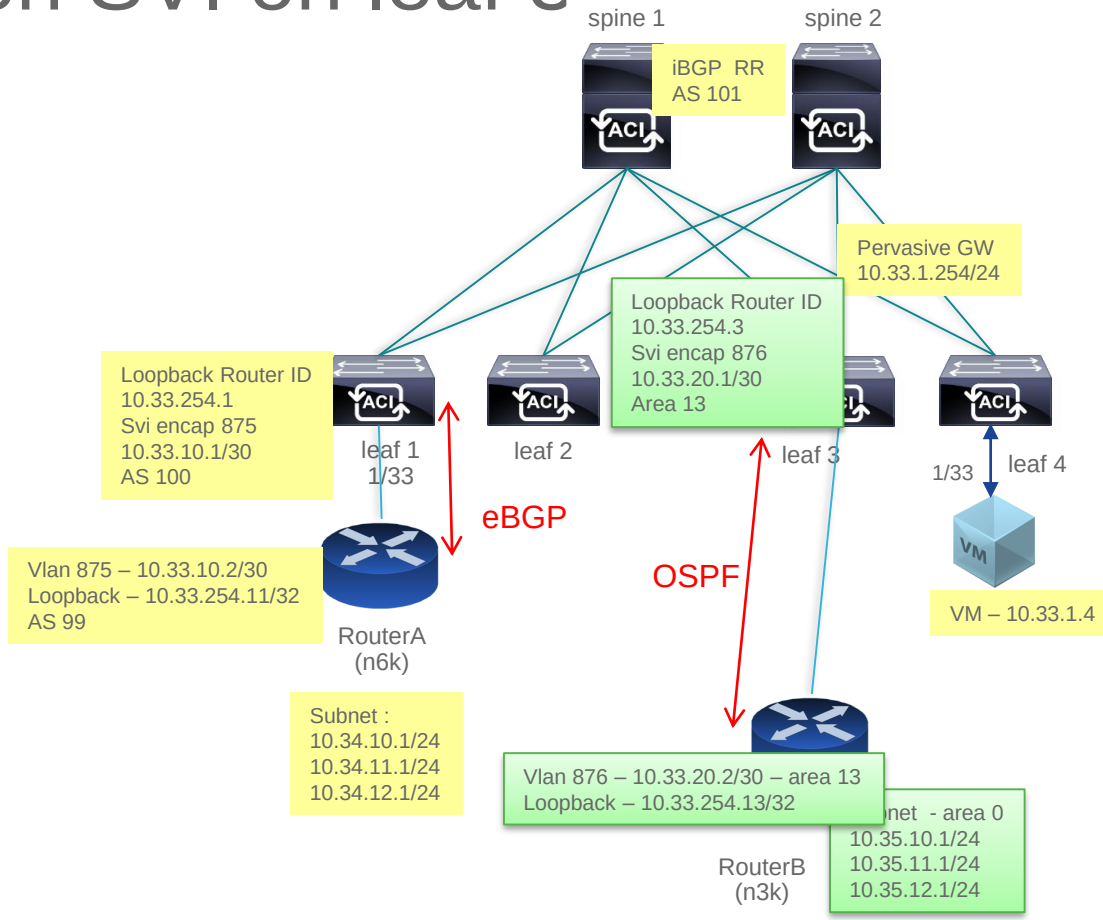
L3 out – Example 2
Adding OSPF L3 out
Configuration and troubleshooting

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Lab setup – L3 out on SVI on leaf 3 using OSPF

Adding a new L3 out for OSPF on leaf 3

Goal is
1/ ping from n3k to VM



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Config– Configuration external router

```
ACI-N3K-1# show run vrf L3:L3

!Command: show running-config vrf L3:L3
!Time: Sat Apr  9 00:54:54 2016

version 6.0(2)U3(3)
vrf context L3:L3
router ospf 103
  vrf L3:L3
    router-id 10.33.254.13

interface loopback13
  vrf member L3:L3
  ip address 10.33.254.12/32

interface Vlan876
  vrf member L3:L3
  ip address 10.33.20.2/30
  ip ospf mtu-ignore
  ip router ospf 103 area 0.0.0.13

interface Vlan2300
  vrf member L3:L3
  ip address 10.35.12.1/24
  ip router ospf 103 area 0.0.0.0

interface Vlan2301
  vrf member L3:L3
  ip address 10.35.10.1/24
  ip router ospf 103 area 0.0.0.0

interface Vlan2302
  vrf member L3:L3
  ip address 10.35.11.1/24
  ip router ospf 103 area 0.0.0.0
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Config– Configuration ACI – L3 out

- Networking
  - Bridge Domains
  - VRFs
  - External Bridged Networks
  - External Routed Networks
    - BGP-Out
      - Logical Node Profiles
      - Networks
      - Route Profiles
    - Match Action Rule Profiles
    - OSPF-Out
      - Logical Node Profiles
        - Node-103
      - Networks
        - epg-l3-ospf
          - L4-L7 Service Parameters
      - Route Profiles
    - Set Action Rule Profiles
  - Route Profiles
  - Protocol Policies
- L4-L7 Service Parameters
- Security Policies
- Troubleshoot Policies
- Monitoring Policies
- L4-L7 Services

Properties

Name: **OSPF-Out**

Description: optional

Tags: enter tags separated by comma

Label:

Target DSCP: unspecified

Route Control Enforcement: ☐ Import ☑ Export

VRF: L3/L3

Resolved VRF: **L3/L3**

External Routed Domain: Phys_L3_out_RD

Route Profile for Interleak: select a value

Route Control For Dampening:

Address Family Type

Enable BGP/EIGRP/OSPF: ☐ BGP ☐ EIGRP

☑ OSPF

OSPF Area ID: 0.0.0.13

OSPF Area Control: ☑ Send redistributed LSAs into NSSA area

☑ Originate summary LSA

☐ Suppress forwarding address in translated LSA

OSPF Area Type: NSSA area | Regular area | Stub area

Similar to example 1
VRF, L3 domain are specified.

Specific to OSPF:
Specify OSPF
Area id and type of Area

Note a Layer 3 can only be part of a single OSPF area. If you want need multiple ospf are in the same vrf in ACI you need multiple L3 out.

ll rights reserved. Cisco Confidential

Config– Configuration ACI – Node and If Profile

- Bridge Domains
- VRFs
- External Bridged Networks
- External Routed Networks
  - BGP-Out
    - Logical Node Profiles
    - Networks
    - Route Profiles
  - Match Action Rule Profiles
  - OSPF-Out
    - Logical Node Profiles
      - Node-103
        - Logical Interface Profiles
          - Ospf-If-Profile

Properties

Name: **Node-103**

Description: optional

Label:

Target DSCP: unspecified

Nodes:

Node ID	Router ID
topology/pod-1/node-103	10.33.254.3

Node profile configured with router ID

Interface profile on 1/33 on node 103
Encap 876 and IP is specified

- Logical Interface Profiles
  - Ospf-If-Profile
    - OSPF Interface Profile
  - Configured Nodes
- Networks
  - epg-l3-ospf
    - L4-L7 Service Parameters
- Route Profiles
- Set Action Rule Profiles

No items have ...
Select Action ... a new item.

SVI:

Path	IP Address	Side A IP	Side B IP	MAC Address	MTU (bytes)	Encap
Node-103/eth1/33	10.33.20.1/30			00:22:BD:F8:19:FF	inherit	vlan-876

Config – L3 out epg

System | Tena

ALL TENANTS | Add Tenant | Search: enter n

Tenant L3

- Quick Start
- Tenant L3
  - Application Profiles
  - Networking
    - Bridge Domains
    - VRFs
    - External Bridged Networks
    - External Routed Networks
      - Set Action Rule Profiles
      - Match Action Rule Profiles
      - BGP-Out
        - Logical Node Profiles
        - Networks
          - epg-l3-bgp
        - Route Profiles
      - OSPF-Out
        - Logical Node Profiles
        - Networks

Subnet - 10.35.10.0/24

Properties

IP Address: **10.35.10.0/24**
address/mask

Scope:
- [ ] Export Route Control Subnet
- [ ] Import Route Control Subnet
- [x] External Subnets for the External EPG
- [ ] Shared Route Control Subnet
- [ ] Shared Security Import Subnet

Aggregate:
- [ ] Aggregate Export
- [ ] Aggregate Import
- [ ] Aggregate Shared Routes

OSPF Route Summarization Policy: select an option

Route Control Profile:

Name	Direction
No items have been found. Select Actions to create a new item.	

- BGP-Out
  - Logical Node Profiles
  - Networks
  - Route Profiles
- Match Action Rule Profiles
- OSPF-Out
  - Logical Node Profiles
    - Node-103
      - Logical Interface Profiles
        - Ospf-If-Profile
          - OSPF Interface Profile
      - Configured Nodes
  - Networks
    - epg-l3-ospf

Creating L3 out Epg
We here use specific subnet received behind this L3 out for OSPF ..
And providing default common contract

Properties

Provided Contracts:

Name	Tenant	Type
default	common	Contract

common

Consumed Contracts:

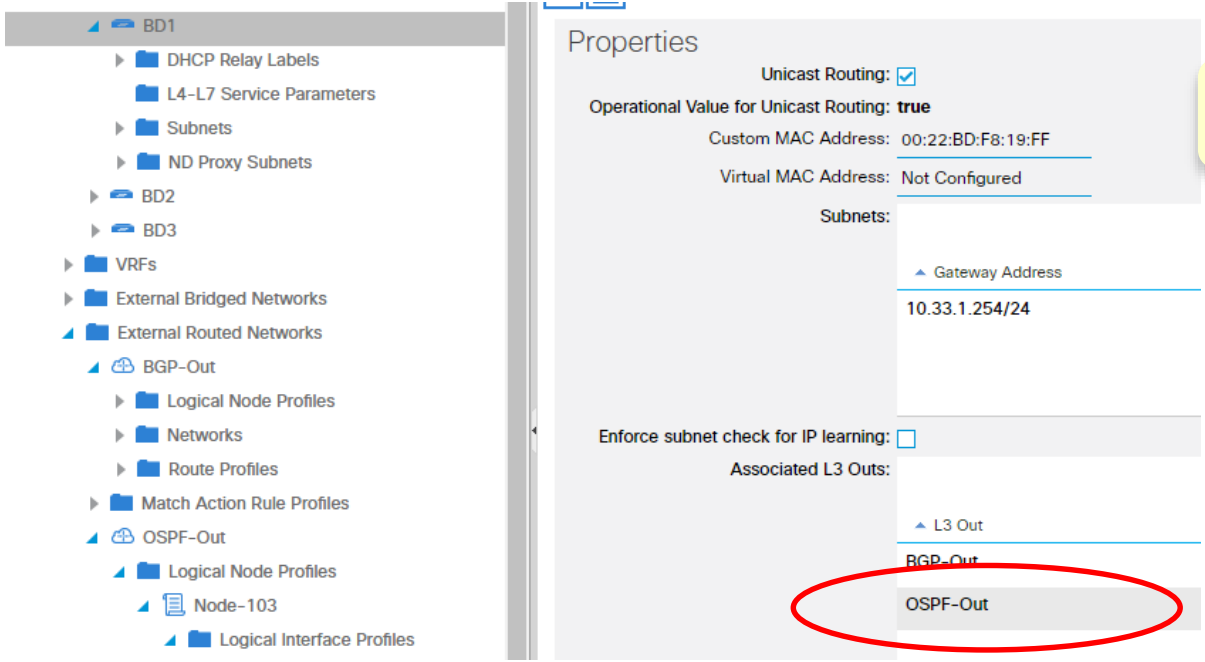
Name	Tenant

Taboo Contracts:

Name	Tenant

cisco

Config – BD settings



Add OSPF-Out to BD associated L3 out

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Verification – ACI OSPF border leaf

```
pod2-leaf3# show ip ospf interface vrf L3:L3
 loopback3 is up, line protocol is up
    IP address 10.33.254.3/32, Process ID default VRF L3:L3, area 0.0.0.13
    Enabled by interface configuration
    State LOOPBACK, Network type LOOPBACK, cost 1

 Vlan46 is up, line protocol is up
    IP address 10.33.20.1/30, Process ID default VRF L3:L3, area 0.0.0.13
    Enabled by interface configuration
    State BDR, Network type BROADCAST, cost 4
    Index 135, Transmit delay 1 sec, Router Priority 1
    Designated Router ID: 10.33.254.13, address: 10.33.20.2
    Backup Designated Router ID: 10.33.254.3, address: 10.33.20.1
    1 Neighbors, flooding to 1, adjacent with 1
    Timer intervals: Hello 10, Dead 40, Wait 40, Retransmit 5
      Hello timer due in 00:00:00
    No authentication
    Number of opaque link LSAs: 0, checksum sum 0

pod2-leaf3# show ip ospf neigh vrf L3:L3
 OSPF Process ID default VRF L3:L3
 Total number of neighbors: 1
 Neighbor ID     Pri State            Up Time  Address         Interface
 10.33.254.13      1 FULL/DR          00:23:41 10.33.20.2      Vlan46
pod2-leaf3#
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Verification – ACI OSPF border leaf

```
pod2-leaf3# show ip ospf database  vrf L3:L3
        OSPF Router with ID (10.33.254.3) (Process ID default VRF L3:L3)

                Router Link States (Area 0.0.0.13)

Link ID         ADV Router      Age         Seq#       Checksum Link Count
10.33.254.3     10.33.254.3     1481        0x80000003 0xb77a   2
10.33.254.13    10.33.254.13    1481        0x80000005 0x8ca9   1

                Network Link States (Area 0.0.0.13)

Link ID         ADV Router      Age         Seq#       Checksum
10.33.20.2      10.33.254.13    1481        0x80000002 0xa3db

                Summary Network Link States (Area 0.0.0.13)

Link ID         ADV Router      Age         Seq#       Checksum
10.35.10.0      10.33.254.13    91          0x80000004 0xb50e
10.35.11.0      10.33.254.13    1793        0x80000003 0xac17
10.35.12.0      10.33.254.13    1803        0x80000003 0xa121

                Type-5 AS External Link States

Link ID         ADV Router      Age         Seq#       Checksum Tag
10.33.1.0       10.33.254.3     599         0x80000002 0x2044    0
10.33.2.0       10.33.254.3     600         0x80000002 0x154e    0
10.33.3.0       10.33.254.3     599         0x80000002 0x0a58    0
```

L3 out subnet and ospf loopback appears as LSA type 1 and 2

Route behind N3k (in area 0) appears as LSA type 3 as expected

ACI pervasive subnet appears as LST type 5 (external LSA) static redistributed to OSPF

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Verification – ACI OSPF border leaf

```
pod2-leaf3# show ip route vrf L3:L3
IP Route Table for VRF "L3:L3"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

10.33.1.0/24, ubest/mbest: 1/0, attached, direct, pervasive
    *via 10.0.232.65%overlay-1, [1/0], 00:10:17, static
10.33.2.0/24, ubest/mbest: 1/0, attached, direct, pervasive
    *via 10.0.232.65%overlay-1, [1/0], 00:10:17, static
10.33.3.0/24, ubest/mbest: 1/0, attached, direct, pervasive
    *via 10.0.232.65%overlay-1, [1/0], 00:10:17, static
10.33.10.0/30, ubest/mbest: 1/0
    *via 10.0.168.95%overlay-1, [200/0], 00:28:11, bgp-101, internal, tag 101
10.33.20.0/30, ubest/mbest: 1/0, attached, direct
    *via 10.33.20.1, vlan46, [1/0], 00:28:12, direct
10.33.20.1/32, ubest/mbest: 1/0, attached
    *via 10.33.20.1, vlan46, [1/0], 00:28:12, local, local
10.33.254.1/32, ubest/mbest: 1/0
    *via 10.0.168.95%overlay-1, [1/0], 00:28:11, bgp-101, internal, tag 101
10.33.254.3/32, ubest/mbest: 2/0, attached, direct
    *via 10.33.254.3, lo3, [1/0], 00:28:12, local, local
    *via 10.33.254.3, lo3, [1/0], 00:28:12, direct
10.34.10.0/24, ubest/mbest: 1/0
    *via 10.0.168.95%overlay-1, [200/0], 00:28:11, bgp-101, internal, tag 100
10.34.11.0/24, ubest/mbest: 1/0
    *via 10.0.168.95%overlay-1, [200/0], 00:28:11, bgp-101, internal, tag 100
10.35.10.0/24, ubest/mbest: 1/0
    *via 10.33.20.2, vlan46, [110/44], 00:24:52, ospf-default, inter
10.35.11.0/24, ubest/mbest: 1/0
    *via 10.33.20.2, vlan46, [110/44], 00:24:52, ospf-default, inter
10.35.12.0/24, ubest/mbest: 1/0
    *via 10.33.20.2, vlan46, [110/44], 00:24:52, ospf-default, inter
```

ACI pervasive subnet appears in OSPF DB as LST type 5 (external LSA) because these ar static redistributed to OSPF

Route behind N3k (in area 0) appears as LSA type 3 as expected

ghts reserved. Cisco Confidential

Verification –OSPF external router

```
ACI-N3K-1# show ip ospf database vrf L3:L3
        OSPF Router with ID (10.33.254.13) (Process ID 103 VRF L3:L3)

..

                Router Link States (Area 0.0.0.13)

Link ID         ADV Router      Age         Seq#       Checksum Link Count
10.33.254.3     10.33.254.3     41          0x80000004 0xb57b   2
10.33.254.13    10.33.254.13    36          0x80000006 0x8aaa   1

                Network Link States (Area 0.0.0.13)

Link ID         ADV Router      Age         Seq#       Checksum
10.33.20.2      10.33.254.13    36          0x80000003 0xa1dc

                Summary Network Link States (Area 0.0.0.13)

Link ID         ADV Router      Age         Seq#       Checksum
10.35.10.0      10.33.254.13    467         0x80000004 0xb50e
10.35.11.0      10.33.254.13    346         0x80000004 0xaa18
10.35.12.0      10.33.254.13    356         0x80000004 0x9f22

                Type-5 AS External Link States

Link ID         ADV Router      Age         Seq#       Checksum Tag
10.33.1.0       10.33.254.3     977         0x80000002 0x2044    0
10.33.2.0       10.33.254.3     977         0x80000002 0x154e    0
10.33.3.0       10.33.254.3     977         0x80000002 0x0a58    0
```

ACI pervasive subnet appears as LST type 5 (external LSA) redistributed to OSPF

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Verification – external router routing table

```
ACI-N3K-1# show ip route vrf L3:L3
IP Route Table for VRF "L3:L3"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

10.33.1.0/24, ubest/mbest: 1/0
    *via 10.33.20.1, Vlan876, [110/20], 00:18:09, ospf-103, type-2
10.33.2.0/24, ubest/mbest: 1/0
    *via 10.33.20.1, Vlan876, [110/20], 00:18:09, ospf-103, type-2
10.33.3.0/24, ubest/mbest: 1/0
    *via 10.33.20.1, Vlan876, [110/20], 00:18:09, ospf-103, type-2
10.33.20.0/30, ubest/mbest: 1/0, attached
    *via 10.33.20.2, Vlan876, [0/0], 01:16:27, direct
10.33.20.2/32, ubest/mbest: 1/0, attached
    *via 10.33.20.2, Vlan876, [0/0], 01:16:27, local
10.33.254.3/32, ubest/mbest: 1/0
    *via 10.33.20.1, Vlan876, [110/41], 00:32:45, ospf-103, intra
10.33.254.12/32, ubest/mbest: 2/0, attached
    *via 10.33.254.12, Lo13, [0/0], 01:18:03, local
    *via 10.33.254.12, Lo13, [0/0], 01:18:03, direct
10.35.10.0/24, ubest/mbest: 1/0, attached
    *via 10.35.10.1, Vlan2301, [0/0], 01:15:11, direct
10.35.10.1/32, ubest/mbest: 1/0, attached
    *via 10.35.10.1, Vlan2301, [0/0], 01:15:11, local
10.35.11.0/24, ubest/mbest: 1/0, attached
    *via 10.35.11.1, Vlan2302, [0/0], 01:08:24, direct
10.35.11.1/32, ubest/mbest: 1/0, attached
    *via 10.35.11.1, Vlan2302, [0/0], 01:08:24, local
10.35.12.0/24, ubest/mbest: 1/0, attached
    *via 10.35.12.1, Vlan2300, [0/0], 01:08:28, direct
10.35.12.1/32, ubest/mbest: 1/0, attached
    *via 10.35.12.1, Vlan2300, [0/0], 01:08:28, local
```

ACI pervasive subnet appears as type 2 ospf routes

Verification – non Border Leaf routing table

```
pod2-leaf4# show ip route vrf L3:L3
IP Route Table for VRF "L3:L3"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

10.33.1.0/24, ubest/mbest: 1/0, attached, direct, pervasive
    *via 10.0.232.65%overlay-1, [1/0], 1d03h, static
10.33.1.254/32, ubest/mbest: 1/0, attached, pervasive
    *via 10.33.1.254, vlan47, [1/0], 1d03h, local, local
10.33.2.0/24, ubest/mbest: 1/0, attached, direct, pervasive
    *via 10.0.232.65%overlay-1, [1/0], 1d03h, static
10.33.2.254/32, ubest/mbest: 1/0, attached, pervasive
    *via 10.33.2.254, vlan49, [1/0], 1d03h, local, local
10.33.3.0/24, ubest/mbest: 1/0, attached, direct, pervasive
    *via 10.0.232.65%overlay-1, [1/0], 1d03h, static
10.33.3.254/32, ubest/mbest: 1/0, attached, pervasive
    *via 10.33.3.254, vlan9, [1/0], 1d03h, local, local
10.33.10.0/30, ubest/mbest: 1/0
    *via 10.0.168.95%overlay-1, [200/0], 1d03h, bgp-101, internal, tag 101
10.33.20.0/30, ubest/mbest: 1/0
    *via 10.0.168.91%overlay-1, [200/0], 00:37:26, bgp-101, internal, tag 101
10.33.254.1/32, ubest/mbest: 1/0
    *via 10.0.168.95%overlay-1, [1/0], 1d03h, bgp-101, internal, tag 101
10.33.254.3/32, ubest/mbest: 1/0
    *via 10.0.168.91%overlay-1, [1/0], 00:37:26, bgp-101, internal, tag 101
10.34.10.0/24, ubest/mbest: 1/0
    *via 10.0.168.95%overlay-1, [200/0], 1d02h, bgp-101, internal, tag 100
10.34.11.0/24, ubest/mbest: 1/0
    *via 10.0.168.95%overlay-1, [200/0], 1d02h, bgp-101, internal, tag 100
10.35.10.0/24, ubest/mbest: 1/0
    *via 10.0.168.91%overlay-1, [200/44], 00:34:06, bgp-101, internal, tag 101
10.35.11.0/24, ubest/mbest: 1/0
    *via 10.0.168.91%overlay-1, [200/44], 00:34:06, bgp-101, internal, tag 101
10.35.12.0/24, ubest/mbest: 1/0
    *via 10.0.168.91%overlay-1, [200/44], 00:34:06, bgp-101, internal, tag 101
```

ACI pervasive subnet

Ebgp subnet coming from border leaf 1 (example 1)

OSPF subnet redistributed to iBGP vpnv4 from leaf3

L3 out – Example 3
Transit Routing
Configuration and troubleshooting

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Lab setup – Transit Routing

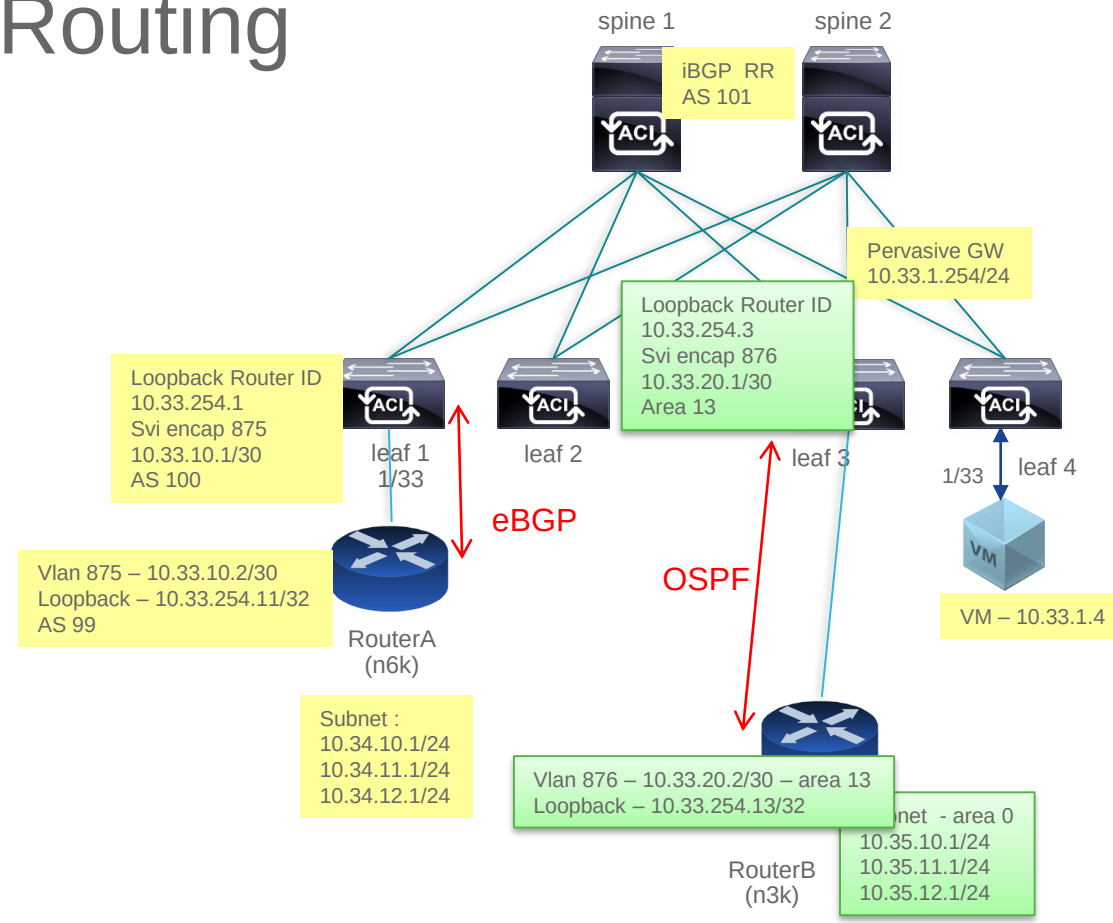
Same setup as example 1 + 2

Now we want to be able to ping from RouteA (n6k) to router B (n3k)

