Troubleshoot ACI L3Out - Directly-Connected Subnet PcTag1

Contents

Introduction

This document describes a scenario where traffic sourced from a directly-connected L3Out subnet without the proper configuration under the external EPG can lead to contract drops.

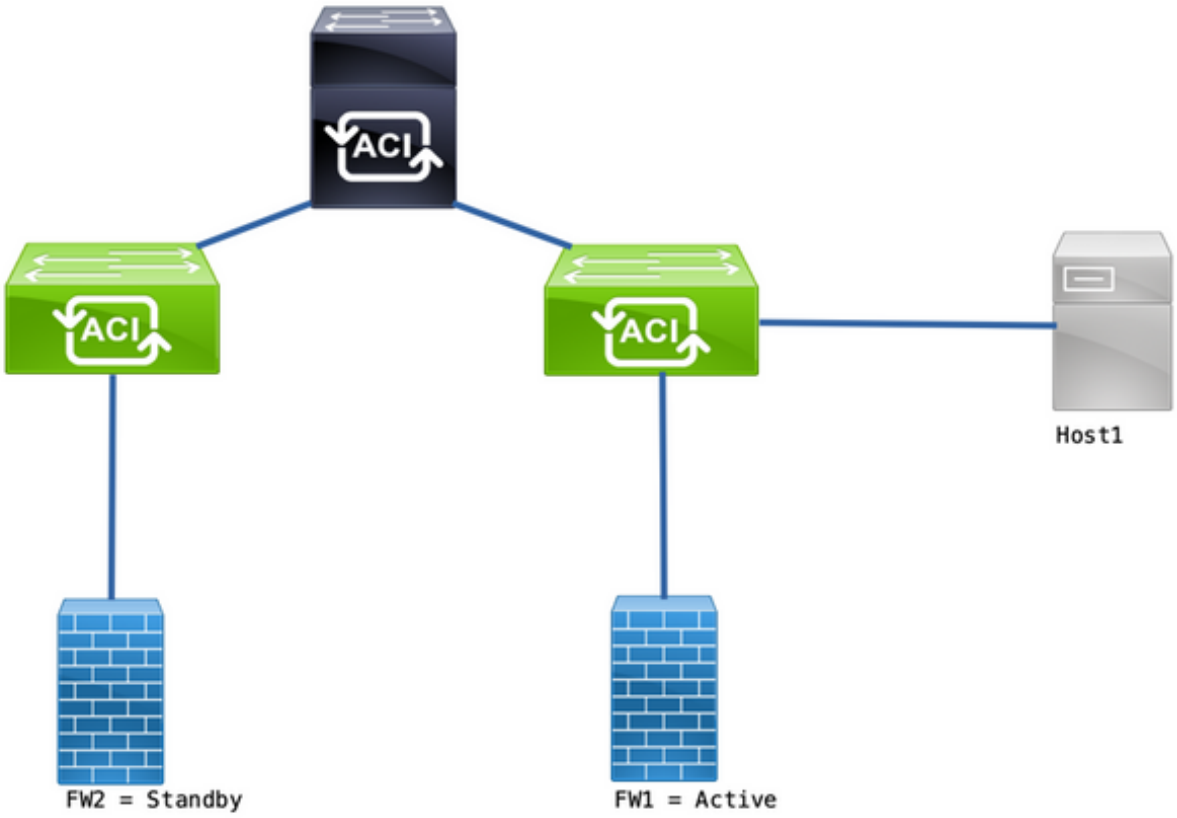
Background Information

The "**An exception for a directly connected subnet with 0.0.0.0/0**" section from the ACI L3out Whitepaper calls out this behavior with regards to pcTag 1:

"...by default, directly connected subnets are assigned pcTag 1, which is a special pcTag to bypass a contract. This is to implicitly allow route protocol communications in a corner case scenario. However... this can cause a security concern instead. Hence, this behavior is explained in detail via Cisco bug ID CSCuz12913 , which also introduces a workaround configuration:"

The Scenario

Topology and Configuration



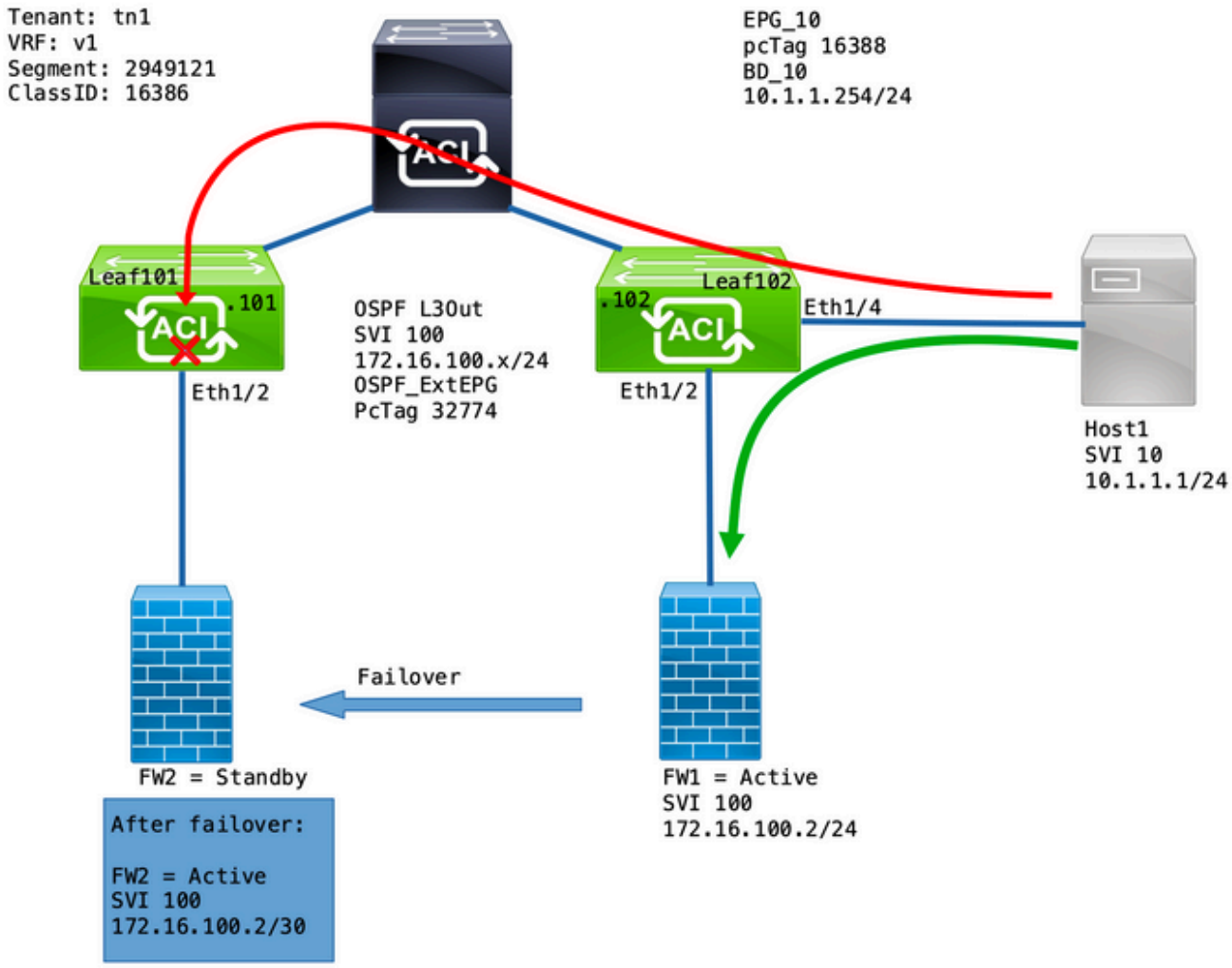
Topology

- The Firewalls (FW) are configured with Network Address Translation (NAT).
- All traffic sent into the ACI fabric is sourced from the IP of the FW that forms the OSPF adjacency with ACI.
- The external EPG has a 0.0.0.0/0 network configured with **External Subnets for the External EPG**.
- A contract is in place for communication between the internal EPG and the external EPG.