Original situation :

Subnet 10.35.x.x are learned from OSPF and are known in ACI but not send out to Router A eBGP

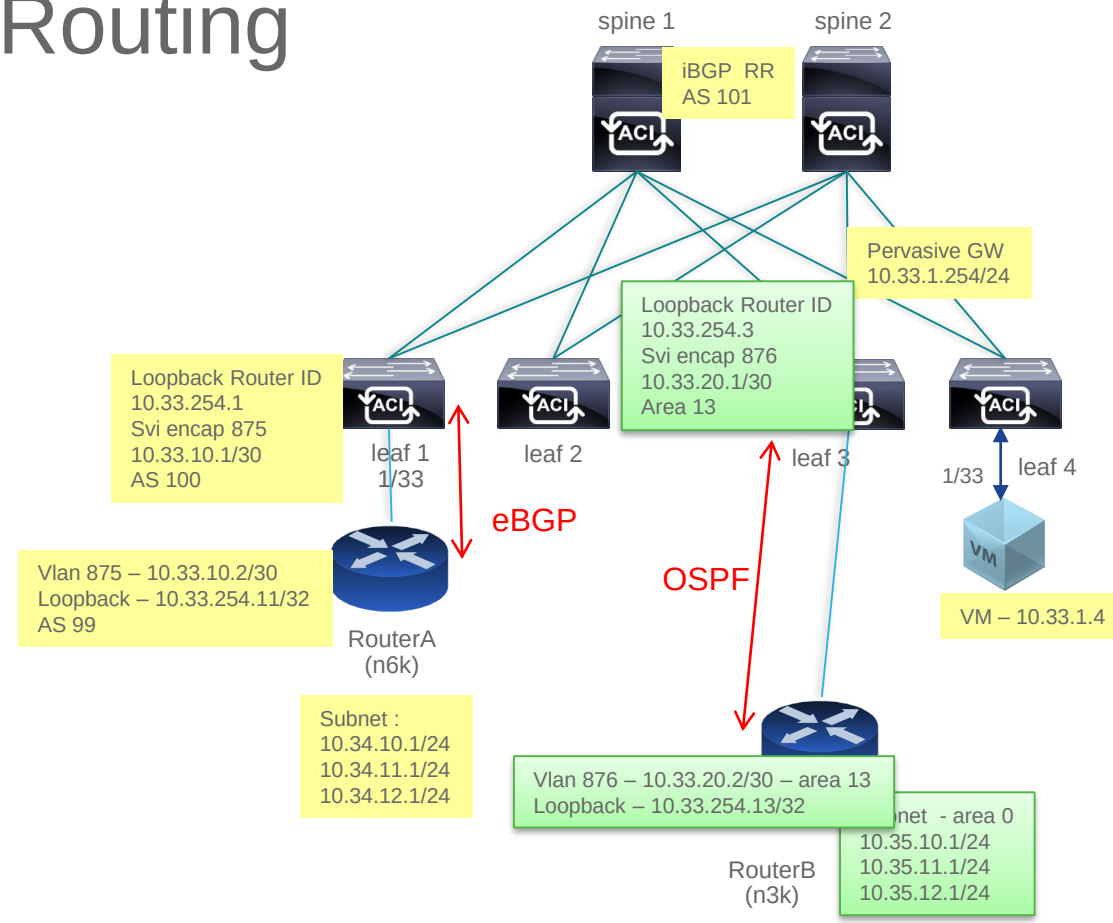
Subnet 10.34.x.x are learned from eBGP and are known in ACI but not send out to Router B via OSPF



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

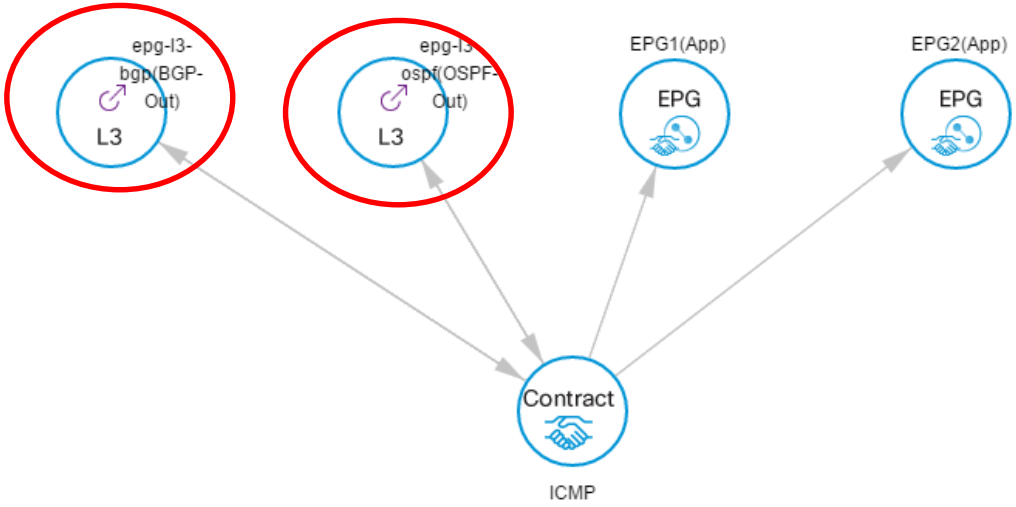
Lab setup – Transit Routing

3 steps:
1. Ensure there is contract between the two L3 out
2. Redistribute 10.35.x.x subnet on leaf 1 between VPNv4 iBGP to eBGP
3. Redistribute 10.34.x.x subnet on leaf 3 between VPNv4 iBGP to OSPF



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 1 – Add a contract between the two L3 out



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 2 – redistribute to eBGP on leaf 1 - Before

```
pod2-leaf1# show ip bgp vrf L3:L3 neighbors 10.33.10.2 advertised-routes

Peer 10.33.10.2 routes for address family IPv4 Unicast:
BGP table version is 25, local router ID is 10.33.254.1
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

   Network            Next Hop            Metric     LocPrf     Weight Path
*>r10.33.1.0/24       0.0.0.0                  0        100      32768 ?
*>r10.33.2.0/24       0.0.0.0                  0        100      32768 ?
*>r10.33.3.0/24       0.0.0.0                  0        100      32768 ?
```

Only local ACI pervasive subnet are send out

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 2 – redistribute to eBGP on leaf 1 - config

ALL TENANTS | Add Tenant | Search: enter name, descr | L3 | common | DC | DC-Test | infra

Tenant L3

- Quick Start
- Tenant L3
  - Application Profiles
  - Networking
    - Bridge Domains
    - VRFs
    - External Bridged Networks
    - External Routed Networks
      - Set Action Rule Profiles
      - Match Action Rule Profiles
      - BGP-Out
        - Logical Node Profiles
        - Networks
          - epg-l3-bgp
        - Route Profiles
      - OSPF-Out
        - Logical Node Profiles
        - Networks
          - epg-l3-ospf
        - Route Profiles
    - Route Profiles

Subnet - 10.35.10.1/24

Properties

IP Address: **10.35.10.1/24**

Scope:
- [x] Export Route Control Subnet
- [ ] Import Route Control Subnet
- [ ] External Subnets for the External EPG
- [ ] Shared Route Control Subnet
- [ ] Shared Security Import Subnet

Aggregate:
- [ ] Aggregate Export
- [ ] Aggregate Import
- [ ] Aggregate Shared Routes

BGP Route Summarization Policy: select an option

Route Control Profile:

We need to add under the BGP l3 out EPG , the subnet received from OSPF on the 2nd L3 out and mark them as

Export Router Control Subnet

epg-l3-bgp
Route Profiles
OSPF-Out
Logical Node Profiles
Networks
epg-l3-ospf
Route Profiles
Route Profiles
Protocol Policies
L4-L7 Service Parameters
Security Policies
Troubleshoot Policies
Monitoring Policies

Resolved VRF: **uni/tn-L3/ctx-L3**

QoS Class: Unspecified

Target DSCP: unspecified

Configuration Status: **applied**

Configuration Issues:

Subnets:

IP Address	Scope	Aggregate
10.35.10.1/24	Export Route Control Subnet	
10.35.11.0/24	Export Route Control Subnet	

Step 2 – redistribute to eBGP on leaf 1 – Verify route-map

```
pod2-leaf1# show ip bgp neighbor vrf L3:L3 | grep route-map
  Inbound route-map configured is imp-l3out-BGP-Out-peer-2785280, handle obtained
  Outbound route-map configured is exp-l3out-BGP-Out-peer-2785280, handle obtained


pod2-leaf1# show route-map  exp-l3out-BGP-Out-peer-2785280
route-map exp-l3out-BGP-Out-peer-2785280, permit, sequence 7801
  Match clauses:
    ip address prefix-lists: IPv4-peer32775-2785280-exc-int-inferred-export-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:
route-map exp-l3out-BGP-Out-peer-2785280, permit, sequence 7802
  Match clauses:
    ip address prefix-lists: IPv4-peer32775-2785280-exc-ext-inferred-export-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:
    tag 4294967295
route-map exp-l3out-BGP-Out-peer-2785280, deny, sequence 8000
  Match clauses:
    route-type: static
  Set clauses:
route-map exp-l3out-BGP-Out-peer-2785280, deny, sequence 8001
  Match clauses:
    route-type: direct
  Set clauses:


pod2-leaf1# show ip prefix-list IPv4-peer32775-2785280-exc-ext-inferred-export-dst
ip prefix-list IPv4-peer32775-2785280-exc-ext-inferred-export-dst: 3 entries
   seq 3 permit 10.35.10.0/24
   seq 4 permit 10.35.12.0/24
   seq 7 permit 10.35.11.0/24
```

Redistribution is done via route-map out on bgp neighbors.

Subnet we flag as Advertize out are showing up in the prefix list.

Step 2 – redistribute to eBGP on leaf 1 – Verify routing

```
pod2-leaf1# pod2-leaf1# show ip bgp vrf L3:L3 neighbors 10.33.10.2 advertised-routes

Peer 10.33.10.2 routes for address family IPv4 Unicast:
BGP table version is 46, local router ID is 10.33.254.1
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

   Network            Next Hop            Metric     LocPrf     Weight Path
*>r10.33.1.0/24       0.0.0.0                  0        100      32768 ?
*>r10.33.2.0/24       0.0.0.0                  0        100      32768 ?
*>r10.33.3.0/24       0.0.0.0                  0        100      32768 ?
*>i10.35.10.0/24      10.0.168.91             44        100          0 ?
*>i10.35.11.0/24      10.0.168.91             44        100          0 ?
*>i10.35.12.0/24      10.0.168.91             44        100          0 ?
```

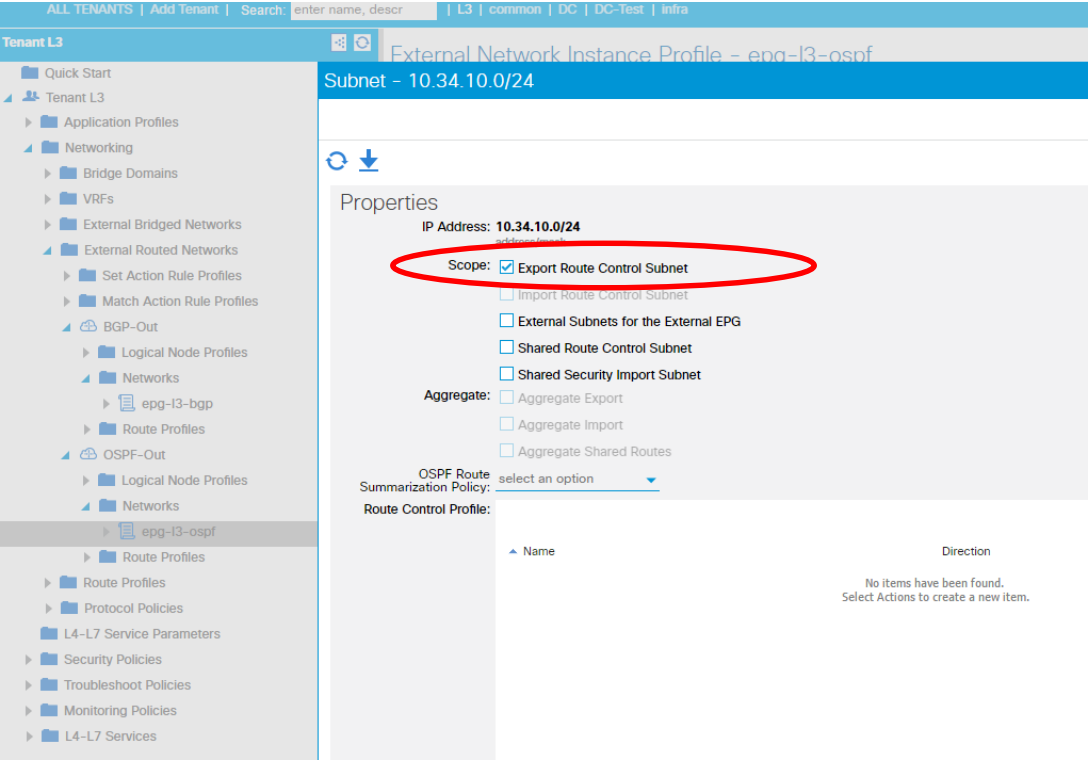
Subnet are advertize via BGP to our Neighbor.

```
bdsol-n6004-01# show ip route vrf L3:L3 | egrep -A 1 "10.35"
10.35.10.0/24, ubest/mbest: 1/0
    *via 10.33.10.1, [20/0], 00:02:55, bgp-99, external, tag 100
10.35.11.0/24, ubest/mbest: 1/0
    *via 10.33.10.1, [20/0], 00:02:43, bgp-99, external, tag 100
10.35.12.0/24, ubest/mbest: 1/0
    *via 10.33.10.1, [20/0], 00:02:31, bgp-99, external, tag 100
```

Nexus 6k receives them and install in routing-table

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 3 – Redistribute subnet to OSPF on leaf 3



Similar as step 2, we add the prefix received by BGP on leaf1 as Export Route control subnet under the epg for OSPF on leaf 3.

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 3 – redistribute to OSPF on leaf 3 – Verify route-map

```
.
pod2-leaf3# show ip ospf vrf L3:L3 | egrep "Area|filter"
    Area (0.0.0.13)
        Area has existed for 2d22h
        Area ranges are
        Area-filter in 'exp-ctx-proto-2785280'

pod2-leaf3# show route-map exp-ctx-proto-2785280
route-map exp-ctx-proto-2785280, permit, sequence 7801
  Match clauses:
    ip address prefix-lists: IPv4-proto49155-2785280-exc-ext-inferred-export-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:
    tag 4294967295

pod2-leaf3# show ip prefix-list IPv4-proto49155-2785280-exc-ext-inferred-export-dst
ip prefix-list IPv4-proto49155-2785280-exc-ext-inferred-export-dst: 2 entries
   seq 3 permit 10.34.10.0/24
   seq 4 permit 10.34.11.0/24
```

Redistribution is done via route-map out per Area (what is allowed inside the area)

Subnet we flag as Advertize out are showing up in the prefix list.

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 3 – redistribute to OPSF on leaf 3 – Verify routing

```
pod2-leaf3# show ip ospf  database external vrf L3:L3
        OSPF Router with ID (10.33.254.3) (Process ID default VRF L3:L3)

                Type-5 AS External Link States

Link ID         ADV Router      Age         Seq#       Checksum Tag
10.33.1.0       10.33.254.3     719         0x8000008d 0x09cf    0
10.33.2.0       10.33.254.3     719         0x8000008d 0xfdd9    0
10.33.3.0       10.33.254.3     719         0x8000008d 0xf2e3    0
10.34.10.0      10.33.254.3     524         0x80000002 0xf17b    4294967295
10.34.11.0      10.33.254.3     49          0x80000002 0xe685    4294967295
```

Prefixes appeared in OSPF Database LSA type 5 (redistributed)

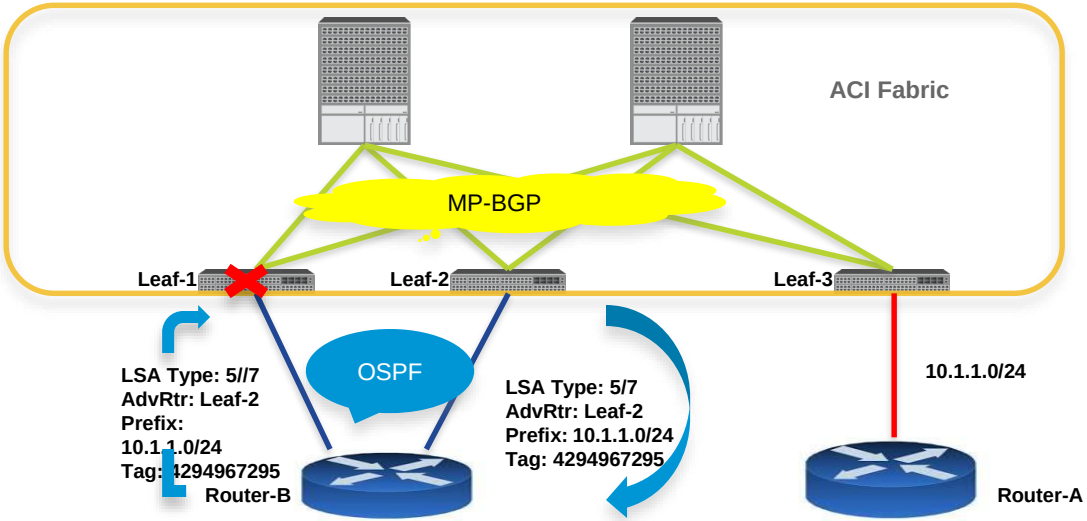
```
ACI-N3K-1# show ip route vrf L3:L3 | egrep -A 1 "10.34"
10.34.10.0/24, ubest/mbest: 1/0
    *via 10.33.20.1, Vlan876, [110/1], 00:11:30, ospf-103, type-2, tag 4294967295
10.34.11.0/24, ubest/mbest: 1/0
    *via 10.33.20.1, Vlan876, [110/1], 00:03:36, ospf-103, type-2, tag 4294967295
ACI-N3K-1#


ACI-N3K-1# ping 10.34.10.1 source 10.35.10.1 vrf L3:L3
PING 10.34.10.1 (10.34.10.1) from 10.35.10.1: 56 data bytes
64 bytes from 10.34.10.1: icmp_seq=0 ttl=252 time=1.669 ms
64 bytes from 10.34.10.1: icmp_seq=1 ttl=252 time=1.193 ms
64 bytes from 10.34.10.1: icmp_seq=2 ttl=252 time=1.154 ms
64 bytes from 10.34.10.1: icmp_seq=3 ttl=252 time=2.312 ms
64 bytes from 10.34.10.1: icmp_seq=4 ttl=252 time=1.185 ms
```

Nexus 3k, gets them as OSPF external route

Ping is flowing end to end !!!

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Routing Loop Avoidance



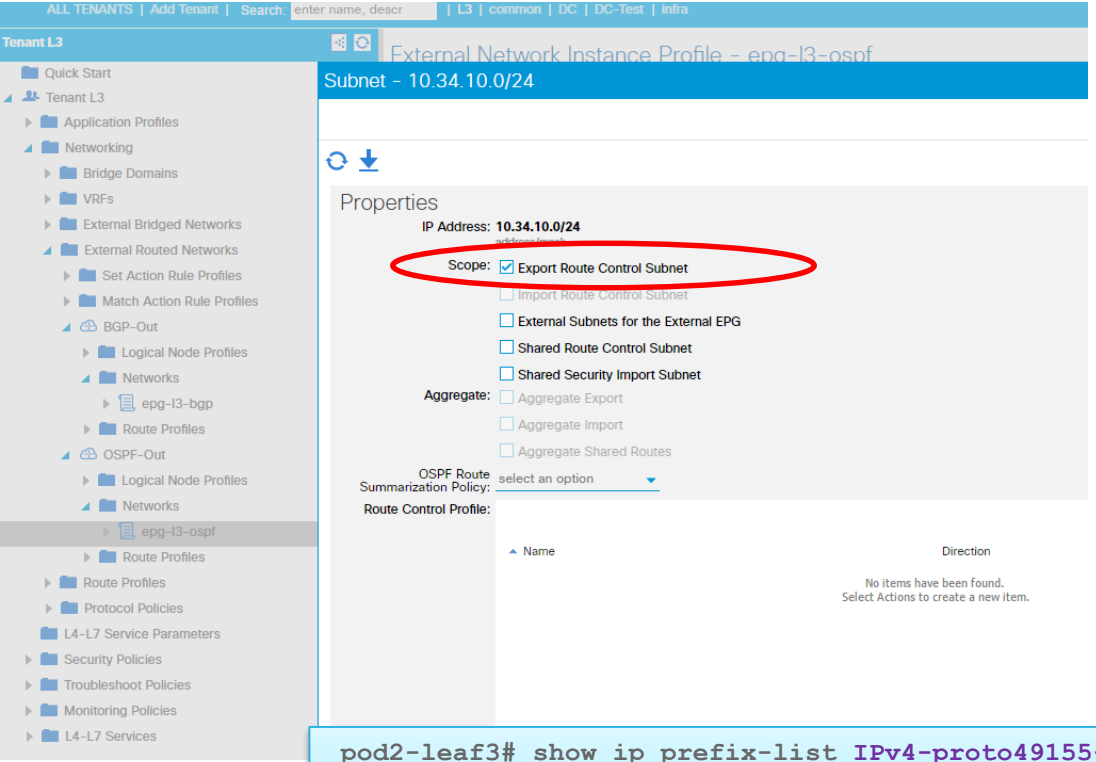
- **Transit prefixes are learnt by OSPF via redistribution from MP-BGP**
- **Do not want transit prefixes advertised by the fabric to be re-learnt by the fabric via external router --- Avoid potential routing loops**
- **Transit prefixes within a VRF are marked with a configurable tag (default: 2^32 – 1)**

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Transit routing – aggregate route

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Aggregate route ?