Observed Issue

With FW1 as the active device, traffic works as expected. There are no drops observed.

After the firewall services fail over to FW2, the connectivity is lost - 10.1.1.1 and 172.16.100.2 are no longer able to communicate.



Issue Deep-Dive

An ELAM capture on Leaf101 allows us to validate if the traffic from Host1 to FW2 is dropped.

These ELAM options were used:

```
leaf101# vsh_lc
module-1# debug platform internal roc elam asic 0
module-1(DBG-elam-insel6)# trigger reset
module-1(DBG-elam)# trigger init in-select 14 out-select 1
module-1(DBG-elam-insel14)# set inner ipv4 src_ip 10.1.1.1 dst_ip 172.16.100.2
module-1(DBG-elam-insel14)# start
module-1(DBG-elam-insel14)# status
```

And when triggered, the e-report allows you to view the lookup results:

```
<snip>
================================================================================================
================================================
Captured Packet
================================================================================================
================================================
<snip>
------------------------------------------------------------------------------------------------
```

```
-----------------------------------------------------
Inner L3 Header
-----------------------------------------------------------------------------------------------
-----------------------------------------------------
L3 Type : IPv4
DSCP : 0
Don't Fragment Bit : 0x0
TTL : 254
IP Protocol Number : ICMP
Destination IP                 : 172.16.100.2   <<<----
Source IP                      : 10.1.1.1   <<<----
<snip>
===============================================================================================
=====================================================
Contract Lookup ( FPC )
===============================================================================================
=====================================================

-----------------------------------------------------------------------------------------------
-----------------------------------------------------
Contract Lookup Key
-----------------------------------------------------------------------------------------------
-----------------------------------------------------
IP Protocol : ICMP( 0x1 )
L4 Src Port : 2048( 0x800 )
L4 Dst Port : 52579( 0xCD63 )
sclass (src pcTag)                      : 16388( 0x4004 )   <<<----
dclass (dst pcTag)                      : 16386( 0x4002 )   <<<----
<snip>

-----------------------------------------------------------------------------------------------
-----------------------------------------------------
Contract Result
-----------------------------------------------------------------------------------------------
-----------------------------------------------------
Contract Drop                          : yes   <<<----
Contract Logging : yes
Contract Applied : no
Contract Hit : yes
Contract Aclqos Stats Index : 81824
( show sys int aclqos zoning-rules | grep -B 9 "Idx: 81824" )
```

This report shows that the flow is Contract Dropped along with these details:

- The SCLASS is 16388 which is the pcTag of EPG_10.
- The DCLASS is 16386 which is the pcTag of the VRF v1.

Next, validate the zoning rules for the VRF:

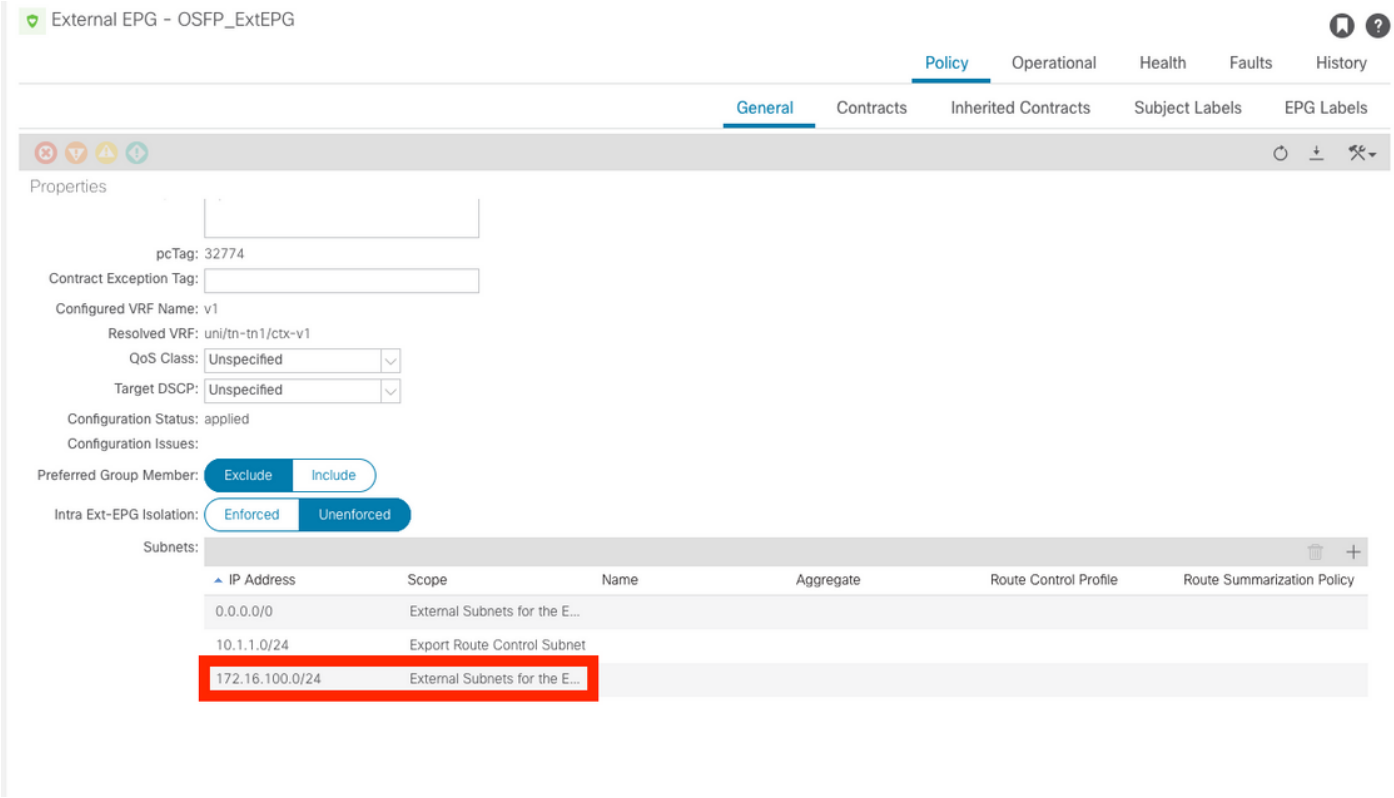
```
leaf102# show zoning-rule scope 2949121
+---------+--------+--------+----------+---------+---------+---------+------------------+-------
---+----------------------+
| Rule ID | SrcEPG | DstEPG | FilterID |   Dir   |  operSt |  Scope  |       Name       |
Action  |       Priority       |
+---------+--------+--------+----------+---------+---------+---------+------------------+-------
---+----------------------+
|   4131  |   0    |   15   | implicit | uni-dir | enabled | 2949121 |                  |
deny,log | any_vrf_any_deny(22) |
|   4130  |   0    |   0    | implarp  | uni-dir | enabled | 2949121 |                  |
permit  |  any_any_filter(17)  |
|   4129  |   0    |   0    | implicit | uni-dir | enabled | 2949121 |                  |
deny,log |   any_any_any(21)    |
|   4132  |   0    | 49155  | implicit | uni-dir | enabled | 2949121 |                  |
```

```
permit  |   any_dest_any(16)   |
|   4112  | 16386  | 16388  | default  | uni-dir | enabled | 2949121 | tn1:EPG-to-L3Out |
permit  |    src_dst_any(9)    |
|   4133  | 16388  |   15   | default  | uni-dir | enabled | 2949121 | tn1:EPG-to-L3Out |
permit  |    src_dst_any(9)    |
+---------+--------+--------+----------+---------+---------+---------+-----------------+-------
---+----------------------+
```

There is a contract in place for communication from EPG_10 (16388) to networks behind the OSPF L3Out (0.0.0.0/0 = 15). However, the traffic from 172.16.100.2 is tagged under the VRF v1's pcTag (16386).