When you mark a subnet as Export Route Control Subnet , see the Aggregate Export is greyed out

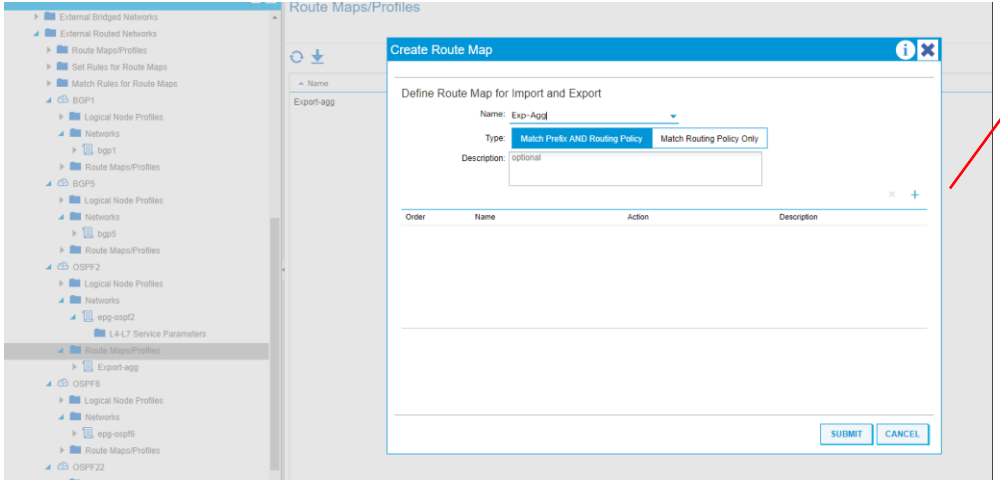
The prefix list is an exact math we do not have the le 32 to match everythinbg between the prefix (/24 and 32)

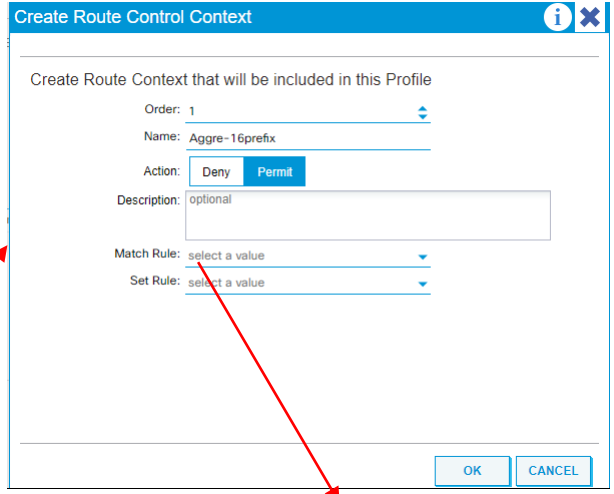
```
pod2-leaf3# show ip prefix-list IPv4-proto49155-2785280-exc-ext-inferred-export-dst
ip prefix-list IPv4-proto49155-2785280-exc-ext-inferred-export-dst: 2 entries
   seq 3 permit 10.34.10.0/24
```

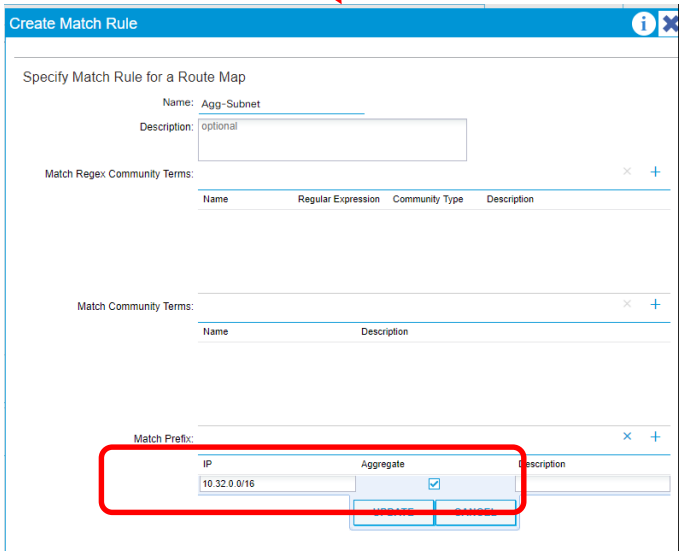
Step 1-

- Configure a route map in the layer 3 to mark the supernet we want as aggregate
- Create new route map
- Create a new sequence in it
- Create the match rule for the supernet
- No need to create set rules

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential


Route Maps/Profiles
Create Route Map
Define Route Map for Import and Export
Name: Exp-Agg
Type: Match Prefix AND Routing Policy
Match Routing Policy Only
Description: optional
Order
Name
Action
Description
SUBMIT
CANCEL
Export-agg

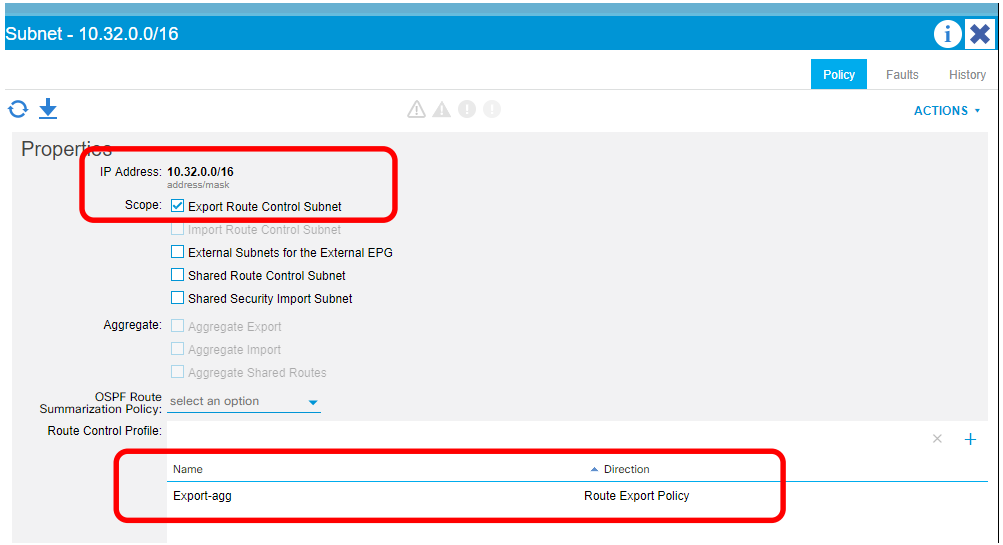

Create Route Control Context
Create Route Context that will be included in this Profile
Order: 1
Name: Aggre-16prefix
Action: Deny Permit
Description: optional
Match Rule: select a value
Set Rule: select a value
OK
CANCEL


Create Match Rule
Specify Match Rule for a Route Map
Name: Agg-Subnet
Description: optional
Match Regex Community Terms:
Name
Regular Expression
Community Type
Description
Match Community Terms:
Name
Description
Match Prefix:
IP
Aggregate
Description
10.32.0.0/16
UPDATE
CANCEL

Step 2

- In l3 out EPG subnet,
- Create a subnet for the supernet to aggregate,
- Mark it as ‘export route control subnet’
- Add route control profile for export and select the created route map in previous step

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Result on the leaf

```
bdsol-aci32-leaf2# show ip ospf vrf RD-BGP:RD | egrep filter
        Area-filter in 'exp-ctx-proto-2654211'

bdsol-aci32-leaf2# show route-map exp-ctx-proto-2654211
route-map exp-ctx-proto-2654211, permit, sequence 4601
route-map exp-ctx-proto-2654211, permit, sequence 4602
  Match clauses:
    ip address prefix-lists: IPv4-proto16388-2654211-exc-ext-out-Export-agg1exp-Agg-32132Subnet-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:
    tag 4294967295
bdsol-aci32-leaf2# show ip prefix-list  IPv4-proto16388-2654211-exc-ext-out-Export-agg1exp-Agg-
32132Subnet-dst
ip prefix-list IPv4-proto16388-2654211-exc-ext-out-Export-agg1exp-Agg-32132Subnet-dst: 1 entries
   seq 1 permit 10.32.0.0/16 le 32
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

OSPD DB now have subnet of supernet as LSA 5

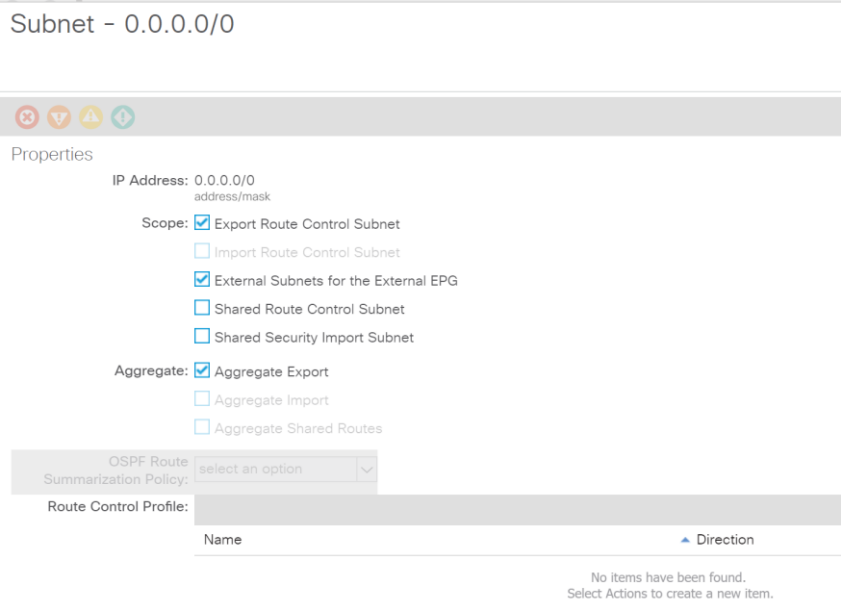
```
bdsol-aci32-leaf2# show ip ospf database external vrf RD-BGP:RD
        OSPF Router with ID (172.16.1.2) (Process ID default VRF RD-BGP:RD)

                Type-5 AS External Link States

Link ID         ADV Router      Age         Seq#       Checksum Tag
10.32.1.0       172.16.1.2      142         0x80000002 0x36af    4294967295
10.32.2.0       172.16.1.2      142         0x80000002 0x2bb9    4294967295
10.32.7.0       172.16.1.2      142         0x80000002 0xf3eb    4294967295
10.32.9.0       172.16.1.2      142         0x80000002 0xddff    4294967295
```

Aggregate supernet was 10.32.0.0/16
We now have all its underlying /24 subnet in OSPF and they are sent to External router

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Transit Aggregate route exception

- The only exception is the route 0.0.0.0/0 which can be set as export route control subnet and export aggregate directly
- This will create prefix list with 0.0.0.0/0 le 32
- Note this can only be done for the 0.0.0.0/0
- It is not recommended as it can easily lead to temporarely routing loop in case of reconvergence if L3 out is shared on multiple router
- (see routing loop example later in this pres)



L3 out
ARP and EP learning

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

ARP on L3 out

- ARP for next-hop must be resolved properly.
- Typically ACI do not store ARP info in show ip arp. However, l2 attached IP on layer 3 (aka next-hops) are an exception and should be seen in show ip arp

```
pod2-leaf1# show ip arp vrf L3:L3

Flags: * - Adjacencies learnt on non-active FHRP router
       + - Adjacencies synced via CFSoE
       # - Adjacencies Throttled for Glean
       D - Static Adjacencies attached to down interface

IP ARP Table for all contexts
Total number of entries: 1
Address         Age       MAC Address     Interface
10.33.10.2      00:08:58  002a.6ab1.917c  vlan65
```

tes. All rights reserved. Cisco Confidential

ARP on L3 out - Troubleshooting

- Step 1 – TCP dump on int kpm_inb on leaf to see if we get ARP (note we only see Rx ARP on this interface, not Tx)

```
pod2-leaf1# tcpdump -i kpm_inb arp
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on kpm_inb, link-type EN10MB (Ethernet), capture size 65535 bytes
23:24:46.926593 ARP, Reply 10.33.10.2 is-at 00:2a:6a:b1:91:7c (oui Unknown), length 42
^C
```

For ARP specifically, tcpdump on kpm_inb
Do only show Rx packet

- Step 2 – Check ARP process debug log

We see (read bottom to top)
- We received arp from 10.33.10.2
- We add the arp entry
- We populate MIT object store with the arp info

```
Yes

pod2-leaf1# show ip arp internal  event-history event  | egrep -B 1 "10.33.10.2" | more
156) Event:E_DEBUG_DSF, length:420, at 927332 usecs after Mon Apr 11 23:24:46 2016
    [116] TID 9422:arp_objst_dynadj_handle_pri_chunk:1279: Object store handling: my_dn - sys/arp/inst/dom-L3:L3/db-ip/adj-[vlan65]-
[10.33.10.2], oper state - 2, arp_cfg_opc - 1, arp_mac - 002a.6ab1.917c, arp_ip
- 10.33.10.2, arp_ifindex - 0, arp_iod - 17, arp_iod - Vlan65, arp_phy_ifindex - 0,arp_phy_iod - 88, arp_phy_iod - Ethernet1/33,
context_id - 14, uptime(sec) - 6737, uptime(nsec) - 702545344, upTS - 1460417087363
--
157) Event:E_DEBUG_DSF, length:212, at 927263 usecs after Mon Apr 11 23:24:46 2016
    [116] TID 9422:arp_add_adj:3898: Entry added for 10.33.10.2, 002a.6ab1.917c, state 2 on interface Vlan65, physical interface
Ethernet1/33, ismct 0. Rearp (interval: 0, count: 0), TTL: 1500 seconds, sent_to_am
: No
--
160) Event:E_DEBUG_DSF, length:134, at 927071 usecs after Mon Apr 11 23:24:46 2016
    [116] TID 9422:arp_process_receive_packet_msg:6681: Check if garp. iod Vlan65 ifmode 128 src 10.33.10.2 dst 10.33.10.1 adj
0x97c0f594
--
```

EPM – EP learning on L3 out

- We do not learn EP received on Layer 3 out
- One Exception only:
  - In case L3 out is using SVI
  - We will learn only Mac address of next-hop in the SVI with no IP address
  - The BD_EXT_VLAN (ext svi BD) is set as Flood and we do not learn IP
  - After CSCvd56870 : timers for those BD (EP retention policy were change to be higher than regular BD and higer than ARP timer.

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

ELTMC Ext SVI vlan info

```
module-1# show system internal eltmc  info  vlan 13
             vlan_id:              13   :::      hw_vlan_id:             26
           vlan_type:     BD_EXT_VLAN   :::         bd_vlan:             13
   access_encap_type:          802.1q   :::    access_encap:           1120
   fabric_encap_type:            VXLAN   :::    fabric_encap:       15040475
              sclass:           49153   :::           scope:              3
            untagged:               0
     acess_encap_hex:           0x460   :::  fabric_enc_hex:       0xe57fdb
        vlan_fd_list:
             context:      RD-OSPF:RD
     pd_vlan_ft_mask:             0x8
       learn_disable:               0
        qos_class_id:               0   :::      qos_pap_id:              0
          qq_met_ptr:              26   :::      ipmc_index:              0
   ingressBdAclLabel:           0x800   ::: ingBdAclLblMask:          0x800
    egressBdAclLabel:               0   ::: egrBdAclLblMask:              0
        rwBdAclLabel:               0   :::  rwBdAclLblMask:              0
         qos_map_idx:               0   :::     qos_map_pri:              0
        qos_map_dscp:               0   :::      qos_map_tc:              0
        vlan_ft_mask:         0x209ff   :::       sisf_bmap:              0
            fwd_mode:  bridge, route
            arp_mode:           flood
            ip_learn:               1
           mac_learn:               1
        unk_mc_flood:               1
     copy_service_bd:               0
    qinq provider bd:               0   :::  pgmQiqEpgLabel:              0
         unk_uc_mode:           flood
          multi_dest:           flood
 ep_move_detect_mode:       undefined
           hw_bd_idx:            4630   :::      hw_epg_idx:          11263
          intf_count:               1   ::: glbl_scp_if_cnt:              1
```

2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

EPMC ext SVI vlan info

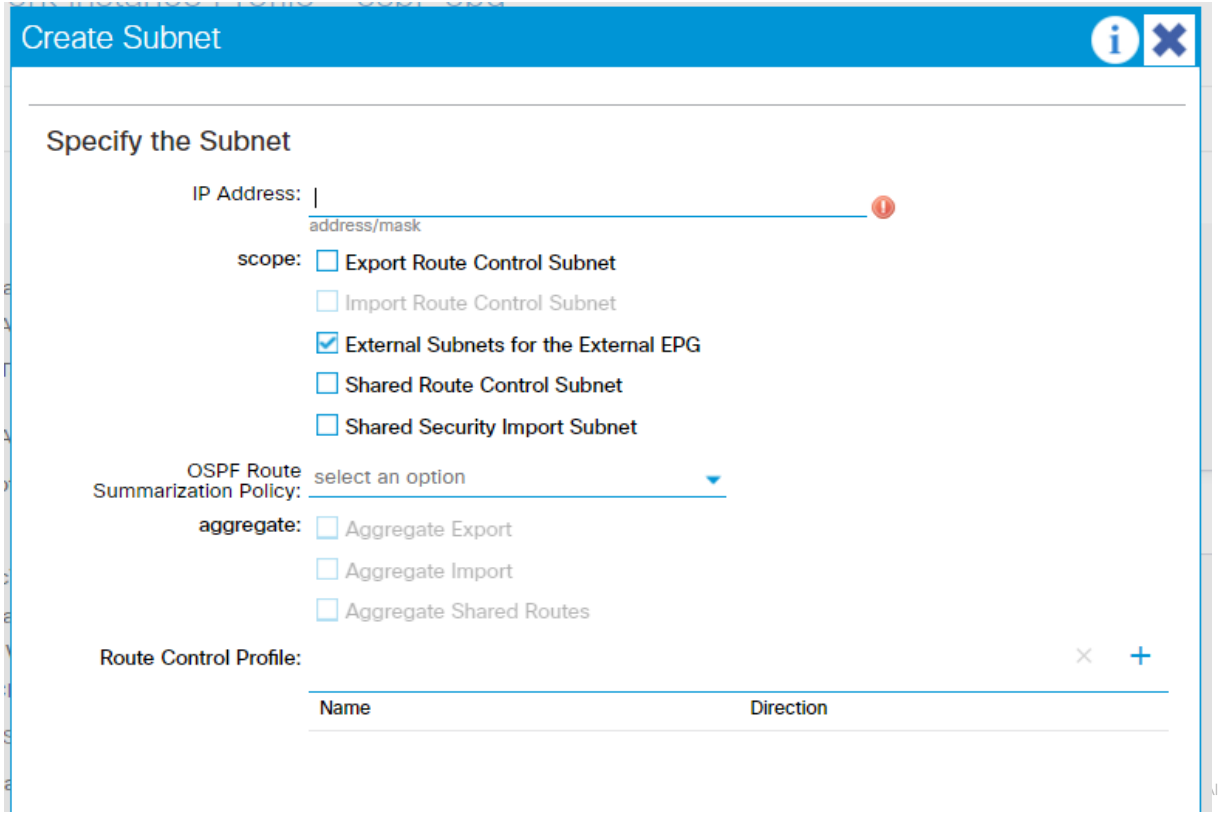
```
module-1# show system internal epmc vlan 13 detail


VLAN 13
VLAN type : Ext. BD
hw id : 26 ::: scope : 3  ::: sclass : 49153
access enc : (802.1Q, 1120)
fabric enc : (VXLAN, 15040475)
BD vlan id : 13 ::: BD vnid : 15040475 ::: VRF vnid : 2621441
EP retention policy valid : Yes
Local EP timeout : 2500 ::: Remote EP timeout : 1200
EP bounce timeout : 2500 ::: EP hold timeout : 300 ::: EP move frequency : 256
fwd_mode : route,bridge ::: fwd_ctrl : mdst-flood,arp-flood, ::: bridge_mode: mac ::: unk_mac_ucast: flood
is_sisf_enabled : No ::: fhs_mode :
Endpoint count : 3
Learning disabled :no
No of ep moves 1 since 03:08:18.520520
BD Subnet ip_pfx-1 : 10.0.9.3/24
MAC CKT HT is empty
::::
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

L3 out – Subnet Flags detail

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

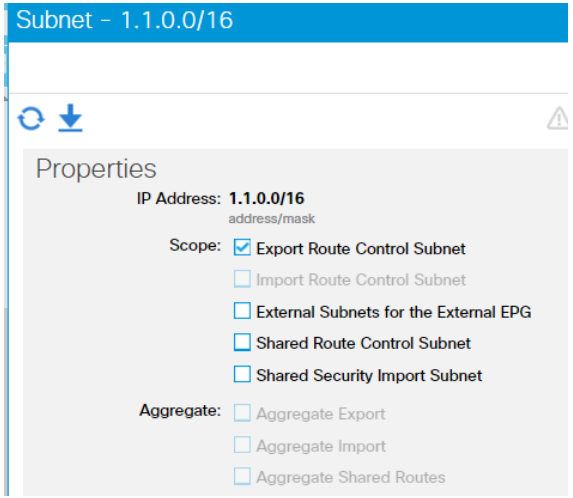
When to use which flags ?



All rights reserved. Cisco Confidential

1. External Route Control Subnet

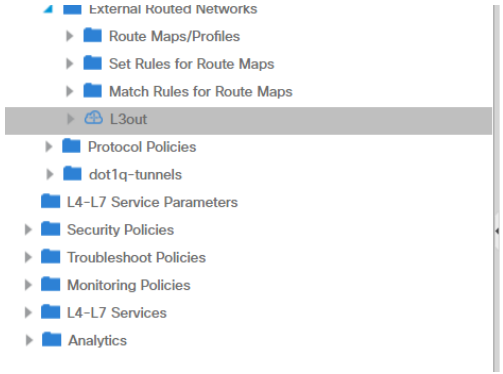
- Only use in transit routing, to advertise prefix received from a first L3 out (or GOLF) down to a second layer 3 out
- Directly adds entry in prefix-list of the outbound route-map
- Should NOT contain the BD prefix or the prefix you are suppose to receive on that Layer 3 out
- Does not matter whether the prefix exist or is received.
- Do not add routes, only plays with route-map
- Route-map is on BGP neighbor or ospf area-filter

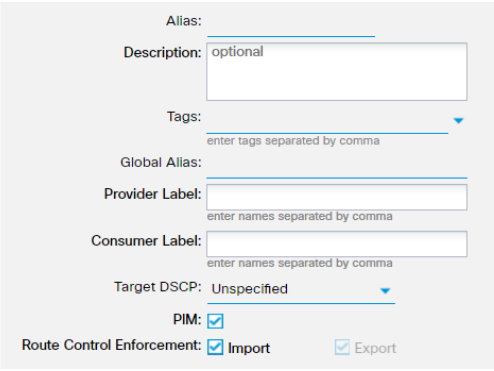


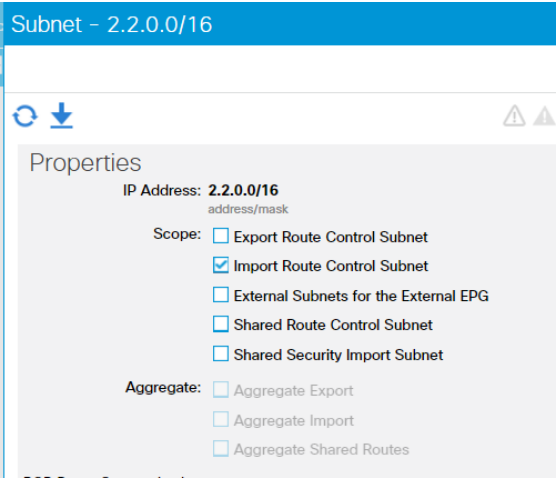

```
bdsol-aci32-leaf2# show ip ospf  vrf DC:DC  | egrep lter
        Area-filter in 'exp-ctx-proto-3014656'
bdsol-aci32-leaf2# show route-map exp-ctx-proto-3014656
route-map exp-ctx-proto-3014656, permit, sequence 7801
  Match clauses:
    ip address prefix-lists: IPv4-proto32772-3014656-exc-ext-inferred-export-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:
    tag 4294967295
bdsol-aci32-leaf2# show ip prefix-list IPv4-proto32772-3014656-exc-ext-inferred-export-dst
ip prefix-list IPv4-proto32772-3014656-exc-ext-inferred-export-dst: 1 entries
   seq 1 permit 1.1.0.0/16
```

2. Import Route Control Subnet

- Disabled by default (greyed out)
- Only enable if the L3 out is set with import security enable
- Works the same as export route control subnet but on the import route-map



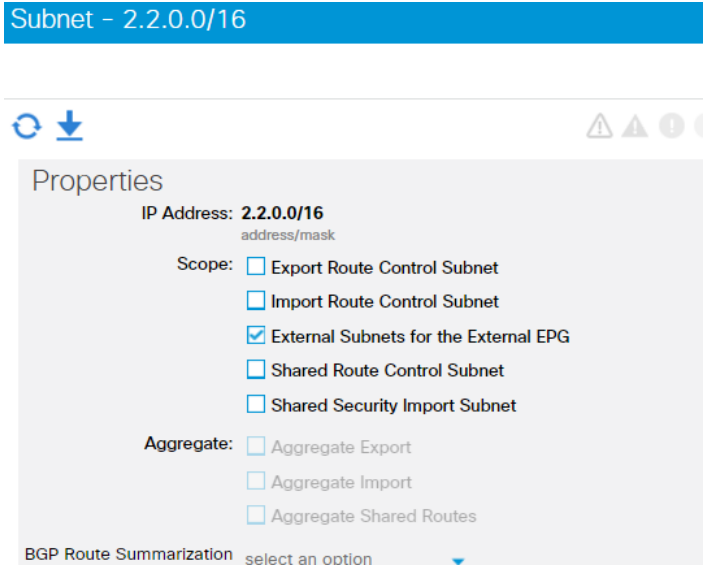





```
bdsol-aci32-leaf2# show ip bgp neigh vrf RD:SB | egrep route-map
  Outbound route-map configured is exp-l3out-L3out-peer-2097153, handle obtained
  Inbound route-map configured is imp-l3out-L3out-peer-2097153, handle obtained
bdsol-aci32-leaf2# show route-map  imp-l3out-L3out-peer-2097153
route-map imp-l3out-L3out-peer-2097153, permit, sequence 7801
  Match clauses:
    ip address prefix-lists: IPv4-peer49153-2097153-exc-ext-inferred-import-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:
bdsol-aci32-leaf2# show ip prefix-list IPv4-peer49153-2097153-exc-ext-inferred-import-dst
ip prefix-list IPv4-peer49153-2097153-exc-ext-inferred-import-dst: 1 entries
   seq 1 permit 2.2.0.0/16
```

3. External subnet for external EPG

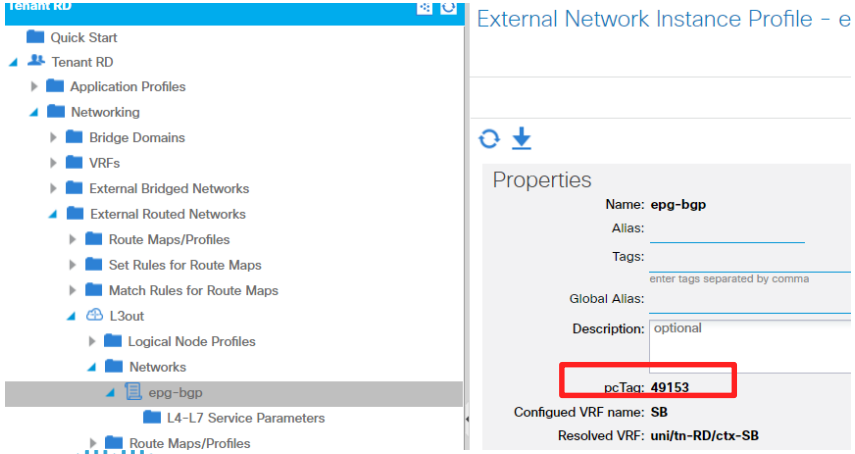
- By far the most important
- Is used for datapath traffic to assign external prefix to the correct l3 epg.
  - Used both for ingress traffic and egress traffic
- Should typically contains the external prefix we expect to receive on that L3 out
- SHOULD NOT contains, BD subnet or subnet received on a different L3 out
- Might contain default internet route 0.0.0.0/0
- Also act on the routing table :
  - On Gen1 – install a route in LPM (longest prefix match) to redirect routes to broadcom)
  - On Gen-2 – also install a route in LPM of the ASIC, but punt all smaller subnet to the EGP
- Keep in mind EPG assignment is based on LPM lookup



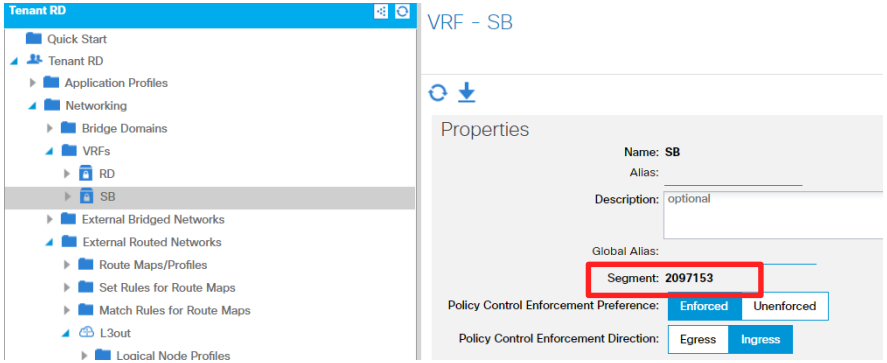
© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

3. External subnet for external EPG finding pcTag and vrf scope

- L3 out epg gives pcTag
- Or pcTag parameter of class l3extInstP



- VRF gives vrf scope
- Or scope parameter of class fvCtx



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

3. External subnet for external EPG Seeing all external subnet list

- Here we lookup all aclqos prefix for vrf with scope 3014656
- We have at least 2 l3 out epg 16390 and 32772
- 10.200.2.0/24 would be assigned to 1390
- 10.200.1.0/24 would be assigned to epg 32772

```
module-1# show system internal  aclqos  prefix  | egrep "==|Scope|3014656"
Vrf-Vni VRF-Id Table-Id          Addr                        Scope Class  Shared Remote Complete
======= ====== ============ ============================ ===== ====== ====== ====== ========

3014656 8      0x7            10.200.2.0/24                 7     16390  0      1      No
3014656 8      0x7            172.16.0.0/16                 7     16390  0      1      No
3014656 8      0x7            10.0.0.0/8                    7     16390  0      1      No
3014656 8      0x7            10.200.1.0/24                 7     32772  0      1      No
```

Deprecated in 3.2
For Gen-2, Gen-3 leaf

© 2017 Cisco and/or its affiliates. All rights reserv

3. External subnet for external EPG Gen-2 : LPM check

- We configured only 10.0.0.0/8 as external subnet on one of the l3 out epg.
- We can see all smaller net of 10.0.0.0/8 that does use it

Deprecated in 3.2
For Gen-2, Gen-3 leaf use
Vsh –c ‘show sys internal policy-mgr prefix’

```
module-1# show system internal  aclqos prefix  | egrep "Vrf|==|10.0.0.0"
Vrf-Vni VRF-Id Table-Id          Addr                        Scope Class  Shared Remote Complete
======= ====== ============ ============================= ===== ====== ====== ====== ========
3014656 7      0x6            10.0.0.0/8                     6     16390  0      1      No
module-1# show forwarding ppfx vrf DC:DC prefix 10.0.0.0/8

IPv4 routes for table DC:DC/base

------------------+------------------+---------------+------------------------
Ppfx              |   prefix         | CLASSID       | INCOMPLETE
------------------+------------------+---------------+------------------------
 10.0.0.0/8                 NO PFX        0x4006           0x1
List of prefix pointing for policy:
 10.200.100.102/32  10.201.101.2/32    10.19.80.1/32      10.19.77.1/32      10.19.61.0/24
 10.19.62.0/24      10.19.81.0/24      10.19.77.0/24      10.19.80.0/24      10.1.1.2/32
 10.200.102.0/24    10.1.1.1/32        10.201.101.255/32  10.201.101.1/32    10.201.101.0/32
 10.201.101.0/24    10.201.100.2/32
```

Show forwarding cli
Still there in 3.2 software

3. External subnet for external EPG Overlapping subnet

On one of the l3 out epg we set 10.0.0.0/8

Subnet - 10.0.0.0/8

Properties

IP Address: **10.0.0.0/8**
address/mask

Scope:
- [ ] Export Route Control Subnet
- [ ] Import Route Control Subnet
- [x] External Subnets for the External EPG
- [ ] Shared Route Control Subnet
- [ ] Shared Security Import Subnet

Aggregate:
- [ ] Aggregate Export

On a second l3 out epg we set 10.19.0.0/16

Subnet - 10.19.0.0/16

Properties

IP Address: **10.19.0.0/16**
address/mask

Scope:
- [ ] Export Route Control Subnet
- [ ] Import Route Control Subnet
- [x] External Subnets for the External EPG
- [ ] Shared Route Control Subnet
- [ ] Shared Security Import Subnet

Aggregate:
- [ ] Aggregate Export

LPM lookup if 10.0.0.0/8 ip we assign to 16390 EPG unless it is a 10.19.0.0/16 which goes to 32772

```
module-1# show system internal  aclqos prefix | egrep "3014656.*10."
3014656 7      0x6            10.0.0.0/8                    6      16390  0      1      No
3014656 7      0x6            10.19.0.0/16                  6      32772  0      0      No
```

Remark

- Note seems that in 2.3 or 3.x software even if you didn’t specify a subnet with Ext Subnet with external EPG we do install always 0.0.0.0/0 in aclqos prefix LPM table by default to sclass 15 even if not configured.

- However it seems we do not push zoning-rule to it . So still need to configure it to get the zoning-rule pushed

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

3.2 and above note

- Starting in 3.2 we do not use anymore aclqos prefix for Gen-2, Gen-3 leaf (EX, FX).
- See CSCvk16258 Aclqos Prefix Output Empty for EX and FX Switches in 3.2
- A replacement Cli is provided in vsh (not vsh_lc) see below

```
bdsol-aci32-leaf2# vsh -c "show system internal policy-mgr prefix"
Requested prefix data

Vrf-Vni VRF-Id Table-Id Table-State  VRF-Name                    Addr                                  Class Shared Remote Complete
======= ======  =========== =======  ============================ ===================================== ====== ====== ====== ========
2654211 7      0x7           Up     RD-BGP:RD                                     170.0.0.0/8   16387   False  True   False
2654211 7      0x7           Up     RD-BGP:RD                                 172.16.99.0/24  16387   False  True   False
2654211 7      0x7           Up     RD-BGP:RD                                 172.16.1.10/32  16387   False  True   False
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Advertise your local BD subnet to a L3 out

- In BD L3 configuration
- Subnet Advertised Externally
- And L3 out associated (here it is called L3out)

Bridge Domain - SB1

Policy | Operational | Stats | Health | Faults

Main | L3 Configurations | Advanced/Trouble

100

Properties

Unicast Routing: ☑

Operational Value for Unicast Routing: **true**

Custom MAC Address: 00:22:BD:F8:19:F1

Virtual MAC Address: Not Configured

Subnets:

Gateway Address	Scope	Primary IP Address	Virtual IP	Subnet Control
10.20.1.254/24	Advertised Externally	False	False	
10.20.11.254/24	Advertised Externally	False	False	

EP Move Detection Mode: ☑ GARP based detection

Associated L3 Outs:

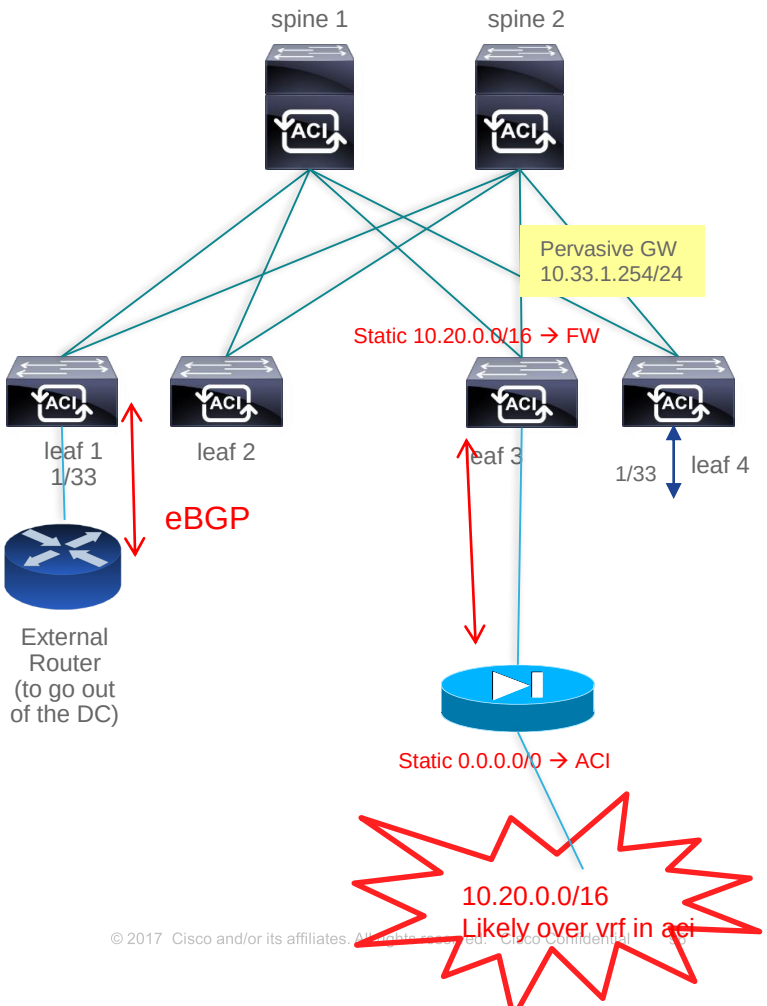
L3 Out
L3out

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Design example

2 layer 3 out in the same VRf
- One BGP session to an external router
To get out of the DC
- One static L3 out to a firewall (going to more secure zone of the DC)

BGP receive all Enterprise prefix
Behind FW we have other prefix from the DC using a supernet of 10.20.0.0/16

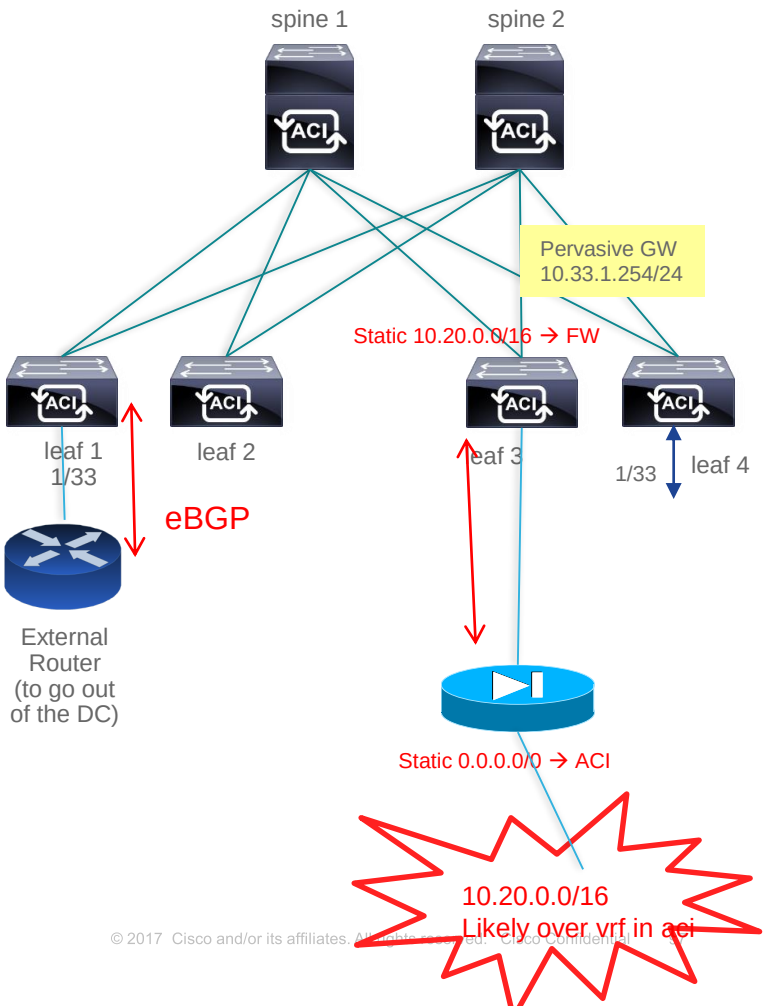


© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Flags to Use

On l3 out bgp :
- 10.20.0.0/16 with only export route control. DO NOT TICK Here external subnet for external EPG
- 0.0.0.0/0 with only external subnet for external EPG

On L3 out to Fw
- 10.20.0.0/16 with external subnet for external EPG



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Single VRF flags best practices

- External subnet for external EPG for the subnet we are expecting to receive from this l3 out
- Export route control subnet for the prefix we want to send on that l3 out
- Normally you never need both flags together
- Keep in mind the LPM routing behavior if you have some overlapping subnet part of a supernet

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Contract verification summary L3 out – vzAny and EPG mix

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Default internet route

- The subnet 0.0.0.0/0 used as External Subnet for external EPG has a special meaning in l3 out epg
- Like other prefix with Ext subnet for Ext Epg, it is inserted with its scope (vrf vnid) in the aclqos prefix LPM routing table But instead of deriving the l3 ext epg pcTag, it derives a special reserved pcTag in aclqos prefix (always 15)
- However zoning-rule for it is a bit more complex (next slides)

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Sclass derivation for default internet route

- If the aclqos tcam lookup in the VRF hits the default entry 0.0.0.0/0
- The current behavior (2.2 and + apparently) is the following
  - For traffic from EPG to L3 out we drive EPG 15
  - For traffic from L3 out to EPG we drive pcTag of the VRF
- Example zoning rule-below in next slide
  - Vrf seg id – 2654211
  - Vrf sclass/pcTag -- 16386
  - Epg sclass/pcTag – 49154
  - L3 out sclass pcTag - 16387

First line goes from vrf sclass to EPG sclass and will be used for L3 out to EPG
2nd line goes from EPG to 15 and will be used for traffic from EPG to L3 out

```
bdsol-aci32-leaf3# show zoning-rule | egrep " 49154 .* 2621441"
4116            49154           15              5               enabled         2654211         permit                          fully_qual(7)
4134            16386           49154           5               enabled         2654211         permit                          fully_qual(7)
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Contract between EPG 49154 abd l3 out only 0.0.0.0/0 as ext sub for ext epg
Line 4116 – from EPG 49154 to 15 (Epg to l3 out drives 15)
Line 4134 -- from 13386 (vrf pcTag) to egp 49154 (l3 out to epg drives VRF pcTag)

```
bdsol-aci32-leaf3# show zoning-rule | egrep " 49155 .* 2621441"
4116            49154           15              5               enabled         2654211         permit                  fully_qual(7)
4134            16386           49154           5               enabled         2654211         permit                  fully_qual(7)
```

Line added when there is a non def internet route in l3 out epg as ext Subnet for external
Here we really drive pcTag of l3 out epg 16387 in both direction…

```
4114            49154           16387           5               enabled         2654211         permit                  fully_qual(7)
4115            16387           49154           5               enabled         2654211         permit                  fully_qual(7)

bdsol-aci32-leaf3# vsh_lc -c 'show sys internal aclqos prefix' | egrep 2654211
2654211 18      0x80000012      0::/0                           15      0       0       No
2654211 18      0x12            0.0.0.0/0                       15      0       0       No
2654211 18      0x12            172.16.1.10/32                  16387   0       1       No
2654211 18      0x12            172.16.99.0/24                  16387   0       1       No
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Reading Zoning-Rule

Rule ID : no meaning, just an index
SrcEPG : sclass/pcTag that will match for the src of the traffic
DstEPG : sclass/pcTag that will match for the dst of the traffic
FilterId : pointer to Filter table (aka protocol to check)
- implicit – not result of config, this is a default rule
- implicate arp – arp is always permitted by default in the vrf
- default – means the filter applied is allow all (or commond/default)

Scope – vrf vnid (or seg)
Action – Permit, deny, log or a combination
Priority – the number is the priority, the lower the better

```
bdsol-aci32-leaf3#  show zoning-rule | egrep "Rule|==|2228225"
Rule ID         SrcEPG          DstEPG          FilterID        operSt          Scope           Action                  Priority
=======         ======          ======          ========        ======          =====           ======                  ========
4159            0               0               implicit        enabled         2228225         deny,log                any_any_any(21)
4197            0               0               implarp         enabled         2228225         permit                  any_any_filter(17)
4202            0               15              implicit        enabled         2228225         deny,log                any_vrf_any_deny(22)
4203            0               49153           implicit        enabled         2228225         permit                  any_dest_any(16)
4206            0               15              default         enabled         2228225         permit                  any_dest_any(16)
4207            32770           0               default         enabled         2228225         permit                  src_any_any(15)
4212            0               16387           implicit        enabled         2228225         permit                  any_dest_any(16)
4217            32772           16386           43              enabled         2228225         permit                  fully_qual(7)
4218            32772           16386           37              enabled         2228225         permit                  fully_qual(7)
4219            16386           32772           38              enabled         2228225         permit                  fully_qual(7)
4216            16386           32772           44              enabled         2228225         permit                  fully_qual(7)
```

Policy rule lookup algorithm

1. Packet enters a leaf , we need to derive pcTAg
    If ingress on regular EPG it will always be the epg pctag
    If ingress on l3 out, it will be
        vrf pcTag if hitting 0.0.0.0/0 in aclqos prefix
        L3 out epg pcTag if hitting a more specific subnet in aclqos prefix table.
2. We make a forwarding decision (route lookup or epm destination)
3. We try to derive dest pcTag
    If dest is l3 out we will again use aclqos prefix to derive dst pcTag
        If hitting 0.0.0.0/0 pctag will be 15
        If hitting more specific subnet it will be the dest l3 out epg pcTag
    If dest is regular EPG, we will know the pctag from EPM lookup
        If EPM is populated we derive directly dst pctag of dest EPG
        If EPM is not populated, we send to fabric (spine or flood) and we insert src pctag in vxlan header to let Dest Leaf to enforce zoning-rule
4. At this stage we should have both src and dst pcTag and we will make zoning-rule lookup by priority either in src leaf or on destination leaf:
    For every line we look in zoning-rule, we consider it to be a hit if either :
        src and dst pcTag resulting from packet lookup do match src and dst epg of the line AND the protocol of the frame is matching the filter of that line. Note src and dst epg in zoning-rule with value 0 will be consider as potential Hit always
    We checked line by priority and scope:
        We first check line in the scope with lowest priority
            If match , check the zoning-filter . If protocol match , apply action
        if no match in lowest priority line or match but protocol do not match the filter , we check every line with the next lowest priority (just higher)

Finding pcTag for a packet (src and dst)

- From Data packet : use ELAM and check sclass and dclass
- Below packet is from pcTag 49155 to pcTag 15
- Next steps :
  - Is it expected pcTag ?
  - What line do we hit in zoning-rule ?

```
module-1(DBG-elam-insel6)# report | egrep class
      sug_lurw_vec.info.ifabric_leaf.dclass: 0xF      - here 15
      sug_lurw_vec.info.ifabric_leaf.sclass: 0xC003   - here 49155
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Verifying expected pcTag from cli – L3 out

- For l3 out you need to use aclqos prefix
  - Below src of packet in previous elam was 10.20.0.1/32 on a l3 out hence using 49155 as src sclass
  - Dst ip was 10.30.0.1 on another l3 out hence using 15

```
module-1(DBG-elam-insel6)# show system internal  aclqos prefix  | egrep "Vrf|==|2228225"
Vrf-Vni VRF-Id Table-Id          Addr                        Class  Shared Remote Complete
======= ====== ============ ============================ ====== ====== ====== ========
2228225 15     0x8000000f     0::/0                         15     0      0      No
2228225 15     0xf            0.0.0.0/0                     15     0      0      No
2228225 15     0xf            10.20.0.1/24                  49155  0      0      No
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

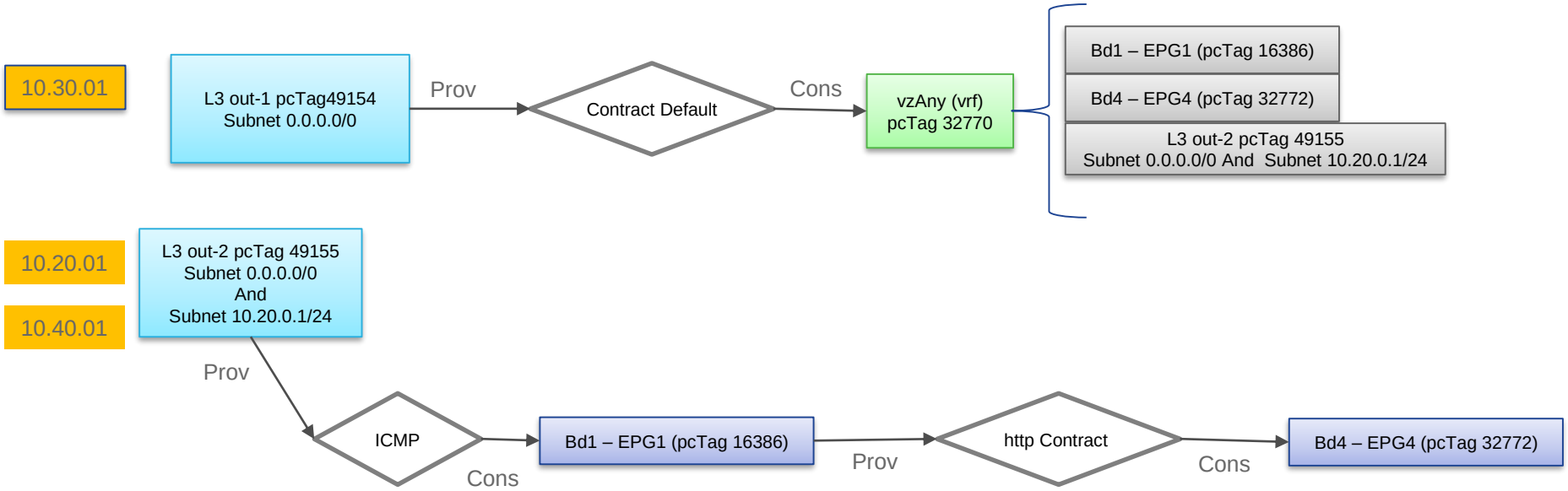
Verifying expected pcTag from cli – regular EPG