Solution

Add the directly-connected subnet of the L3Out under the OSPF Ext_EPG.



This addition has 2 effects:

1. Traffic from the directly connected subnet is tagged under the OSPF_ExtEPG pcTag (32774)
2. Rules are added to permit the flow to and from EPG_10 and OSPF_ExtEPG

```
leaf102# show zoning-rule scope 2949121
+---------+--------+--------+----------+----------------+---------+---------+-----------------
+----------+----------------------+ | Rule ID | SrcEPG | DstEPG | FilterID | Dir | operSt |
Scope | Name | Action | Priority | +---------+--------+--------+----------+----------------+----
-----+---------+-----------------+----------+----------------------+ | 4131 | 0 | 15 | implicit
| uni-dir | enabled | 2949121 | | deny,log | any_vrf_any_deny(22) | | 4130 | 0 | 0 | implarp |
uni-dir | enabled | 2949121 | | permit | any_any_filter(17) | | 4129 | 0 | 0 | implicit | uni-
dir | enabled | 2949121 | | deny,log | any_any_any(21) | | 4132 | 0 | 49155 | implicit | uni-dir
| enabled | 2949121 | | permit | any_dest_any(16) | | 4112 | 16386 | 16388 | default | uni-dir |
enabled | 2949121 | tn1:EPG-to-L3Out | permit | src_dst_any(9) | | 4133 | 16388 | 15 | default |
uni-dir | enabled | 2949121 | tn1:EPG-to-L3Out | permit | src_dst_any(9) | |   4134  | 16388  |
32774  | default  |     bi-dir     | enabled | 2949121 | tn1:EPG-to-L3Out |  permit  |
src_dst_any(9)    |  <<<----
```

```
|   4135  | 32774  | 16388  | default  | uni-dir-ignore | enabled | 2949121 | tn1:EPG-to-L3Out |
permit  |    src_dst_any(9)    |  <<<----
+---------+--------+--------+----------+----------------+---------+---------+-----------------
+----------+----------------------+
```

Explanation

The reason this works when the FW and Host are connected to the same leaf (without the L3Out subnet addition) is because directly connected subnets use a special pcTag of 1 which bypass all contracts. This is to implicitly allow route protocol communications in a corner case scenario.

With these triggers we can catch a traffic flow from 172.16.100.2 to 10.1.1.1 while on Leaf102:

```
leaf102# vsh_lc
module-1# debug platform internal roc elam asic 0
module-1(DBG-elam)# trigger reset
module-1(DBG-elam)# trigger init in-select 6 out-select 1
module-1(DBG-elam-insel6)# set outer ipv4 src_ip 172.16.100.2 dst_ip 10.1.1.1
module-1(DBG-elam-insel6)# start
module-1(DBG-elam-insel6)# status
 ELAM STATUS
===========
Asic 0 Slice 0 Status Triggered
```

This ereport shows the lookup results:

```
module-1(DBG-elam-insel6)# ereport
Python available. Continue ELAM decode with LC Pkg
 ELAM REPORT
================================================================================================================
=====================================================
                                                                Captured Packet
================================================================================================================
=====================================================

----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------
Outer L3 Header
----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------
L3 Type                       : IPv4
IP Version                    : 4
DSCP                          : 0
IP Packet Length              : 84 ( = IP header(28 bytes) + IP payload )
Don't Fragment Bit            : not set
TTL                           : 255
IP Protocol Number            : ICMP
IP CheckSum                   : 32320( 0x7E40 )
Destination IP                : 10.1.1.1      <<<----
Source IP                     : 172.16.100.2  <<<----

================================================================================================================
=====================================================
                                                                    Contract Lookup ( FPC )
================================================================================================================
```

```
=====================================================

-------------------------------------------------------------------------------------------------
-----------------------------------------------------
Contract Lookup Key
-------------------------------------------------------------------------------------------------
-----------------------------------------------------
IP Protocol                           : ICMP( 0x1 )
L4 Src Port                           : 0( 0x0 )
L4 Dst Port                           : 19821( 0x4D6D )
sclass (src pcTag)                    : 1( 0x1 )          <<<----
dclass (dst pcTag)                    : 16388( 0x4004 )   <<<----
src pcTag is from local table         : yes
derived from a local table on this node by the lookup of src IP or MAC
Unknown Unicast / Flood Packet        : no
If yes, Contract is not applied here because it is flooded

-------------------------------------------------------------------------------------------------
-----------------------------------------------------
Contract Result
-------------------------------------------------------------------------------------------------
-----------------------------------------------------
Contract Drop                         : no  <<<----
Contract Logging                      : no
Contract Applied                      : no  <<<----
Contract Hit                          : yes
Contract Aclqos Stats Index           : 81903
```

To validate the return flow:

```
module-1(DBG-elam-insel6)# trigger reset
module-1(DBG-elam)# trigger init in-select 6 out-select 1
module-1(DBG-elam-insel6)# set outer ipv4 src_ip 10.1.1.1 dst_ip 172.16.100.2
module-1(DBG-elam-insel6)# start
module-1(DBG-elam-insel6)# status
 ELAM STATUS
===========
Asic 0 Slice 0 Status Triggered
```

The lookup results of the return flow:

```
module-1(DBG-elam-insel6)# ereport
Python available. Continue ELAM decode with LC Pkg
 ELAM REPORT
=================================================================================================
=====================================================
                                                                  Captured Packet
=================================================================================================
=====================================================
-------------------------------------------------------------------------------------------------
-----------------------------------------------------
Outer L3 Header
-------------------------------------------------------------------------------------------------
-----------------------------------------------------
L3 Type                     : IPv4
IP Version                  : 4
DSCP                        : 0
IP Packet Length            : 84 ( = IP header(28 bytes) + IP payload )
Don't Fragment Bit          : not set
TTL                         : 255
```

```
IP Protocol Number            : ICMP
IP CheckSum                   : 32198( 0x7DC6 )
Destination IP                : 172.16.100.2  <<<----
Source IP                     : 10.1.1.1      <<<----

================================================================================================
=================================================
                                                                    Contract Lookup ( FPC )
================================================================================================
=================================================

------------------------------------------------------------------------------------------------
-------------------------------------------------
Contract Lookup Key
------------------------------------------------------------------------------------------------
-------------------------------------------------
IP Protocol                           : ICMP( 0x1 )
L4 Src Port                           : 2048( 0x800 )
L4 Dst Port                           : 18134( 0x46D6 )
sclass (src pcTag)                    : 16388( 0x4004 ) <<<----
dclass (dst pcTag)                    : 1( 0x1 )        <<<----
src pcTag is from local table         : yes
derived from a local table on this node by the lookup of src IP or MAC
Unknown Unicast / Flood Packet        : no
If yes, Contract is not applied here because it is flooded

------------------------------------------------------------------------------------------------
-------------------------------------------------
Contract Result
------------------------------------------------------------------------------------------------
-------------------------------------------------
Contract Drop                         : no    <<<----
Contract Logging                      : no
Contract Applied                      : no    <<<----
Contract Hit                          : yes
Contract Aclqos Stats Index           : 81903
```

This table summarizes the expected behavior on Gen2 switches:

Scenario	Directionality	Contract Drop	No Contract Dro
Across the Same Leaf VRF Policy Enforcement: Both	X to L3Out		X
	L3Out to X		X
Across 2 Leaf Nodes VRF Policy Enforcement: Ingress	X to L3Out	X	
	L3Out to X		X
Across 2 Leaf Nodes VRF Policy Enforcement: Egress	X to L3Out		X
	L3Out to X		X