```
bdsol-aci32-leaf3# show system internal epm  endpoint  ip
10.143.1.1

MAC : 0050.56a4.0017 ::: Num IPs : 1
IP# 0 : 10.143.1.1 ::: IP# 0 flags :
Vlan id : 55 ::: Vlan vnid : 8412 ::: VRF name : RD-TR:RD
BD vnid : 15826915 ::: VRF vnid : 2228225
Phy If : 0x1a008000 ::: Tunnel If : 0
Interface : Ethernet1/9
Flags : 0x80004c04 ::: sclass : 16386 ::: Ref count : 5
EP Create Timestamp : 06/06/2018 05:41:06.782885
EP Update Timestamp : 06/06/2018 05:48:49.495349
EP Flags : local|IP|MAC|sclass|timer|
::::


bdsol-aci32-leaf3# show system internal epm  endpoint  ip
10.143.2.1

MAC : 0050.56a4.5e29 ::: Num IPs : 1
IP# 0 : 10.143.2.1 ::: IP# 0 flags :
Vlan id : 61 ::: Vlan vnid : 8413 ::: VRF name : RD-TR:RD
BD vnid : 16580494 ::: VRF vnid : 2228225
Phy If : 0x1a008000 ::: Tunnel If : 0x1801002c
Interface : Tunnel44
Flags : 0x80000c80 ::: sclass : 32772 ::: Ref count : 5
EP Create Timestamp : 06/06/2018 05:41:25.617174
EP Update Timestamp : 06/06/2018 05:45:25.516971
EP Flags : on-peer|IP|MAC|sclass|
::::
```

- For regular EPG
- you can always derive src pcTag (src sclass) using epm local learning
- Dest pcTag (dclass in elam) it will be in EPM if EP is learned, if not we will use a default pcTag (1 I believe) and enforcement will be done on egress leaf)

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Lab Contract layout



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

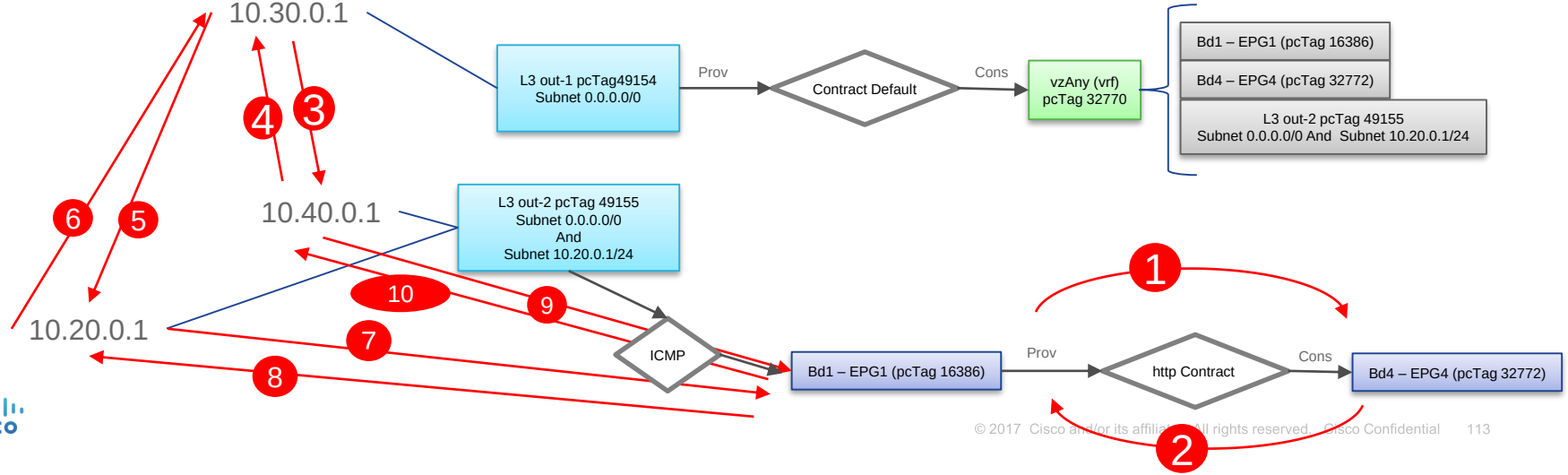
Walking down Zoning-rule example

Order of line check :
4212, 4218,4219,4216 (prio7) - epg1 to epg4
4220, 4221,4222,4223 (prio 7 same level) – l3out-2 to EPG1
4207 (prio 15) - ingress 0.0.0.0/0 L3 out to ANY
4203, 4206 and 4212 (prio 16) - any to egress 0.0.0.0/0 L3 out
4197 (prio 17)
4159 (prio 21)
4202 (prio 22)

32770 is vrf pcTag
16386 is regular EPG1
32772 is regular EPG4
49154 is L3out-1 EPG
49155 is L3out-2 EPG
49153 is BD1 pcTag (not used)
16387 is BD4 pcTag (not used)

```
bdsol-aci32-leaf3#  show zoning-rule | egrep "Rule|==|2228225"
Rule ID         SrcEPG          DstEPG          FilterID        operSt          Scope           Action          Priority
=======         ======          ======          ========        ======          =====           ======          ========
4159            0               0               implicit        enabled         2228225         deny,log        any_any_any(21)
4197            0               0               implarp         enabled         2228225         permit          any_any_filter(17)
4202            0               15              implicit        enabled         2228225         deny,log        any_vrf_any_deny(22)
4203            0               49153           implicit        enabled         2228225         permit          any_dest_any(16)
4206            0               15              default         enabled         2228225         permit          any_dest_any(16)
4207            32770           0               default         enabled         2228225         permit          src_any_any(15)
4212            0               16387           implicit        enabled         2228225         permit          any_dest_any(16)
4217            32772           16386           43              enabled         2228225         permit          fully_qual(7)
4218            32772           16386           37              enabled         2228225         permit          fully_qual(7)
4219            16386           32772           38              enabled         2228225         permit          fully_qual(7)
4216            16386           32772           44              enabled         2228225         permit          fully_qual(7)
4220            16386           15              5               enabled         2228225         permit          fully_qual(7)
4221            32770           16386           5               enabled         2228225         permit          fully_qual(7)
4222            16386           49155           5               enabled         2228225         permit          fully_qual(7)
4223            49155           16386           5               enabled         2228225         permit          fully_qual(7)
```

Rule ID	SrcEPG	DstEPG	FilterID	operSt	Scope	Action	Priority
4159	0	0	implicit	enabled	2228225	deny,log	any_any_any(21)
4197	0	0	implarp	enabled	2228225	permit	any_any_filter(17)
4202	0	15	implicit	enabled	2228225	deny,log	any_vrf_any_deny(22)
4203	0	49153	implicit	enabled	2228225	permit	any_dest_any(16)
4206	0	15	default	enabled	2228225	permit	any_dest_any(16)
4[illegible]	32770	0	default	enabled	2228225	permit	src_any_any(15)
4212	0	16387	implicit	enabled	2228225	permit	any_dest_any(16)
4217	32772	16386	43	enabled	2228225	permit	fully_qual(7)
4218	32772	16386	37	enabled	2228225	permit	fully_qual(7)
4219	16386	32772	38	enabled	2228225	permit	fully_qual(7)
4216	16386	32772	44	enabled	2228225	permit	fully_qual(7)
4220	16386	15	5	enabled	2228225	permit	fully_qual(7)
4221	32770	16386	5	enabled	2228225	permit	fully_qual(7)
4222	16386	49155	5	enabled	2228225	permit	fully_qual(7)
4223	49155	16386	5	enabled	2228225	permit	fully_qual(7)



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Verify your suspicion with policy-mgr stats

```
bdsol-aci32-leaf3# show system internal  policy-mgr stats | egrep "2228225"
Rule (4159) DN (sys/actrl/scope-2228225/rule-2228225-s-any-d-any-f-implicit) Ingress: 0, Egress: 0, Pkts: 22  RevPkts: 0
Rule (4197) DN (sys/actrl/scope-2228225/rule-2228225-s-any-d-any-f-implarp) Ingress: 0, Egress: 0, Pkts: 9  RevPkts: 0
Rule (4202) DN (sys/actrl/scope-2228225/rule-2228225-s-any-d-15-f-implicit) Ingress: 0, Egress: 0, Pkts: 0  RevPkts: 0
Rule (4203) DN (sys/actrl/scope-2228225/rule-2228225-s-any-d-49153-f-implicit) Ingress: 0, Egress: 0, Pkts: 17  RevPkts: 0
Rule (4206) DN (sys/actrl/scope-2228225/rule-2228225-s-any-d-15-f-default) Ingress: 0, Egress: 0, Pkts: 1771210  RevPkts: 0
Rule (4207) DN (sys/actrl/scope-2228225/rule-2228225-s-32770-d-any-f-default) Ingress: 0, Egress: 0, Pkts: 8309  RevPkts: 0
Rule (4212) DN (sys/actrl/scope-2228225/rule-2228225-s-any-d-16387-f-implicit) Ingress: 0, Egress: 0, Pkts: 17  RevPkts: 0
Rule (4216) DN (sys/actrl/scope-2228225/rule-2228225-s-16386-d-32772-f-44) Ingress: 0, Egress: 0, Pkts: 20  RevPkts: 0
Rule (4217) DN (sys/actrl/scope-2228225/rule-2228225-s-32772-d-16386-f-43) Ingress: 0, Egress: 0, Pkts: 0  RevPkts: 0
Rule (4218) DN (sys/actrl/scope-2228225/rule-2228225-s-32772-d-16386-f-37) Ingress: 0, Egress: 0, Pkts: 0  RevPkts: 0
Rule (4219) DN (sys/actrl/scope-2228225/rule-2228225-s-16386-d-32772-f-38) Ingress: 0, Egress: 0, Pkts: 1  RevPkts: 0
Rule (4220) DN (sys/actrl/scope-2228225/rule-2228225-s-16386-d-15-f-5) Ingress: 0, Egress: 0, Pkts: 2  RevPkts: 0
Rule (4221) DN (sys/actrl/scope-2228225/rule-2228225-s-32770-d-16386-f-5) Ingress: 0, Egress: 0, Pkts: 0  RevPkts: 0
Rule (4222) DN (sys/actrl/scope-2228225/rule-2228225-s-16386-d-49155-f-5) Ingress: 0, Egress: 0, Pkts: 17925  RevPkts: 0
Rule (4223) DN (sys/actrl/scope-2228225/rule-2228225-s-49155-d-16386-f-5) Ingress: 0, Egress: 0, Pkts: 17925  RevPkts: 0
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Finding all pcTag of a VRF in GUI

cisco APIC System Tenants Fabric Virtual Networking L4-L7 Services Admin Operations Apps admin

ALL TENANTS | Add Tenant | Tenant Search: Enter name, alias, descr | RD-TR | common | infra | mgmt | RD-BGP

Tenant RD-TR

- Quick Start
- Tenant RD-TR
  - Application Profiles
    - App
  - Networking
    - Bridge Domains
    - VRFs
    - External Bridged Networks
    - External Routed Networks
      - Route Maps/Profiles
      - Set Rules for Route Maps
      - Match Rules for Route Maps
      - L3-1
        - Logical Node Profiles
          - node4
            - Logical Interface Profiles
              - if11
            - Configured Nodes
        - Networks
          - L3-1-EPG
            - L4-L7 Service Parameters
        - Route Maps/Profiles
      - L3-2
        - Logical Node Profiles
          - node3
            - Logical Interface Profiles
              - if11

Tenant - RD-TR

Dashboard Policy Operational Stats Health Faults History Contrac

Flows Packets Resc

Bridge Domain Resource IDs

BD Name	BD Alias	Class ID	Segment ID
BD1		49153	15826915
BD2		49156	15925213

Page 1 Of 1 Objects Per Page: 15 Displaying Objects 1

VRF Resource IDs

VRF Name	VRF Alias	Class ID	Segment ID	Scope
RD		32770	2228225	2228225

Page 1 Of 1 Objects Per Page: 25 Displaying Objects 1

Endpoint Group Resource IDs

Application Profile Name	AP Alias	EPG Name	Class ID	Scope
App		EPG3	49157	2228225
App		EPG4	32772	2228225

Page 1 Of 1 Objects Per Page: 15 Displaying Objects 1

External Networks Instance Profile Resource IDs

EPG Name	EPG Alias	Class ID	Scope
L3-1-EPG		49154	2228225
l3-2-epg		49155	2228225

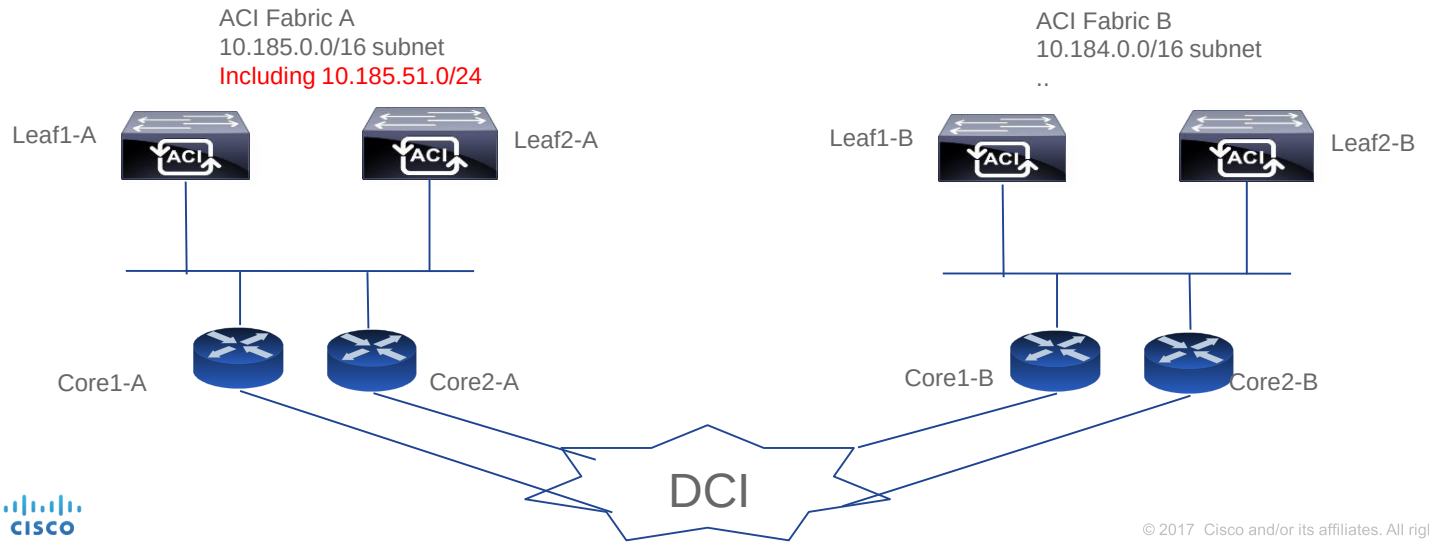
Page 1 Of 1 Objects Per Page: 15 Displaying Objects 1

Routing loop Case Study

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Topology

- Two different ACI fabric
- All leaves and core in OSPF Area 0
- BD subnet advertised by Fabric A in range 10.185/16 and by fabric B in range 10.184/16
- All subnet are LSA type 5 injected by ACI into OSPF domain



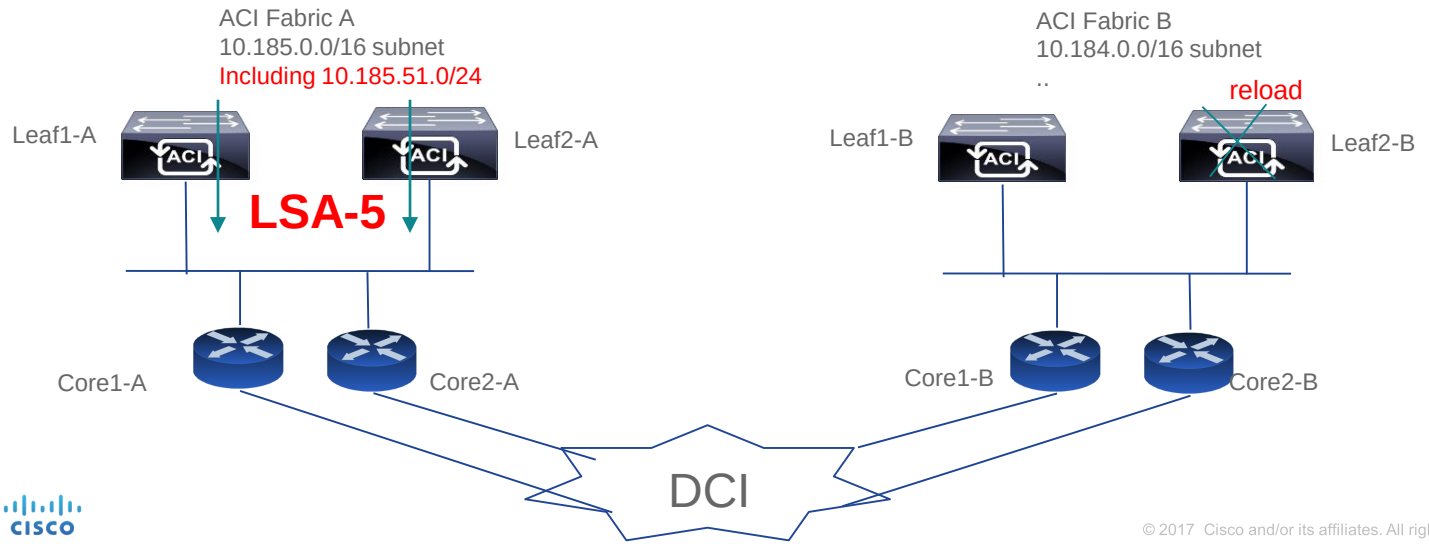
© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Problem description

- Reload of leaf1-B led to loss of routing for prefix 10.185.xxx in core1-A and core2-A !!!

Seen on coreA when leafB reload :

"ospf-1": 10.185.51.0/24, deleting nh 10.185.5.30%Vlan400, metric [110/20] route-type type-2 tag 0x00000000 flags 0x00000000

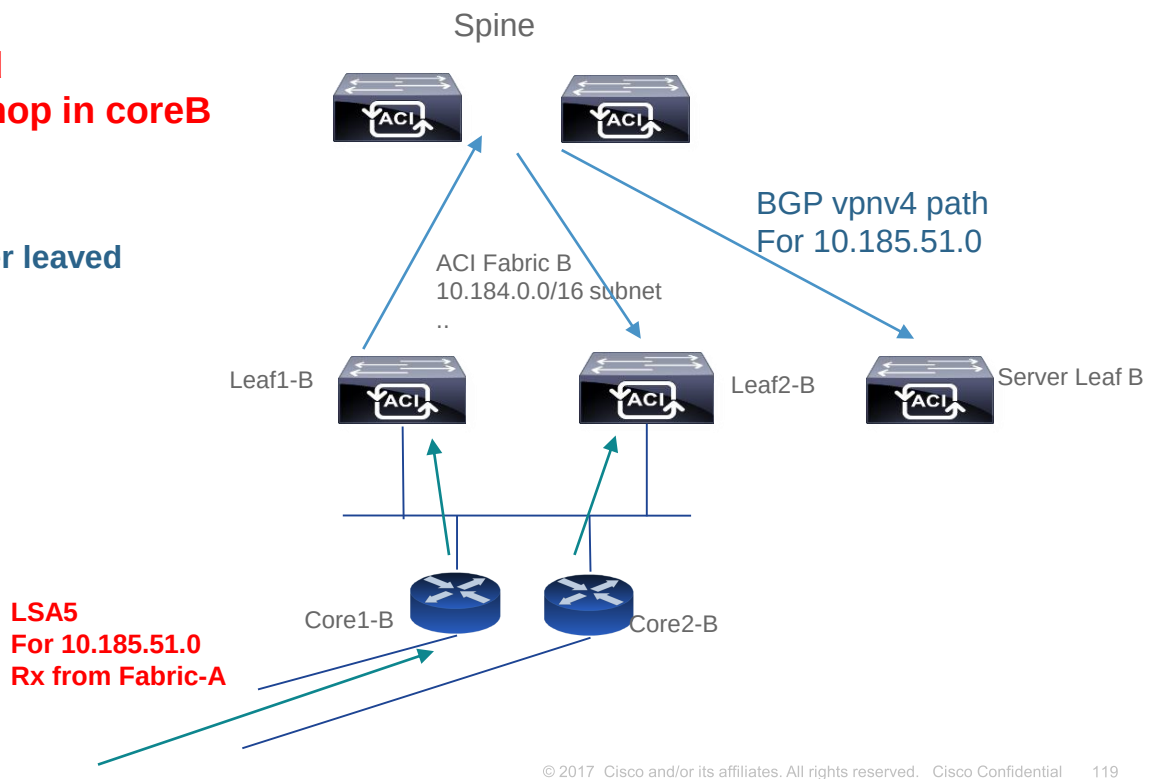


© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

What happens in ACI fabric-B ?

LSA5 is received from Fabric-A
Flooded in OSPF area and received
By both leaf in Fabric-B with Next-hop in coreB

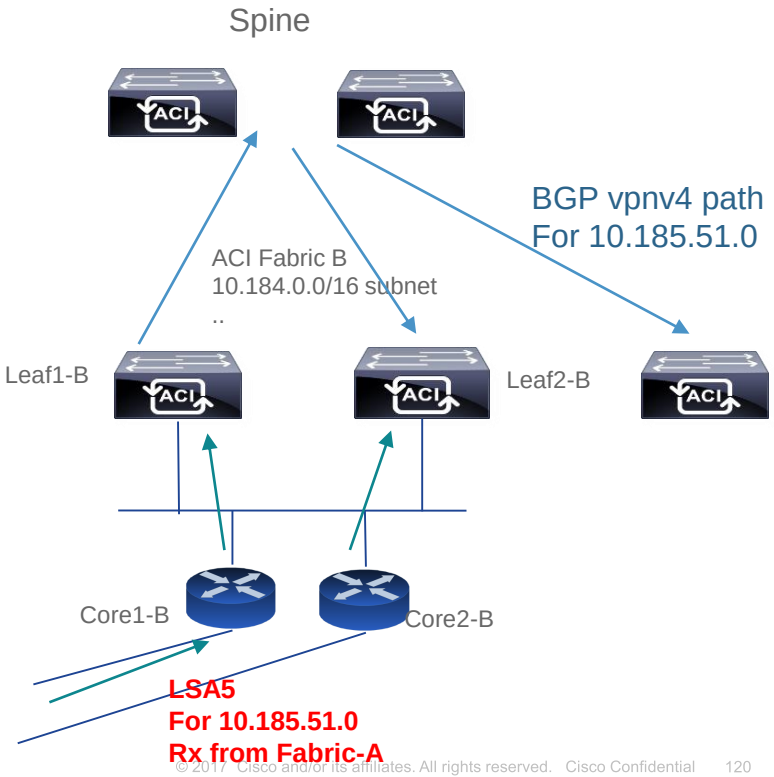
Leaf-B redistribute the OSPF prefix to
BGP VPNv4 to send in ACI fabric to other leaved
Through spine RR



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Result entry in Fabric-B

```
pod2-leaf2# show bgp vpnv4 unicast 10.185.51.0 vrf DC:DC
BGP routing table information for VRF overlay-1, address family VPNv4
Unicast
Route Distinguisher: 10.0.168.93:26    (VRF DC:DC)
BGP routing table entry for 10.185.51.0/24, version 61
Paths: (2 available, best #2)
Flags: (0x80c0002 0x000004) on xmit-list, is not in urib, exported, is in
objstore
  vpn: version 511, (0x100002) on xmit-list
Multipath: eBGP iBGP

  VPN AF advertised path-id 2
  Path type: internal 0xc0000018 0x40040 ref 1, path is valid, not best
reason: Weight
             Imported from 10.0.168.95:35:10.185.51.0/24
 AS-Path: NONE, path sourced internal to AS
!!!!! PATH FROM iBGP coming from neighbor leaf
    10.0.168.95 (metric 3) from 10.0.168.92 (10.0.168.92)
      Origin incomplete, MED 20, localpref 100, weight 0
      Received label 0
      Received path-id 1
      Extcommunity:
          RT:101:2097157
          VNID:2097157
          COST:pre-bestpath:162:110
      Originator: 10.0.168.95 Cluster list: 10.0.168.92

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: redist 0x408 0x40001 ref 0, path is valid, is best path
!!!!! PATH FROM OSPF
  AS-Path: NONE, path locally originated
    0.0.0.0 (metric 0) from 0.0.0.0 (10.0.168.93)
      Origin incomplete, MED 20, localpref 100, weight 32768
      Extcommunity:
          RT:101:2097157
          VNID:2097157
          COST:pre-bestpath:162:110
```



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

What happens when leaf reload ?

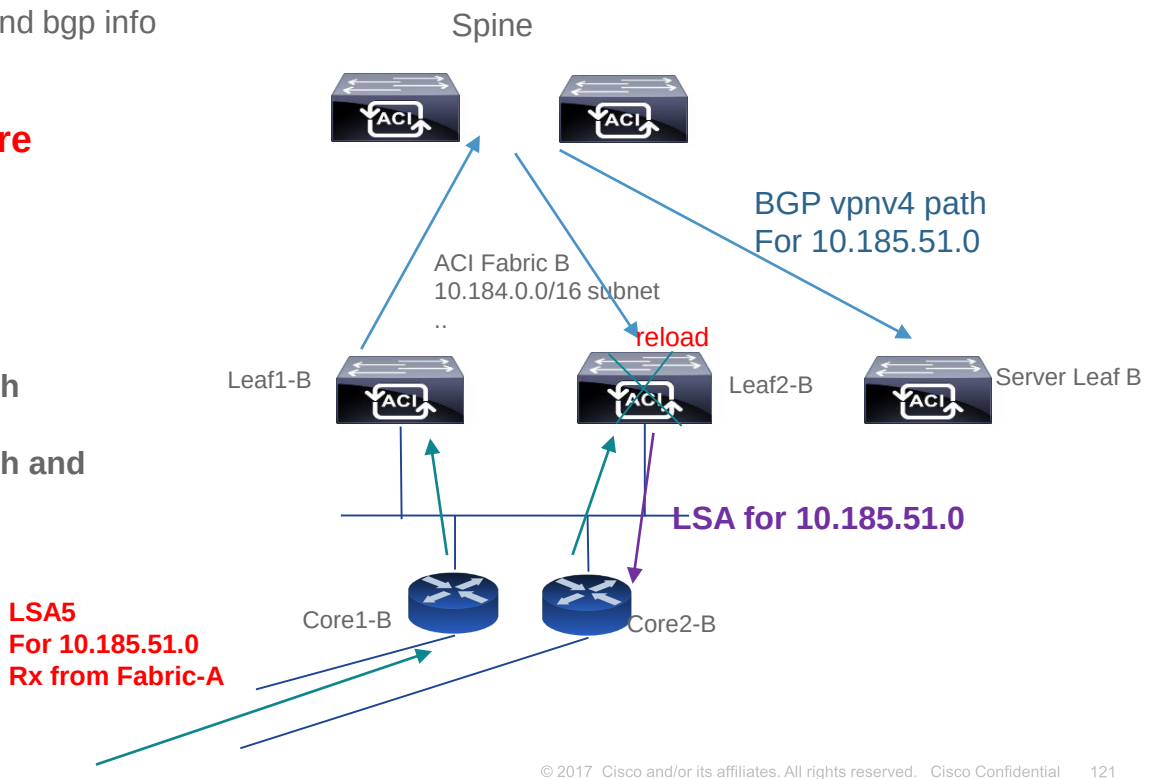
Leaf2-B got reloaded, hence lose all ospf and bgp info
What happens when it comes back up ?

It will get both OSPF path from core
And BGP path from leaf1-B

He may get OSPF before BGP or BGP before OSPF

Assuming BGP path is Rx before OSPF path

Leaf 2-B may redistribute to OSPF BGP path and Inject it in Area 0 (Based on route-map) !



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Reloading leaf output

Route-map that determine
What we redistribute in ospf area

```
pod2-leaf2# show ip ospf  vrf DC:DC
..
   Area BACKBONE(0.0.0.0)
        Area has existed for 00:20:55
        Interfaces in this area: 2 Active interfaces: 2
        Passive interfaces: 1  Loopback interfaces: 1
        No authentication available
        SPF calculation has run 9 times
         Last SPF ran for 0.002716s
        Area ranges are
        Area-filter in 'exp-ctx-proto-2097157'
        Number of LSAs: 14, checksum sum 0x62e65

The above route-map is used to decide what we redistribute to OSPF :
pod2-leaf2# show  route-map exp-ctx-proto-2097157
route-map exp-ctx-proto-2097157, permit, sequence 19801
  Match clauses:
    ip address prefix-lists: IPv4-proto49154-2097157-agg-ext-inferred-export-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:
    tag 4294967295
pod2-leaf2# show ip prefix-list IPv4-proto49154-2097157-agg-ext-inferred-export-dst
ip prefix-list IPv4-proto49154-2097157-agg-ext-inferred-export-dst: 1 entries
   seq 1 permit 0.0.0.0/0 le 32
```

We have a prefix-list that allow everything !!!

Resulting wrong routing entry in core

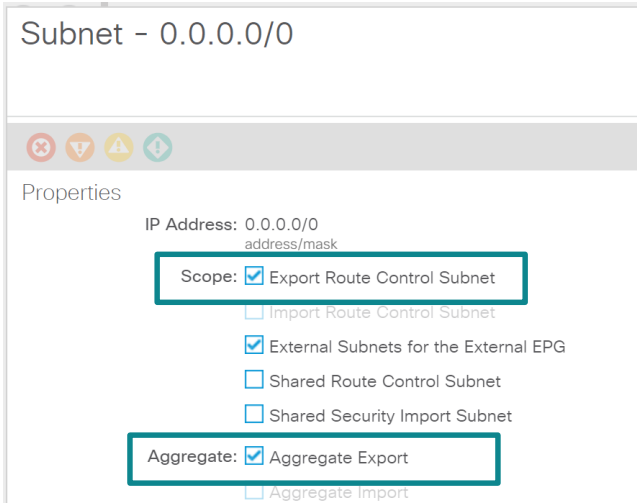
```
CORE# show ip route 10.185.51.0/24 vrf DC
IP Route Table for VRF "BT"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>
10.185.51.0/24, ubest/mbest: 1/0
    *via 10.184.3.4, Vlan883, [110/1], 00:00:01, ospf-1, type-2, tag 4294967295
bdsol-n6004-02#
```

Wrong next-hop and route tag from ACI

Problem correct itself after couple of seconds (when reloading leaf got correct LSA it stops send BGP path to OSPF)

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Root case



- **Subnet flag in l3 out , export route control with aggregate export is set for al routes**
- External subnet for external EPG is used for what we will accept from that L3 out
- Export route control subnet is used to send transit routing to that l3 out
- They should normally not be set together

If 0.0.0.0/0 is external subnet
If we need transit routing, we need to set specific prefix from different L3 out as export route control

Export Route control – What to remember

- Aci do not have any mechanism such as SOO in MPLS VPN to prevent a route received from a site (a l3 out here) to send it back to the same site on a different leaf on same l3 out
- Export route control SHOULD exclude any subnet which may be receive on that L3 out !

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

MP-BGP review
iBGP in ACI detail

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

MP-BGP update packet format (RFC 4760)

MP-BGP allows to send
Any type of protocol in the NLRI.
MP_REACH_NLRI attribute
Specifies next-hop and which address family (AFI/SAFI)

Extended community carries the Route-target information

Route Distinguisher are part of the NLRI prefix

NLRI

```
+-------------------------------------------------------+
| Address Family Identifier (2 octets)                  |
+-------------------------------------------------------+
| Subsequent Address Family Identifier (1 octet)        |
+-------------------------------------------------------+
| Length of Next Hop Network Address (1 octet)          |
+-------------------------------------------------------+
| Network Address of Next Hop (variable)                |
+-------------------------------------------------------+
| Number of SNPAs (1 octet)                             |
+-------------------------------------------------------+
| Length of first SNPA(1 octet)                         |
+-------------------------------------------------------+
| First SNPA (variable)                                 |
+-------------------------------------------------------+
```

```
+-----------------------------------------------------+
|   Withdrawn Routes Length (2 octets)                |
+-----------------------------------------------------+
|   Withdrawn Routes (variable)                       |
+-----------------------------------------------------+
|   Total Path Attribute Length (2 octets)            |
+-----------------------------------------------------+
|   Path Attributes (variable)                        |
+-----------------------------------------------------+
|   Network Layer Reachability Information (variable) |
+-----------------------------------------------------+
```

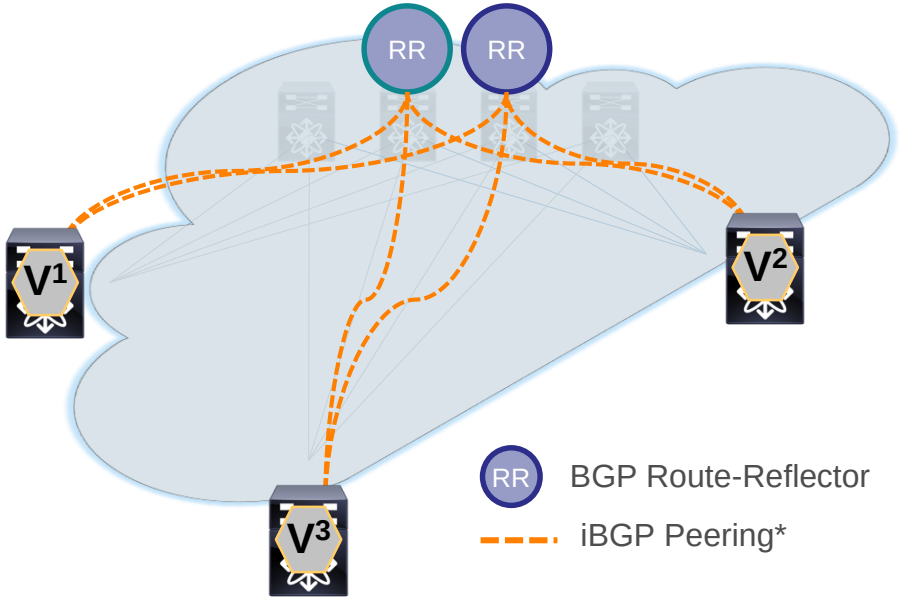
Attribute can be :
ORIGIN, AS_PATH, NEXT_HOP,MED ,
LOCAL_PREF, ATOMIC_AGGREGATE,
AGGREGATOR,COMMUNITY,
EXTENDED_COMMUNITY (contains
Route-target amont other things ,
MP_REACH_NLRI

NLRI for VPNv4 for example is
extended to contains :
- Label stack
- Route Distinguisher
- IPv4 address and masks,

Multiprotocol BGP (MP-BGP) Primer

For Your Reference

- Multiprotocol BGP (MP-BGP)
- Extension to Border Gateway Protocol (BGP) - RFC 4760
- VPN Address-Family:
  - Allows different types of address families (e.g. VPNv4, VPNv6, L2VPN EVPN (RFC 7432), MVPN)
  - Information transported across single BGP peering



*eBGP supported without BGP Route-Reflector

IETF

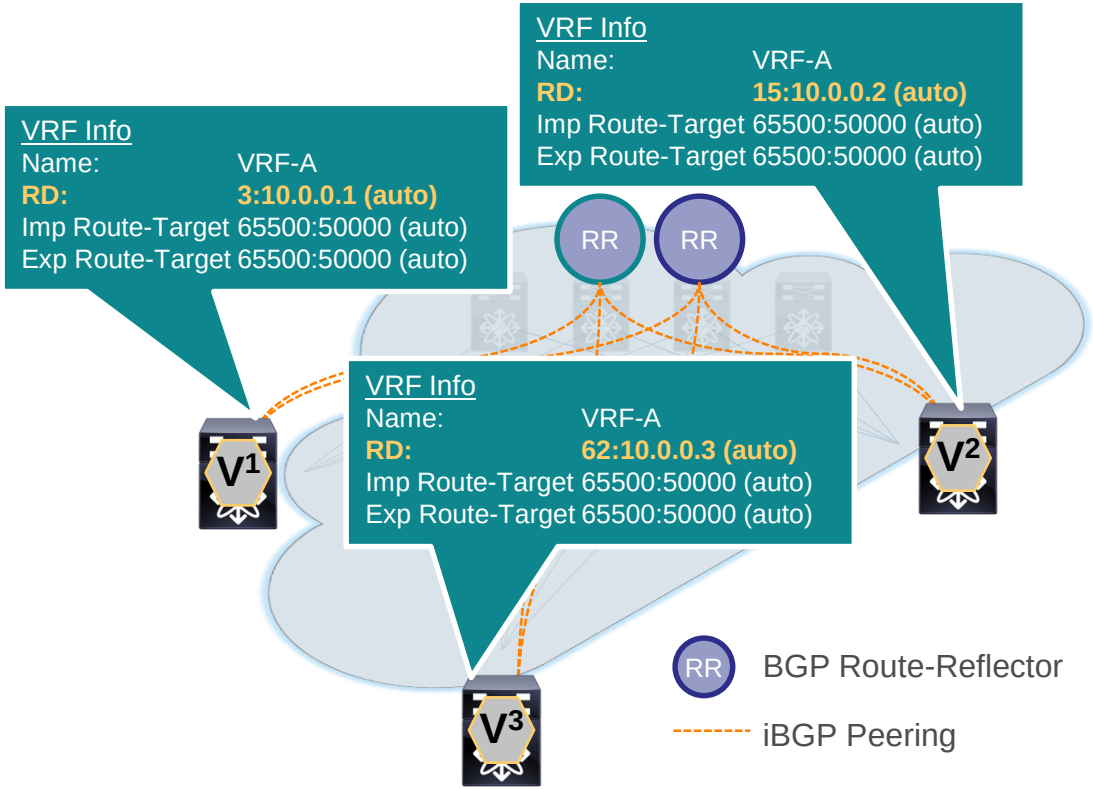
https://tools.ietf.org/html/draft-ietf-bess-evpn-overlay-04

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Multiprotocol BGP (MP-BGP) Primer

For Your Reference

- VPN segmentation for tenant routing (Multi-Tenancy)
  - Route Distinguisher (RD)
  - 8-byte field of VRF parameters
  - value to make VPN prefix unique:
    - RD + VPN prefix

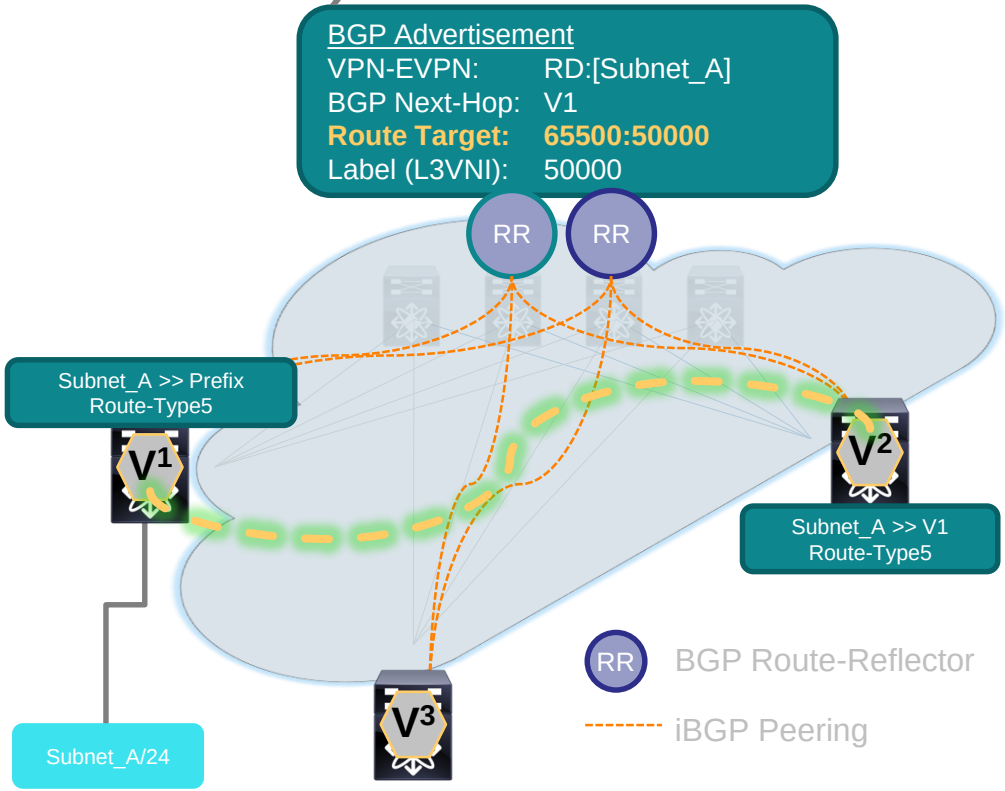


© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Multiprotocol BGP (MP-BGP) Primer

For Your Reference

- VPN Segmentation for tenant routing (Multi-Tenancy)
- Selective distribute VPN routes - Route Target (RT)
  - 8-byte field of VRF parameter
  - unique value to define the import/export rules for VPN prefix



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

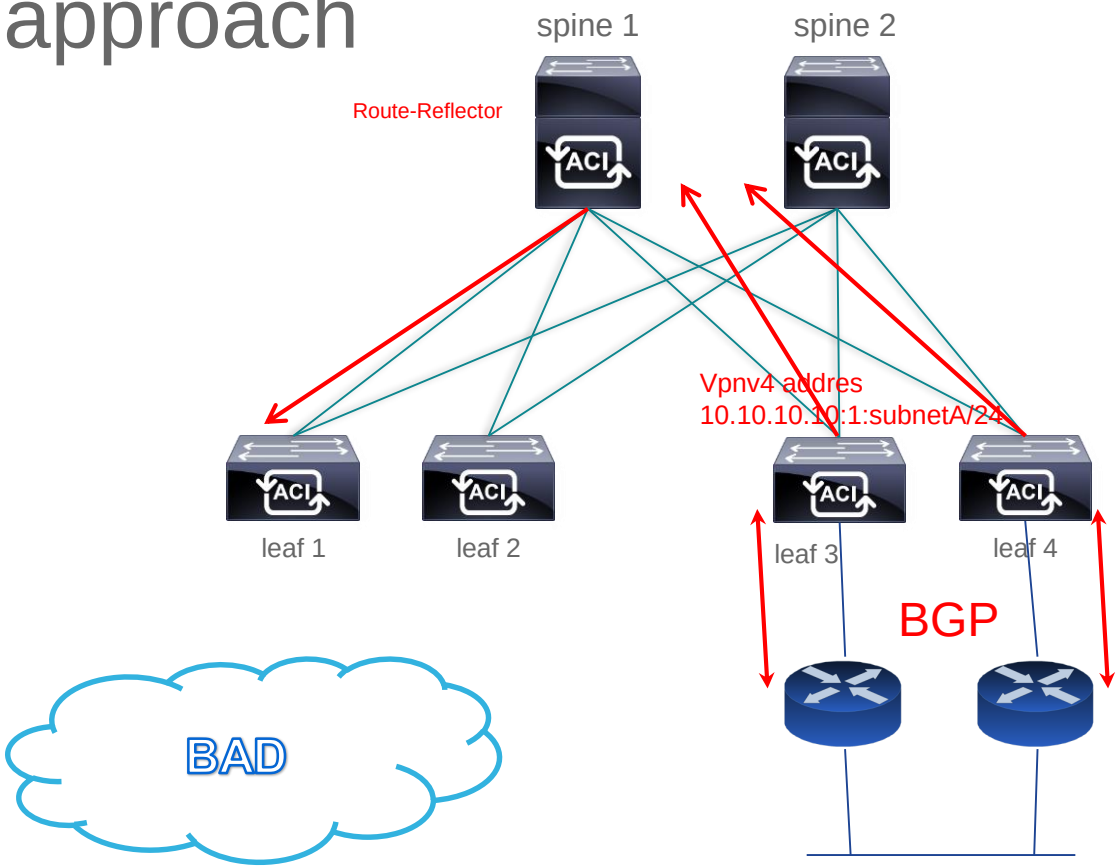
Route Distinguisher in ACI

- In regular vpnv4 network we can either use same RD on every node for the same vrf or use a unique RD per node in the same VRF.

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Same RD per node approach

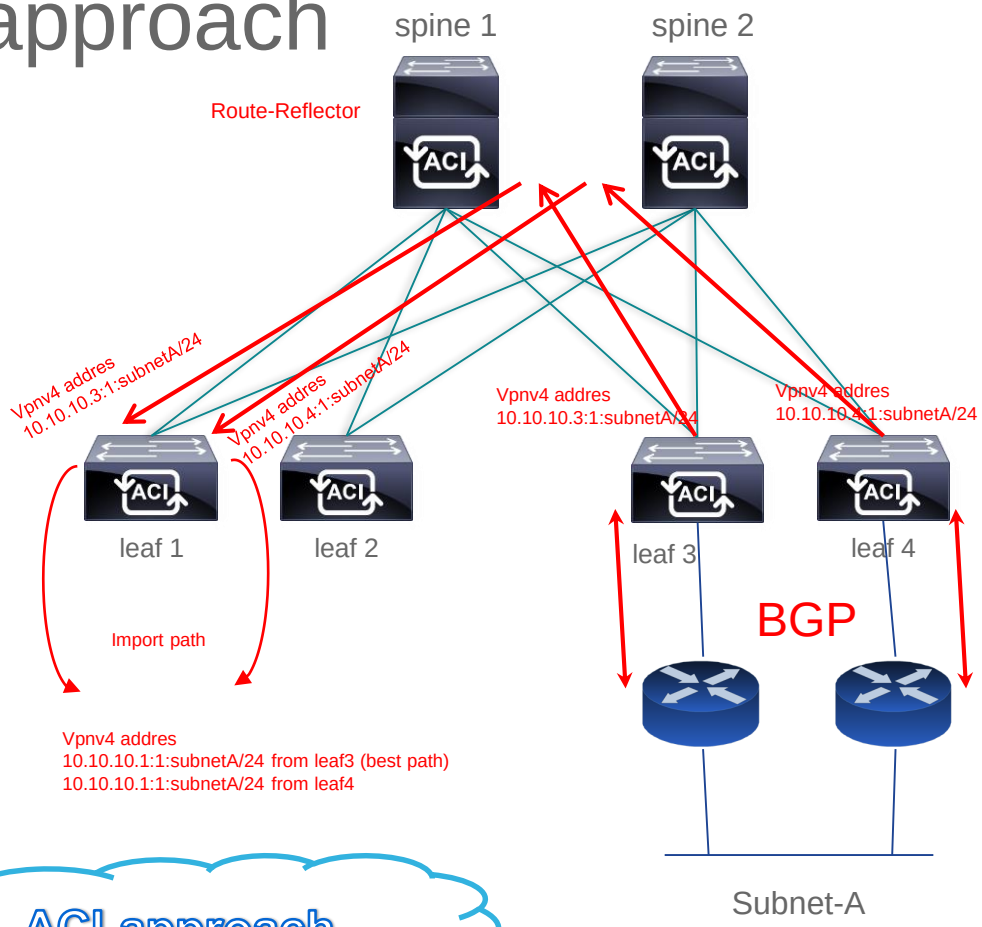
- We suppose all leaf in vrf uses same RD (say 10.10.10.10:1)
- Both Leaf 3 and 4 do send
- Vpnv4 10.10.10.10:1:subnetA/24 to Route Reflector
- Route reflector gets twice same exact vpnv4 prefix and apply best path, it only reflect one Path to Leaf 1 (say from leaf3)
- Leaf 1 gets vpnv4 address with NH leaf3 TEP.
- If leaf3 goes down no backup till we receive new prefix from RR



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Different RD per node approach

- We suppose all leaf in vrf uses same RD (say 10.10.10.leafID:1)
- Leaf 3 do send Vpnv4 10.10.10.3:1:subnetA/24 to Route Reflector
- Leaf 4 do send Vpnv4 10.10.10.4:1:subnetA/24 to Route Reflector
- Route reflector gets two different vpnv4 prefix and reflect both to leaf 1
- Leaf 1 gets two vpnv4 address with one from leaf3, one from leaf4. It does import both to its own RD (10.10.10.1:1:subnetA/24)
- Then leaf 1 applies best path selection and only install one in RIB (say leaf 3)
- If leaf3 goes down we can select directly back path to leaf4

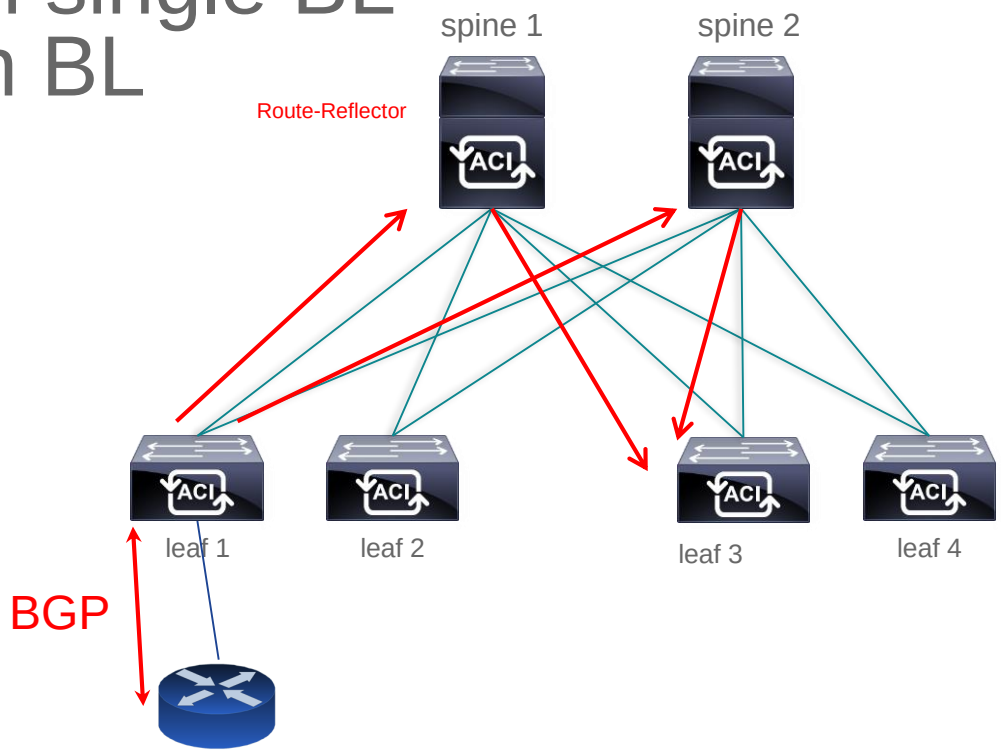


ACI approach

s affiliates. All rights reserved. Cisco Confidential

Example 1 – BGP with single BL Prefix received on non BL

- Leaf 1 has eBGP connection and gets prefix 172.16.1.0/24 in vrf L3:L3.Let's look on leaf 3
- Path send to both spine , each spine reflect it.




```
pod2-leaf3# show bgp vpnv4 unicast 172.16.1.0 vrf overlay-1  | egrep "Route Dist"
Route Distinguisher: 10.0.168.91:4    (VRF L3:L3)
Route Distinguisher: 10.0.168.95:10
```

```
pod2-leaf3# acidiag fnvread | egrep "leaf1|leaf3"
     101       pod2-leaf1      SAL1820SMHV     10.0.168.95/32    leaf       1         active   0
     103       pod2-leaf3      SAL1818RUHM     10.0.168.91/32    leaf       1         active   0
```

Example 1 - Reading BGP – Rx path

```
pod2-leaf3# show bgp vpnv4 unicast 172.16.1.0 vrf overlay-1
BGP routing table information for VRF overlay-1, address family VPNv4
Unicast
Route Distinguisher: 10.0.168.95:10
BGP routing table entry for 172.16.1.0/24, version 943
Paths: (2 available, best #1)
Flags: (0x000002) on xmit-list, is not in urib
Multipath: eBGP iBGP

  Advertised path-id 1
  Path type: internal, path is valid, is best path
  AS-Path: 100 99 , path sourced external to AS
    10.0.168.95 (metric 3) from 10.0.168.92 (10.0.168.92)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Extcommunity:
          RT:101:2785280
          VNID:2785280
      Originator: 10.0.168.95 Cluster list: 10.0.168.92

  Path type: internal, path is valid, not best reason: Neighbor Address
  AS-Path: 100 99 , path sourced external to AS
    10.0.168.95 (metric 3) from 10.0.168.94 (10.0.168.94)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Extcommunity:
          RT:101:2785280
          VNID:2785280
      Originator: 10.0.168.95 Cluster list: 10.0.168.94

  Path-id 1 not advertised to any peer
```

This is the prefix we receive .
It is vpnv4 prefix in RD 10.0.168.95:10 (leaf1 RD) for 172.16.1.0/24

We receive that path twice (one from each Route-Reflector)

```
pod2-leaf3# acidiag fnvread | egrep "spine"
 201    pod2-spine1    10.0.168.92/32   spine    1          active    0
 202    pod2-spine2    10.0.168.94/32   spine    1          active    0
```

In each we have same route-target : 101:2785280

BGP got two path identical (same RD) hence it runs best path and choose the first

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Example 1 - Reading BGP (cont.) imported path

```
BGP routing table information for VRF overlay-1, address family VPNv4
Unicast
Route Distinguisher: 10.0.168.91:4     (VRF L3:L3)
BGP routing table entry for 172.16.1.0/24, version 341
Paths: (1 available, best #1)
Flags: (0x08001a) on xmit-list, is in urib, is best urib route
  vpn: version 944, (0x100002) on xmit-list
Multipath: eBGP iBGP

  Advertised path-id 1
  Path type: internal, path is valid, is best path
             Imported from 10.0.168.95:10:172.16.1.0/24
  AS-Path: 100 99 , path sourced external to AS
    10.0.168.95 (metric 3) from 10.0.168.92 (10.0.168.92)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Extcommunity:
          RT:101:2785280
          VNID:2785280
      Originator: 10.0.168.95 Cluster list: 10.0.168.92

  VRF advertise information:
  Path-id 1 not advertised to any peer

  VPN AF advertise information:
```

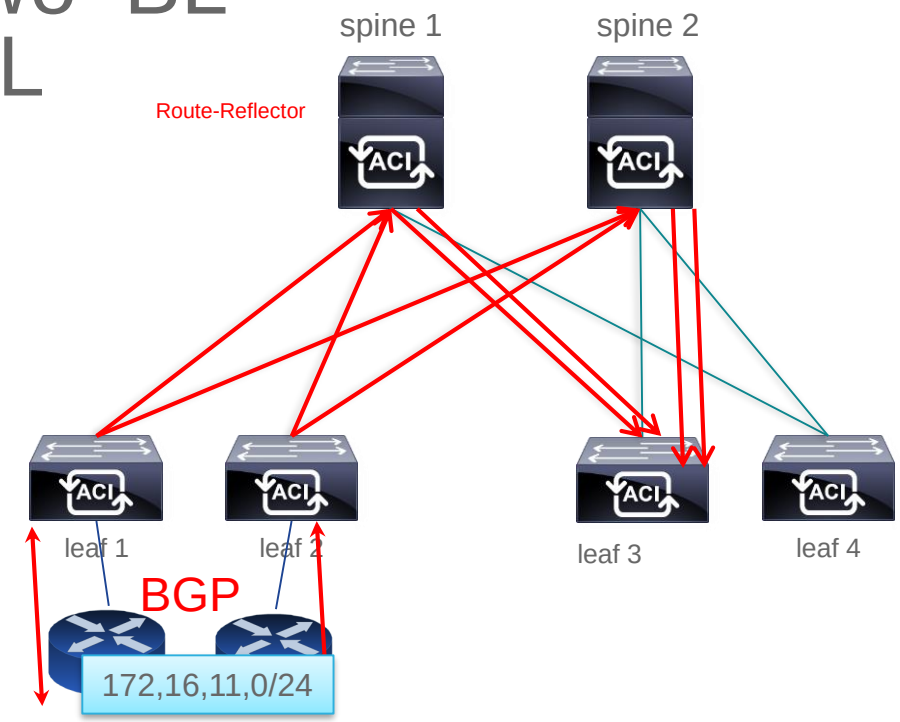
This is the prefix we imported in our own RD 10.0.168.91:4 (here locally we know it is vrf L3:L3) .
We see it is imported prefix from 10.0.168.95:10:172.16.1.0/24 (leaf1-RD:subnetA/24)

Note that only the best path from previous slide is imported (from Spine with ip 10.0.168.92)

We can import it because the Route-Target do match what we allow for import in the vrf

```
pod2-leaf3# show bgp process vrf L3:L3 | egrep -A 1
"RT"
    Export RT list:
        101:2785280
    Import RT list:
        101:2785280

RT is AS nr + ctx vnid
```

Example 2 – BGP with two BL Prefix received on non BL



- Leaf 1 and leaf 2 has eBGP connection and gets prefix 172.16.11.0/24 in vrf LB:LB.Let's look on leaf 3
- Path send to both spine from both leaf , each spine reflect it.

```
pod2-leaf3# show bgp vpnv4 unicast 172.16.11.0 vrf
overlay-1  | egrep "Route Dist"
Route Distinguisher: 10.0.168.91:11     (VRF LB:LB)
Route Distinguisher: 10.0.168.93:21
Route Distinguisher: 10.0.168.95:31
```

```
pod2-leaf3# acidiag fnvread | egrep "leaf1|leaf3"
     101       pod2-leaf1      SAL1820SMHV     10.0.168.95/32    leaf       1         active    0
     102       pod2-leaf2      SAL1816QVBC     10.0.168.93/32    leaf       1         active    0
     103       pod2-leaf3      SAL1818RUHM     10.0.168.91/32    leaf       1         active    0
```

Example 2 – BGP path on leaf 3

Path received with RD of leaf 1

```
BGP routing table information for VRF overlay-1, address
family VPNv4 Unicast
Route Distinguisher: 10.0.168.95:31
BGP routing table entry for 172.16.11.0/24, version 1197
Paths: (2 available, best #2)
Flags: (0x000002) on xmit-list, is not in urib
Multipath: eBGP iBGP

  Path type: internal, path is valid, not best reason:
Neighbor Address
  AS-Path: 99 , path sourced external to AS
    10.0.168.95 (metric 3) from 10.0.168.94 (10.0.168.94)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Extcommunity:
          RT:101:2588672
          VNID:2588672
      Originator: 10.0.168.95 Cluster list: 10.0.168.94

  Advertised path-id 1
  Path type: internal, path is valid, is best path
  AS-Path: 99 , path sourced external to AS
    10.0.168.95 (metric 3) from 10.0.168.92 (10.0.168.92)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Extcommunity:
          RT:101:2588672
          VNID:2588672
      Originator: 10.0.168.95 Cluster list: 10.0.168.92

  Path-id 1 not advertised to any peer
```

In each RD as we have Two path with same vpnv4 address(one from each RR), hence we run Best path selection for each RD.

Path received with RD of leaf 2

```
BGP routing table information for VRF overlay-1, address
family VPNv4 Unicast
Route Distinguisher: 10.0.168.93:21
BGP routing table entry for 172.16.11.0/24, version 1018
Paths: (2 available, best #2)
Flags: (0x000002) on xmit-list, is not in urib
Multipath: eBGP iBGP

  Path type: internal, path is valid, not best reason:
ighbor Address
  AS-Path: 99 , path sourced external to AS
    10.0.168.93 (metric 3) from 10.0.168.94 (10.0.168.94)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Extcommunity:
          RT:101:2588672
          VNID:2588672
      Originator: 10.0.168.93 Cluster list: 10.0.168.94

  Advertised path-id 1
  Path type: internal, path is valid, is best path
  AS-Path: 99 , path sourced external to AS
    10.0.168.93 (metric 3) from 10.0.168.92 (10.0.168.92)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Extcommunity:
          RT:101:2588672
          VNID:2588672
      Originator: 10.0.168.93 Cluster list: 10.0.168.92

  Path-id 1 not advertised to any peer
```

Example 2 – BGP path on leaf 3 (cont.)

```
pod2-leaf3# show bgp vpnv4 unicast 172.16.11.0 vrf overlay-1
BGP routing table information for VRF overlay-1, address family VPNv4 Unicast
Route Distinguisher: 10.0.168.91:11     (VRF LB:LB)
BGP routing table entry for 172.16.11.0/24, version 61
Paths: (2 available, best #2)
Flags: (0x08001a) on xmit-list, is in urib, is best urib route
  vpn: version 1196, (0x100002) on xmit-list
Multipath: eBGP iBGP

  Path type: internal, path is valid, not best reason: Router Id, multipath
             Imported from 10.0.168.95:31:172.16.11.0/24
  AS-Path: 99 , path sourced external to AS
    10.0.168.95 (metric 3) from 10.0.168.92 (10.0.168.92)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Extcommunity:
          RT:101:2588672
          VNID:2588672
      Originator: 10.0.168.95 Cluster list: 10.0.168.92

  Advertised path-id 1
  Path type: internal, path is valid, is best path
             Imported from 10.0.168.93:21:172.16.11.0/24
  AS-Path: 99 , path sourced external to AS
    10.0.168.93 (metric 3) from 10.0.168.92 (10.0.168.92)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Extcommunity:
          RT:101:2588672
          VNID:2588672
      Originator: 10.0.168.93 Cluster list: 10.0.168.9

  VRF advertise information:
  Path-id 1 not advertised to any peer
```

Path imported with RD of leaf 3

Imported from leaf 1 RD

We again have two path in RD 10.0.168.91:11 for the same prefix so we run best path again See the **Multipath** keyword. ACI does iBGP multipath by default so we will install both path in Rib

Imported from leaf 2 RD

```
pod2-leaf3# show ip route vrf LB:LB 172.16.11.0

172.16.11.0/24, ubest/mbest: 2/0
    *via 10.0.168.93%overlay-1, [200/0], 16:47:30, bgp-101, internal, tag 99 (mpls-vpn)
    *via 10.0.168.95%overlay-1, [200/0], 00:42:13, bgp-101, internal, tag 99 (mpls-vpn)
```

RD and Route-Target in ACI

- RD format :
  - Leaf-PTEP:id – id is a number identifying a vrf
- Route-Target format
  - BGP_ASN:vrf_vnid

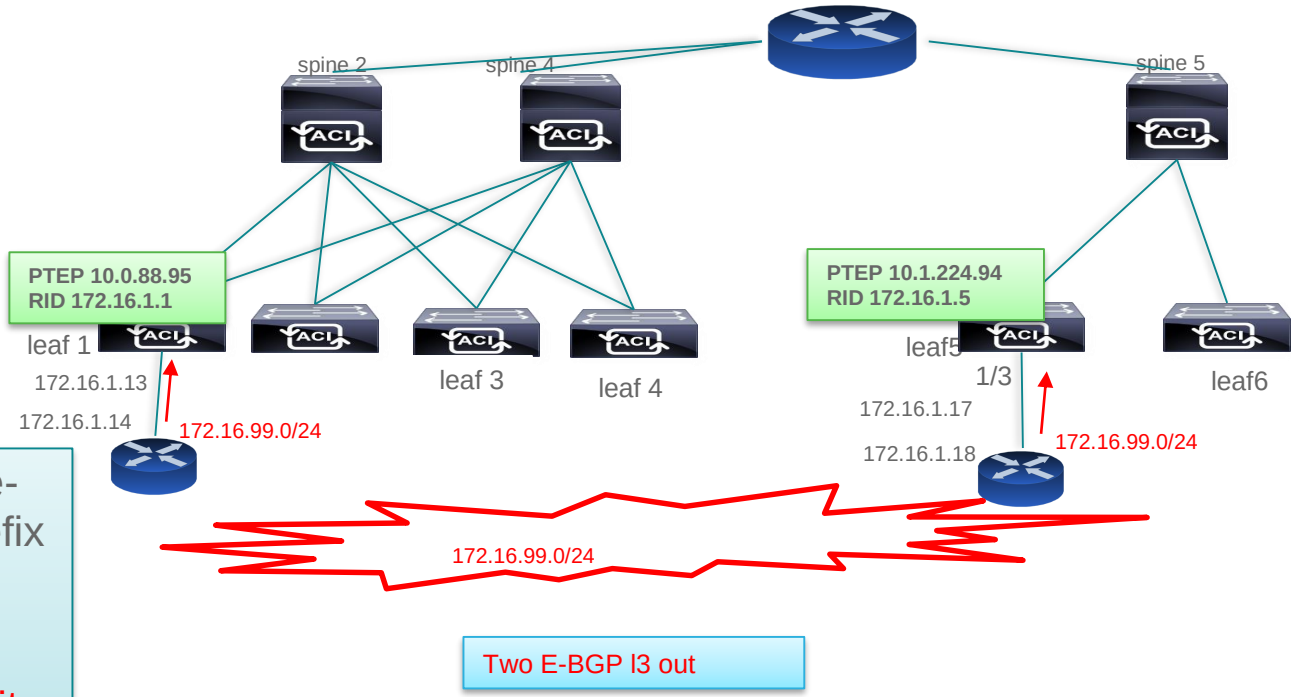
© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Selecting preferred egress leaf to leave Fabric

Import Route Control for BGP

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Example 1 – BGP route control



- Fabric have two l3 out using e-bgp and receive the same prefix from both (172.16.99.0/24)
- Can we force the traffic out of ACI to this prefix to always exit from leaf 1 and only use leaf 5 if we lose primary exit point ?

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Routing Table with basic config

Leaf 1 and leaf 5 both prefer the External BGP route with admin distance 20

```
bdsol-aci32-leaf1#  show ip route 172.16.99.0 vrf RD-BGP:RD
172.16.99.0/24, ubest/mbest: 1/0
    *via 172.16.1.14, [20/0], 00:14:03, bgp-132, external, tag
200
```

```
bdsol-aci32-leaf5#  show ip route 172.16.99.0 vrf RD-BGP:RD

172.16.99.0/24, ubest/mbest: 1/0
    *via 172.16.1.18, [20/0], 00:01:21, bgp-132, external, tag 200
```

Leaf 3 selects ibgp route to leaf1

```
bdsol-aci32-leaf3#  show ip route 172.16.99.0 vrf RD-BGP:RD
172.16.99.0/24, ubest/mbest: 1/0
    *via 10.0.88.95%overlay-1, [200/0], 00:02:58, bgp-132,
internal, tag 200
         recursive next hop: 10.0.88.95/32%overlay-1
```

Leaf 6 selects ibgp route to leaf5

```
bdsol-aci32-leaf6#  show ip route 172.16.99.0 vrf RD-BGP:RD

172.16.99.0/24, ubest/mbest: 1/0
    *via 10.1.224.94%overlay-1, [200/0], 00:15:01, bgp-132,
internal, tag 200 (mpls-vpn)
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

BGP table on border with basic config

Leaf 1 and leaf 5 both prefer the External BGP path over the iBGP path (RX through vpnv4)

```
bdsol-aci32-leaf1# show ip bgp 172.16.99.0 vrf RD-BGP:RD
BGP routing table information for VRF RD-BGP:RD, address family
IPv4 Unicast
BGP routing table entry for 172.16.99.0/24, version 42 dest ptr
0xab12a650
Paths: (2 available, best #2)
Flags: (0x80c001a 00000000) on xmit-list, is in urib, is best
urib route, exported
  vpn: version 989, (0x100002) on xmit-list
Multipath: eBGP iBGP

  VPN AF advertised path-id 2
  Path type: internal 0xc0000018 0x40 ref 1, path is valid, not
best reason: Internal path
             Imported from 10.1.224.94:16:172.16.99.0/24
  AS-Path: 200 , path sourced external to AS
    10.1.224.94 (metric 64) from 10.0.88.93 (10.10.32.214)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Received path-id 1
      Extcommunity:
          RT:132:2654211
          COST:pre-bestpath:165:2415919104
          VNID:2654211
      Originator: 10.1.224.94 Cluster list: 10.10.32.214
10.10.32.202

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: external 0x28 0x0 ref 0, path is valid, is best
path
  AS-Path: 200 , path sourced external to AS
    172.16.1.14 (metric 0) from 172.16.1.14 (172.16.1.21)
      Origin IGP, MED not set, localpref 100, weight 0
      Extcommunity:
          RT:132:2654211
          VNID:2654211
```

```
bdsol-aci32-leaf5#  show ip bgp 172.16.99.0 vrf RD-BGP:RD
BGP routing table information for VRF RD-BGP:RD, address family
IPv4 Unicast
BGP routing table entry for 172.16.99.0/24, version 26 dest ptr
0xab5bd250
Paths: (2 available, best #2)
Flags: (0x80c001a 00000000) on xmit-list, is in urib, is best urib
route, exported
  vpn: version 513, (0x100002) on xmit-list
Multipath: eBGP iBGP

  VPN AF advertised path-id 2
  Path type: internal 0xc0000018 0x40 ref 1, path is valid, not
best reason: Internal path
             Imported from 10.0.88.95:18:172.16.99.0/24
  AS-Path: 200 , path sourced external to AS
    10.0.88.95 (metric 64) from 10.1.224.95 (10.10.32.215)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Received path-id 2
      Extcommunity:
          RT:132:2654211
          COST:pre-bestpath:165:2415919104
          VNID:2654211
      Originator: 10.0.88.95 Cluster list: 10.10.32.215
10.10.32.201

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: external 0x28 0x0 ref 0, path is valid, is best path
  AS-Path: 200 , path sourced external to AS
    172.16.1.18 (metric 0) from 172.16.1.18 (172.16.1.25)
      Origin IGP, MED not set, localpref 100, weight 0
      Extcommunity:
          RT:132:2654211
          VNID:2654211
```

Show bgp vpnv4 unicast 172.16.99.0 vrf XXx

Leaf 1 – we prefer external path

```
VPN AF advertised path-id 2
  Path type: internal 0xc0000018 0x40 ref 1, path is valid, not best reason: Internal
path
              Imported from 10.1.224.94:16:172.16.99.0/24
  AS-Path: 200 , path sourced external to AS
    10.1.224.94 (metric 64) from 10.0.88.93 (10.10.32.214)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Received path-id 1
      Extcommunity:
          RT:132:2654211
          COST:pre-bestpath:165:2415919104
          VNID:2654211
      Originator: 10.1.224.94 Cluster list: 10.10.32.214 10.10.32.202

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: external 0x28 0x0 ref 0, path is valid, is best path
  AS-Path: 200 , path sourced external to AS
    172.16.1.14 (metric 0) from 172.16.1.14 (172.16.1.21)
      Origin IGP, MED not set, localpref 100, weight 0
      Extcommunity:
          RT:132:2654211
          VNID:2654211
```

Leaf 3 we prefer path from leaf1 as leaf5 has worst next-hop metric (isis metric)

```
VPN AF advertised path-id 2
  Path type: internal 0xc0000018 0x40 ref 1, path is valid, not best reason: NH
metric
              Imported from 10.1.224.94:16:172.16.99.0/24
  AS-Path: 200 , path sourced external to AS
    10.1.224.94 (metric 64) from 10.0.88.93 (10.10.32.214)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Received path-id 1
      Extcommunity:
          RT:132:2654211
          COST:pre-bestpath:165:2415919104
          VNID:2654211
      Originator: 10.1.224.94 Cluster list: 10.10.32.214 10.10.32.202

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: internal 0xc0000018 0x40 ref 1, path is valid, is best path
              Imported from 10.0.88.95:18:172.16.99.0/24
  AS-Path: 200 , path sourced external to AS
    10.0.88.95 (metric 3) from 10.0.88.93 (10.10.32.214)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Received path-id 1
      Extcommunity:
          RT:132:2654211
          VNID:2654211
      Originator: 10.0.88.95 Cluster list: 10.10.32.214
```

Leaf 5 – we prefer external path

```
VPN AF advertised path-id 2
  Path type: internal 0xc0000018 0x40 ref 1, path is valid, not best reason: Internal
path
              Imported from 10.0.88.95:18:172.16.99.0/24
  AS-Path: 200 , path sourced external to AS
    10.0.88.95 (metric 64) from 10.1.224.95 (10.10.32.215)
      Origin IGP, MED not set, localpref 100, weight 0
      Received label 0
      Received path-id 2
      Extcommunity:
          RT:132:2654211
          COST:pre-bestpath:165:2415919104
          VNID:2654211
      Originator: 10.0.88.95 Cluster list: 10.10.32.215 10.10.32.201

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: external 0x28 0x0 ref 0, path is valid, is best path
  AS-Path: 200 , path sourced external to AS
    172.16.1.18 (metric 0) from 172.16.1.18 (172.16.1.25)
      Origin IGP, MED not set, localpref 100, weight 0
      Extcommunity:
          RT:132:2654211
          VNID:2654211
```

Leaf 6 we prefer path from leaf5 as leaf1 path has worst next-hop metric (ISIS metric)

```
Advertised path-id 1, VPN AF advertised path-id 1
 Path type: internal 0xc0000018 0x40 ref 1, path is valid, is best path
              Imported from 10.1.224.94:16:172.16.99.0/24
 AS-Path: 200 , path sourced external to AS
   10.1.224.94 (metric 3) from 10.1.224.95 (10.10.32.215)
     Origin IGP, MED not set, localpref 100, weight 0
     Received label 0
     Received path-id 1
     Extcommunity:
         RT:132:2654211
         VNID:2654211
     Originator: 10.1.224.94 Cluster list: 10.10.32.215

 VPN AF advertised path-id 2
 Path type: internal 0xc0000018 0x40 ref 1, path is valid, not best reason: NH metric
              Imported from 10.0.88.95:18:172.16.99.0/24
 AS-Path: 200 , path sourced external to AS
   10.0.88.95 (metric 64) from 10.1.224.95 (10.10.32.215)
     Origin IGP, MED not set, localpref 100, weight 0
     Received label 0
     Received path-id 2
     Extcommunity:
         RT:132:2654211
         COST:pre-bestpath:165:2415919104
         VNID:2654211
     Originator: 10.0.88.95 Cluster list: 10.10.32.215 10.10.32.201
```

BGP path selection

1. Prefer the path with the highest **WEIGHT**. (local to router – cisco specific)
2. Prefer the path with the highest **LOCAL_PREFERENCE**. (value of 100 by default if no Local pref set)
3. Prefer the path that was **locally** originated via a network or aggregate BGP subcommand or through redistribution from an IGP.
4. Prefer the path with the shortest **AS_PATH**.
5. Prefer the path with the lowest **origin type**. (IGP is lower than Exterior Gateway Protocol (EGP), and EGP is lower than INCOMPLETE)
6. Prefer the path with the lowest multi-exit discriminator – MED . (Paths received with no MED are assigned a MED of 0
7. Prefer **eBGP over iBGP** paths.
8. Prefer the path with the lowest **IGP metric** to the BGP next hop. Continue, even if bestpath is already selected.
9. Determine if multiple paths require installation in the routing table for **BGP Multipath**. Continue, if bestpath is not yet selected.
10. When both paths are external, prefer the path that was received first (**the oldest one**).
11. Prefer the route that comes from the BGP router with the lowest **router ID**.
12. Prefer the path that comes from the lowest neighbor address. (This address is the IP address that is used in the BGP **neighbor** configuration)

Criteria used by leaf 1 and 5

Criteria used by leaf 3 and 6

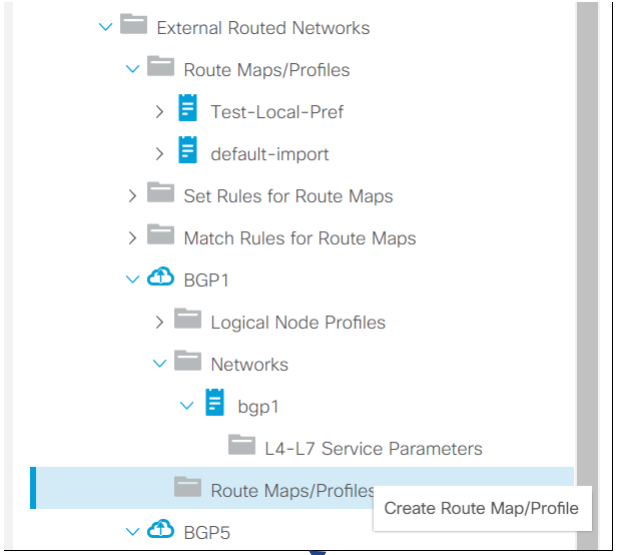
© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

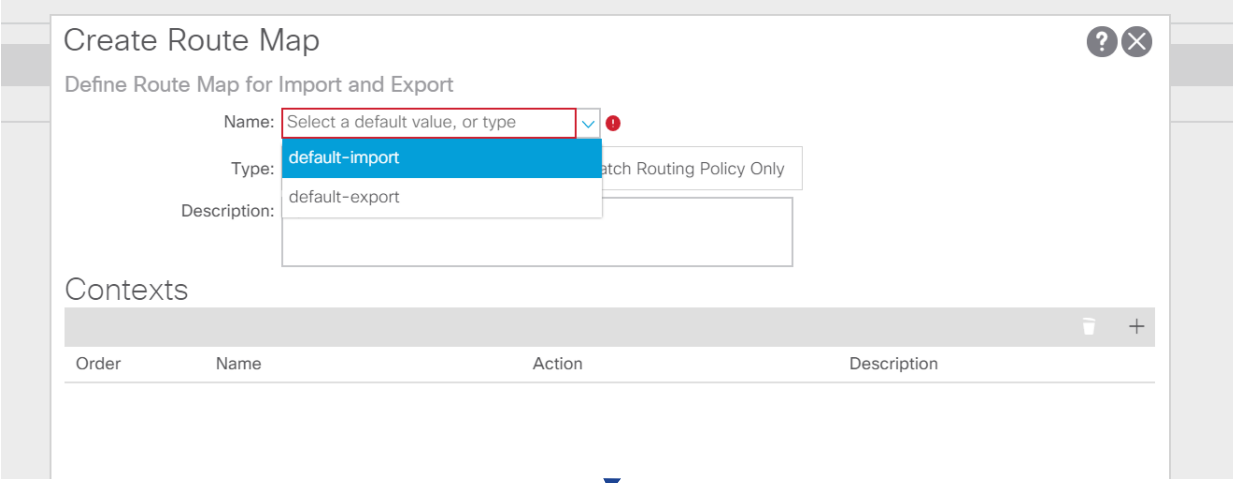
Changing best path

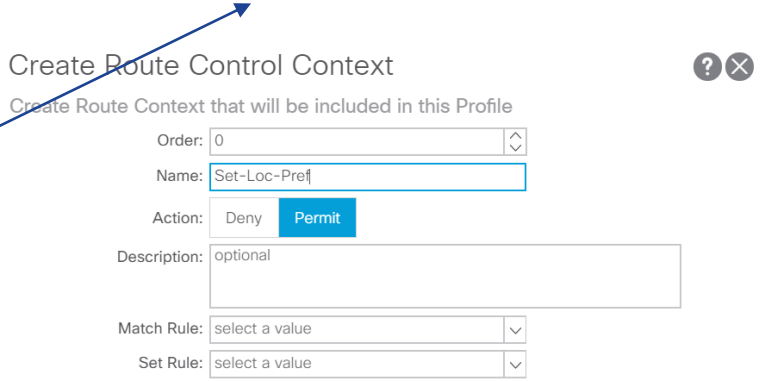
- To be sure the selected best path is the one receive from leaf 1 on every leaf, we will add a local preference (highest than 100) on Leaf 1 layer 3 out
- This is criteria 2 in BGP path selection and will come before eBGP<>iBGP or metric to NH !

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Config







Step 1. in Ext Routed Net, right click on Route-map

Step 2. Select default-import and click the + to create a new Context

Step 3. Specify the order (route-map seq can be 0) and a name
Leave the action for the sequence as permit.
Then you will need to select (or create a new) match rule and set rule (next slides)

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Config Match Rule

Step 4 – configure the match rule
Match can be done on Community
Or on prefix
Here we match on prefix 172.16.99.0

Match Rule - 99subnet

Policy History

Properties

Name: 99subnet

Description: optional

Match Regex Community Terms:

Name	Regular Expression	Community Type	Description

No items have been found.
Select Actions to create a new item.

Match Community Terms:

Name	Description

No items have been found.
Select Actions to create a new item.

Match Prefix:

IP	Aggregate	Description
172.16.99.0/24	True	

Config set Rule

Action Rule Profile - SetLocalPrefHigh

Policy Histor

Properties

Rule Name: SetLocalPrefHigh

Description: optional

Set Communities:

Set Route Tag:

Set Dampening:

Set Weight:

Set Next Hop:

Set Preference: Preference: 150

Set Metric:

Set Metric Type:

Additional Communities:

Community	Set Criteria	Description

No items have been found.
Select Actions to create a new item

Step 5 –in Set Rule, we can set many parameter, here we set Local Pref to 150

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Create Route Control Context

Create Route Context that will be included in this Profile

Order: 0

Name: Set-Loc-Pref

Action: Deny | Permit

Description: optional

Match Rule: 99subnet

Set Rule: SetLocalPrefHigh

Create Route Map

Define Route Map for Import and Export

Name: default-import

Type: Match Prefix AND Routing Policy | Match Routing Policy Only

Description: optional

Contexts

Order	Name	Action	Description
0	Set-Loc-Pref	Permit	

```
bdsol-aci32-leaf1# show ip bgp neighbors 172.16.1.14 vrf RD-BGP:RD | egrep route-map
  Inbound route-map configured is imp-l3out-BGP1-peer-2654211, handle obtained
  Outbound route-map configured is exp-l3out-BGP1-peer-2654211, handle obtained
bdsol-aci32-leaf1# show route-map  imp-l3out-BGP1-peer-2654211
route-map imp-l3out-BGP1-peer-2654211, permit, sequence 8201
  Match clauses:
    ip address prefix-lists: IPv4-peer16387-2654211-exc-ext-in-default-import2Set-Loc-Pref099subnet-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:
    local-preference 150
bdsol-aci32-leaf1# show ip prefix-list  IPv4-peer16387-2654211-exc-ext-in-default-import2Set-Loc-Pref099subnet-dst
ip prefix-list IPv4-peer16387-2654211-exc-ext-in-default-import2Set-Loc-Pref099subnet-dst: 1 entries
   seq 1 permit 172.16.99.0/24 le 32
```

Note that le 32 in prefix-list, this is due to the aggregate flag we set in match rule

ates. All rights reserved. Cisco Confidential

Leaf 1 – ip bgp path – Local pref is set

```
bdsol-aci32-leaf1#  show ip bgp  172.16.99.0 vrf RD-BGP:RD
BGP routing table information for VRF RD-BGP:RD, address family IPv4 Unicast
BGP routing table entry for 172.16.99.0/24, version 57 dest ptr 0xab12a650
Paths: (1 available, best #1)
Flags: (0x80c001a 00000000) on xmit-list, is in urib, is best urib route, exported
  vpn: version 1396, (0x100002) on xmit-list
Multipath: eBGP iBGP

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: external 0x28 0x0 ref 0, path is valid, is best path
  AS-Path: 200 , path sourced external to AS
    172.16.1.14 (metric 0) from 172.16.1.14 (172.16.1.21)
      Origin IGP, MED not set, localpref 150, weight 0
      Extcommunity:
          RT:132:2654211
          VNID:2654211

  VRF advertise information:
  Path-id 1 not advertised to any peer

  VPN AF advertise information:
  Path-id 1 advertised to peers:
    10.0.88.93         10.0.88.94
```

Show bgp vpnv4 unicast 172.16.99.0 vrf XXx

Leaf 1 – we only see the locally rx path with local pref 150

```
bdsol-aci32-leaf1#  show  bgp vpnv4 unicast  172.16.99.0 vrf RD-BGP:RD
BGP routing table information for VRF overlay-1, address family VPNv4 Unicast
Route Distinguisher: 10.0.88.95:18    (VRF RD-BGP:RD)
BGP routing table entry for 172.16.99.0/24, version 57 dest ptr 0xab12a650
Paths: (1 available, best #1)
Flags: (0x80c001a 00000000) on xmit-list, is in urib, is best urib route, exported
  vpn: version 1396, (0x100002) on xmit-list
Multipath: eBGP iBGP

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: external 0x28 0x0 ref 0, path is valid, is best path
  AS-Path: 200 , path sourced external to AS
    172.16.1.14 (metric 0) from 172.16.1.14 (172.16.1.21)
      Origin IGP, MED not set, localpref 150, weight 0
      Extcommunity:
          RT:132:2654211
          VNID:2654211
```

Leaf 5 – we have both the local ebgp path which is not send to RR as we have best path with higher local pref

```
BGP routing table information for VRF overlay-1, address family VPNv4 Unicast
Route Distinguisher: 10.1.224.94:16    (VRF RD-BGP:RD)
BGP routing table entry for 172.16.99.0/24, version 29 dest ptr 0xab5bd250
Paths: (2 available, best #1)
Flags: (0x808001a 00000000) on xmit-list, is in urib, is best urib route
  vpn: version 519, (0x100002) on xmit-list
Multipath: eBGP iBGP

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: internal 0xc0000018 0x40 ref 1, path is valid, is best path
             Imported from 10.0.88.95:18:172.16.99.0/24
  AS-Path: 200 , path sourced external to AS
    10.0.88.95 (metric 64) from 10.1.224.95 (10.10.32.215)
      Origin IGP, MED not set, localpref 150, weight 0
      Received label 0
      Received path-id 2
      Extcommunity:
          RT:132:2654211
          COST:pre-bestpath:165:2415919104
          VNID:2654211
      Originator: 10.0.88.95 Cluster list: 10.10.32.215 10.10.32.201

  VPN AF advertised path-id 2
  Path type: external 0x28 0x0 ref 0, path is valid, not best reason: Local Preference
  AS-Path: 200 , path sourced external to AS
    172.16.1.18 (metric 0) from 172.16.1.18 (172.16.1.25)
      Origin IGP, MED not set, localpref 100, weight 0
      Extcommunity:
          RT:132:2654211
```

Leaf 3 and leaf 6 only have the path with pref 150 from Leaf 1

```
bdsol-aci32-leaf3# show  bgp vpnv4 unicast  172.16.99.0 vrf RD-BGP:RD
BGP routing table information for VRF overlay-1, address family VPNv4 Unicast
Route Distinguisher: 10.0.88.91:15    (VRF RD-BGP:RD)
BGP routing table entry for 172.16.99.0/24, version 47 dest ptr 0xaaa4a450
Paths: (1 available, best #1)
Flags: (0x08001a 00000000) on xmit-list, is in urib, is best urib route
  vpn: version 748, (0x100002) on xmit-list
Multipath: eBGP iBGP

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: internal 0xc0000018 0x40 ref 1, path is valid, is best path
             Imported from 10.0.88.95:18:172.16.99.0/24
  AS-Path: 200 , path sourced external to AS
    10.0.88.95 (metric 3) from 10.0.88.93 (10.10.32.214)
      Origin IGP, MED not set, localpref 150, weight 0
      Received label 0
      Received path-id 1
      Extcommunity:
          RT:132:2654211
          VNID:2654211
      Originator: 10.0.88.95 Cluster list: 10.10.32.214
```

```
bdsol-aci32-leaf6# show  bgp vpnv4 unicast  172.16.99.0 vrf RD-BGP:RD
BGP routing table information for VRF overlay-1, address family VPNv4 Unicast
Route Distinguisher: 10.1.168.95:15    (VRF RD-BGP:RD)
BGP routing table entry for 172.16.99.0/24, version 35 dest ptr 0xab266450
Paths: (1 available, best #1)
Flags: (0x08001a 00000000) on xmit-list, is in urib, is best urib route
  vpn: version 538, (0x100002) on xmit-list
Multipath: eBGP iBGP

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: internal 0xc0000018 0x40 ref 1, path is valid, is best path
             Imported from 10.0.88.95:18:172.16.99.0/24
  AS-Path: 200 , path sourced external to AS
    10.0.88.95 (metric 64) from 10.1.224.95 (10.10.32.215)
      Origin IGP, MED not set, localpref 150, weight 0
      Received label 0
      Received path-id 2
      Extcommunity:
          RT:132:2654211
          COST:pre-bestpath:165:2415919104
          VNID:2654211
      Originator: 10.0.88.95 Cluster list: 10.10.32.215 10.10.32.201
```

Routing Table with Local Pref set

Leaf 1 prefer the External BGP route with admin distance 20

```
bdsol-aci32-leaf1#  show ip route 172.16.99.0 vrf RD-BGP:RD
172.16.99.0/24, ubest/mbest: 1/0
    *via 172.16.1.14, [20/0], 00:14:03, bgp-132, external, tag
200
```

Leaf 5 only have the iBGP route to Leaf1 (admin dist 200) And RIB do not have the eBPG route as it was rejected by BGP best path selection

```
bdsol-aci32-leaf5#  show ip route 172.16.99.0 vrf RD-BGP:RD

172.16.99.0/24, ubest/mbest: 1/0
    *via 10.0.88.95%overlay-1, [200/0], 00:38:11, bgp-132,
internal, tag 200 (mpls-vpn)
```

Leaf 3 selects ibgp route to leaf1

```
bdsol-aci32-leaf3# show ip route 172.16.99.0 vrf RD-BGP:RD

172.16.99.0/24, ubest/mbest: 1/0
    *via 10.0.88.95%overlay-1, [200/0], 00:38:55, bgp-132,
internal, tag 200 (mpls-vpn)
```

Leaf 6 selects ibgp route to leaf1

```
bdsol-aci32-leaf6#  show ip route 172.16.99.0 vrf RD-BGP:RD

172.16.99.0/24, ubest/mbest: 1/0
    *via 10.0.88.95%overlay-1, [200/0], 00:37:49, bgp-132,
internal, tag 200 (mpls-vpn)
```

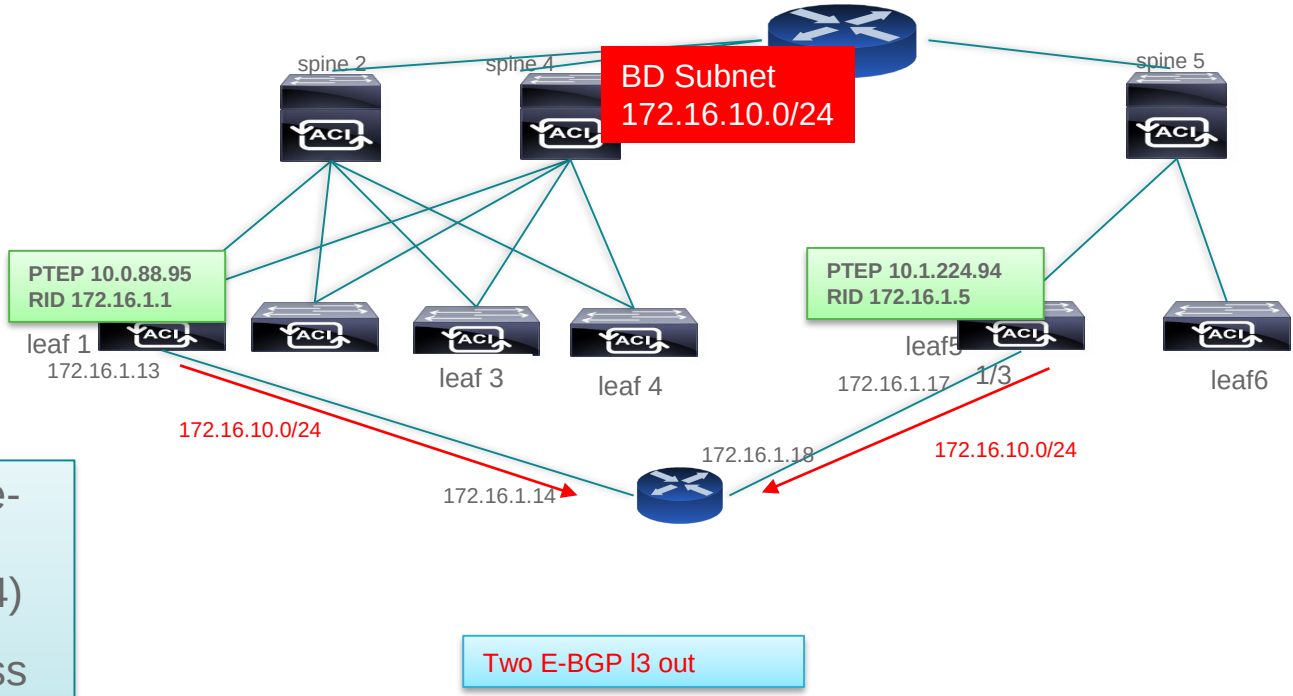
© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Selecting preferred ingress Border leaf

Export Route Control for BGP

© 2015 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

BGP route control



- Fabric have two l3 out using e-bgp and send the same BD prefix to both (172.16.99.0/24)
- Can we force the traffic ingress of ACI to this prefix to always ingress on leaf 5 and only use leaf 1 if we lose primary entry point ?

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

BGP path by default on external router

```
bdsol-aci32-n3k-1# show ip route vrf RD-BGP 172.16.10.0
IP Route Table for VRF "RD-BGP"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.10.0/24, ubest/mbest: 1/0
    *via 172.16.1.13, [20/0], 00:02:04, bgp-200, external, tag 132


bdsol-aci32-n3k-1# show ip bgp  172.16.10.0  vrf RD-BGP
BGP routing table information for VRF RD-BGP, address family IPv4 Unicast
BGP routing table entry for 172.16.10.0/24, version 30
Paths: (2 available, best #2)
Flags: (0x800001a) on xmit-list, is in urib, is best urib route, is in HW

  Path type: external, path is valid, not best reason: newer EBGP path, no labeled nexthop
  AS-Path: 132 , path sourced external to AS
    172.16.1.17 (metric 0) from 172.16.1.17 (172.16.1.5)
      Origin incomplete, MED 0, localpref 100, weight 0

  Advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, in rib
  AS-Path: 132 , path sourced external to AS
    172.16.1.13 (metric 0) from 172.16.1.13 (172.16.1.1)
      Origin incomplete, MED 0, localpref 100, weight 0
```

Here we select newest eBGP path → underministic,

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

BGP path selection

1. Prefer the path with the highest **WEIGHT**. (local to router – cisco specific)
2. Prefer the path with the highest **LOCAL_PREFERENCE**. (value of 100 by default if no Local pref set. Local to AS)
3. Prefer the path that was **locally** originated via a network or aggregate BGP subcommand or through redistribution from an IGP.
4. Prefer the path with the shortest **AS_PATH**.
5. Prefer the path with the lowest **origin type**. (IGP is lower than Exterior Gateway Protocol (EGP), and EGP is lower than INCOMPLETE)
6. Prefer the path with the lowest multi-exit discriminator – MED . (Paths received with no MED are assigned a MED of 0
7. Prefer **eBGP over iBGP** paths.
8. Prefer the path with the lowest **IGP metric** to the BGP next hop. Continue, even if bestpath is already selected.
9. Determine if multiple paths require installation in the routing table for **BGP Multipath.** Continue, if bestpath is not yet selected.
10. When both paths are external, prefer the path that was received first (**the oldest one**).
11. Prefer the route that comes from the BGP router with the lowest **router ID**.
12. Prefer the path that comes from the lowest neighbor address. (This address is the IP address that is used in the BGP **neighbor** configuration)

Criteria we would use in larger bgp network, ext router
Would pick up one or the other depending on igp metric

Criteria used here

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

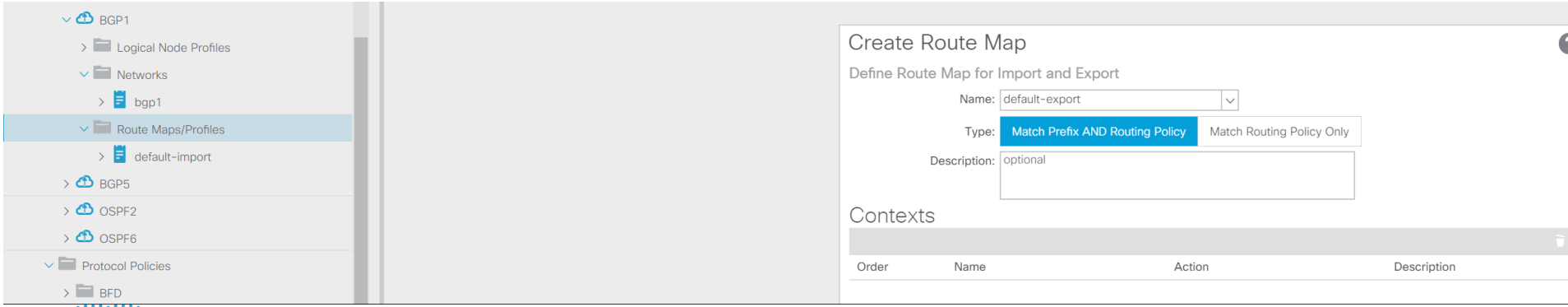
BGP influencing outbound path

- Here we can set MED out on prefix send from leaf 1. MED will be set to 50. The default being 0 we will prefer path from Leaf 5 (lower MED)
- This can be done in 2 way : using BD config (L3 out for Route Profile and route profile under L3 config of BD) . Currently broken in 3.0 due to CSCvg23213
- Or by applying the default route-export route-map under the L3 out for BGP1 (leaf 1 BGP) . We will do that here.
- Note : we could as well have made the as-path longer

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

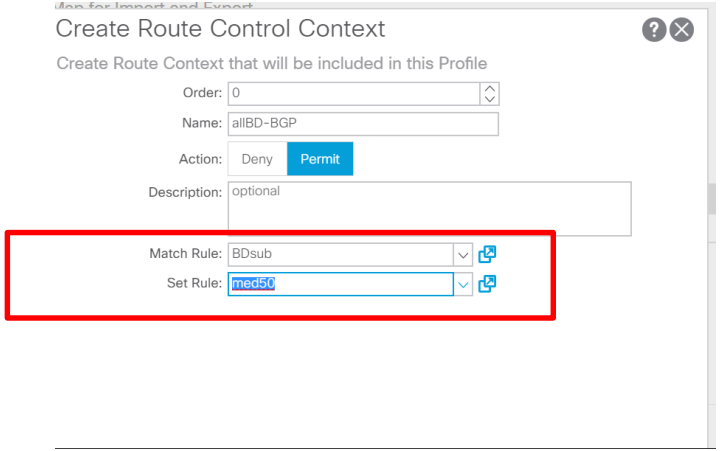
Step 1 – Configure default route-export on L3 out BGP 1

Right click on Route Maps/Profile under the L3 out we want to tune
And create new route-map, default-export will be in the drop down by default.
Then add a Context to it



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Step 2 – Create the context to set MED to 50



© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Route-map on BL before and after

```
bdsol-aci32-leaf1# show ip bgp neighbors 172.16.1.14  vrf RD-BGP:RD | egrep route
  BGP version 4, remote router ID 172.16.1.25
  Inbound route-map configured is imp-l3out-BGP1-peer-2654211, handle obtained
  Outbound route-map configured is exp-l3out-BGP1-peer-2654211, handle obtained
```

```
bdsol-aci32-leaf1# show route-map exp-l3out-BGP1-peer-2654211
route-map exp-l3out-BGP1-peer-2654211, permit, sequence 15801
  Match clauses:
    ip address prefix-lists: IPv4-peer16387-2654211-exc-int-inferred-export-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:
route-map exp-l3out-BGP1-peer-2654211, deny, sequence 16000
  Match clauses:
    route-type: direct
  Set clauses:
```

Before

```
bdsol-aci32-leaf1# show route-map exp-l3out-BGP1-peer-2654211
route-map exp-l3out-BGP1-peer-2654211, permit, sequence 8201
  Match clauses:
    ip address prefix-lists: IPv4-peer16387-2654211-exc-ext-out-default-export2allBD-BGP0BDsub-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:
    tag 4294967295
    metric 50
    community none
    extcommunity 4byteas-generic none
route-map exp-l3out-BGP1-peer-2654211, permit, sequence 8202
  Match clauses:
    ip address prefix-lists: IPv4-peer16387-2654211-exc-int-out-default-export2allBD-BGP0BDsub-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:
    metric 50
    community none
    extcommunity 4byteas-generic none
route-map exp-l3out-BGP1-peer-2654211, deny, sequence 16000
  Match clauses:
    route-type: direct
  Set clauses:
bdsol-aci32-leaf1# show ip prefix-list IPv4-peer16387-2654211-exc-ext-out-default-export2allBD-BGP0BDsub-dst
ip prefix-list IPv4-peer16387-2654211-exc-ext-out-default-export2allBD-BGP0BDsub-dst: 1 entries
   seq 1 permit 172.16.10.0/24 le 32
bdsol-aci32-leaf1#
```

After

Result in external router

```
bdsol-aci32-n3k-1# show ip bgp 172.16.10.0 vrf RD-BGP
BGP routing table information for VRF RD-BGP, address family IPv4 Unicast
BGP routing table entry for 172.16.10.0/24, version 45
Paths: (2 available, best #2)
Flags: (0x800001a) on xmit-list, is in urib, is best urib route, is in HW

  Path type: external, path is valid, not best reason: MED, no labeled nexthop
  AS-Path: 132 , path sourced external to AS
    172.16.1.13 (metric 0) from 172.16.1.13 (172.16.1.1)
      Origin incomplete, MED 50, localpref 100, weight 0

  Advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, in rib
  AS-Path: 132 , path sourced external to AS
    172.16.1.17 (metric 0) from 172.16.1.17 (172.16.1.5)
      Origin incomplete, MED 0, localpref 100, weight 0

bdsol-aci32-n3k-1# show ip route 172.16.10.0 vrf RD-BGP

172.16.10.0/24, ubest/mbest: 1/0
    *via 172.16.1.17, [20/0], 17:35:31, bgp-200, external, tag 132
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential