QoS in Cisco ACI within a POD

Contents

Introduction

In Cisco's Application Centric Infrastructure (ACI) we have various options available to classify traffic to be serviced in specific ways within the fabric. These rules are broadly known as Quality of Service (QoS). QoS is achieved mainly by setting certain values on packets at the Ethernet (Layer 2) or IP (internet protocol, Layer 3) header - known as Class of Service (COS) and Differentiated Services Code Point (DSCP) respectively.

ACI also allows the user to honor, ignore or modify these QOS markings on data traffic coming in or exiting the fabric. We will look at these in detail.

For the scope of this document we will limit ourselves to a single Pod setup in an ACI fabric.

Setup and Topology

The tests and captures were done on Generation 2 hardware in 3.2.x version.

For the purpose of this document we will work with the following setup (indicative diagram).
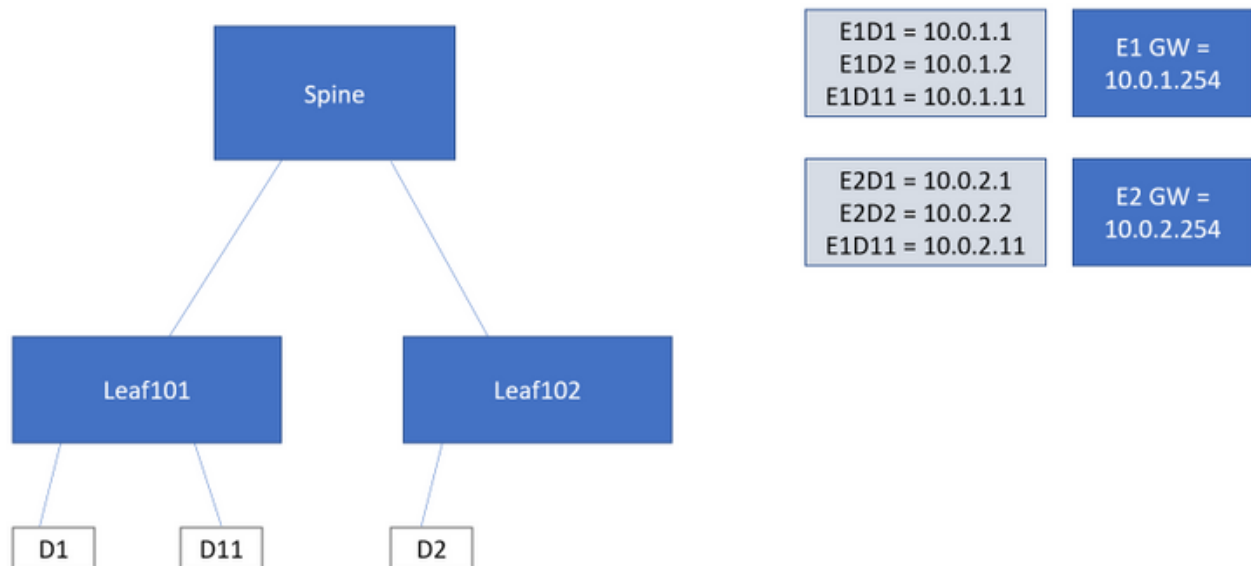
We have a fabric with two End Point Groups (EPG): EPG-1 and EPG-2. Each EPG is linked to its own Bridge Domain (BD).

BD for EPG-1 has subnet 10.0.1.254/24

BD for EPG-2 has subnet 10.0.2.254/24

End points for both EPG's are present on Leaf 1 and 2.

For convenience we will briefly go over the different QOS configurations we will look at in detail:

Scenario 1

In this scenario we will keep the fabric clean of any QOS policies. This is to check the default behavior of the fabric when handling traffic pre-marked with different COS or/and DSCP values.

Scenario 2

In this scenario we will enable 'Dot1p Preserve' option:

Global – QOS Class

Properties

Preserve COS: ☑ Dot1p Preserve

▲ Name	Admin State	Priority Flow Control Admin State	No-Drop-Cos	MTU	Minimum Buffers	Congestion Algorithm
Level1	Enabled	false		9216	0	Tail Drop
Level2	Enabled	false		9216	0	Tail Drop
Level3 (Default)	Enabled	false		9216	0	Tail Drop

We will then repeat some of the traffic streams from Scenario 1 and compare/contrast the fabric's handling of the traffic

Scenario 3

In this scenario we will employ the 'QoS Class' option available in EPG Policy and set it to the different levels available. Then, we repeat the traffic flows and compare the fabric's handling of this traffic.

Scenario 4

This is a repeat of Scenario 3 with the 'Dot1p Preserve' option enabled.

Scenario 5

In this scenario we will define 4 custom QoS policies and then call them on our EPG policy.

Example of one such policy:

Custom QOS Policy - MatchDSCP

Policy | History

Properties

Name: matchDSCP

Description: optional

DSCP to priority map:

Priority	DSCP Range From	DSCP Range To	DSCP Target	▲ Target Cos
Level2	CS1	CS1	CS5	3

Dot1P Classifiers:

Priority	Dot1P Range From	Dot1P Range To	DSCP Target	Target Cos

Show Usage | Close | Submit

These custom QoS policies will help in understanding the different ways of remarking of COS / DSCP on data traffic.

Scenario 1: No QoS policies enabled on ACI

This scenario is to observe default behaviour for traffic pre-marked with some COS or DSCP values.
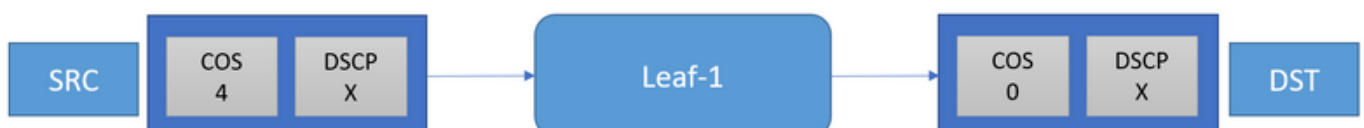
Only two behaviors of concern -

1) Is COS preserved?
2) Is DSCP preserved?

COS is not preserved by default in any condition. The value is lost when the VLAN header is removed at ingress leaf and at egress the cos value is not marked (cos 0 is used)

EXAMPLE 1

Here we send Traffic from E1D1 to E1D11. Traffic at E1D1 is marked with Cos = 4.
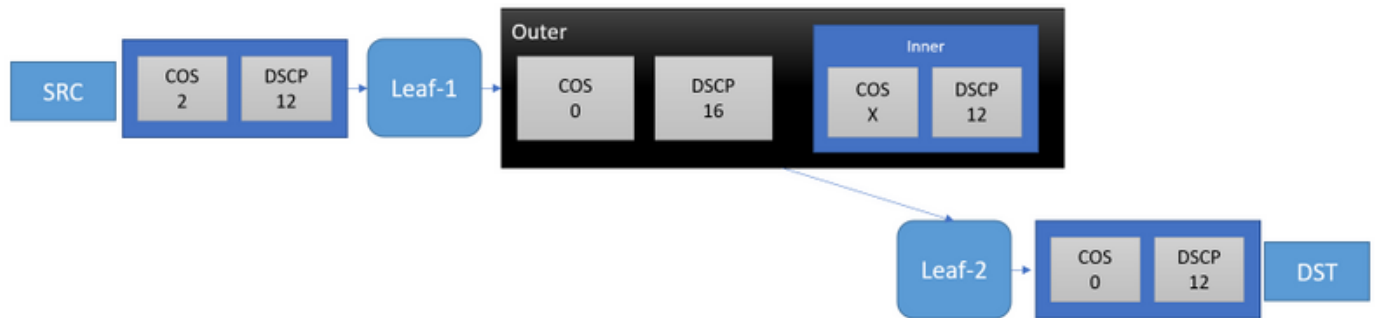

Traffic comes out of Leaf-1 and is received by E1D11, but it has lost its cos marking.

DSCP is preserved by default

EXAMPLE 2

Here we send traffic from E1D1 to E1D2. Traffic at E1D1 is marked with Cos = 2 and DSCP = 12

Traffic exits Leaf 2 with a 0 Cos and same DSCP (12). The outer header has DSCP (16) is explained in the following sections.

Scenario 2: Dot1p Preserve enabled

'Dot1P' is short for 'IEEE 801.1p' - a quality of service prioritisation scheme ; this is part of the the IEEE 802.1Q "Dot1Q" - the networking standard that supports VLAN

Dot1Q header:

TPID : Tag Protocol Identifier- set to a value 0x8100 to identify the frame as an Dot1Q tagged frame

TCI : Tag Control Information , contains the following sub-fields:

PCP : Priority Code Point, a 3 bit field which refers to the Dot1P class of service and maps to the frame priority level

DEI : Drop Eligibility Indicator, a 1 bit field that can be used in conjunction with PCP to indicate frames eligible to be dropped during congestion.

VID : VLAN ID , a 12 bit field specifying the VLAN to which the frame belongs.

By default (with or without 'Dot1p preserve') the COS value on an incoming data packet (ingress into the fabric) is encoded onto the outer header (iVXLAN header) DSCP. 6 bits of DSCP are mapped as follows (pre 4.0):

Significant 3 bits = cos value

Lower 3 bits = class used on traffic (Level 3 by default)

Here is a table with some example DSCP values:

Incoming COS	Level 1	Level 2	Level 3 (default)
0	2	1	0
1	10	9	8
2	18	17	16
3	26	25	24

When 'Dot1p Preserve' is enabled, the outer header DSCP value is decoded to find out the original COS value on data traffic. This is then written to the Dot1P portion of VLAN header on egress from Leaf.

EXAMPLE 3

Here we send traffic from E1D1 to E2D2. Traffic at E1D1 is marked with Cos = 1 and DSCP = 8. With dot1p preserve enabled, both of these values are retained when checked on the destination E2D2.
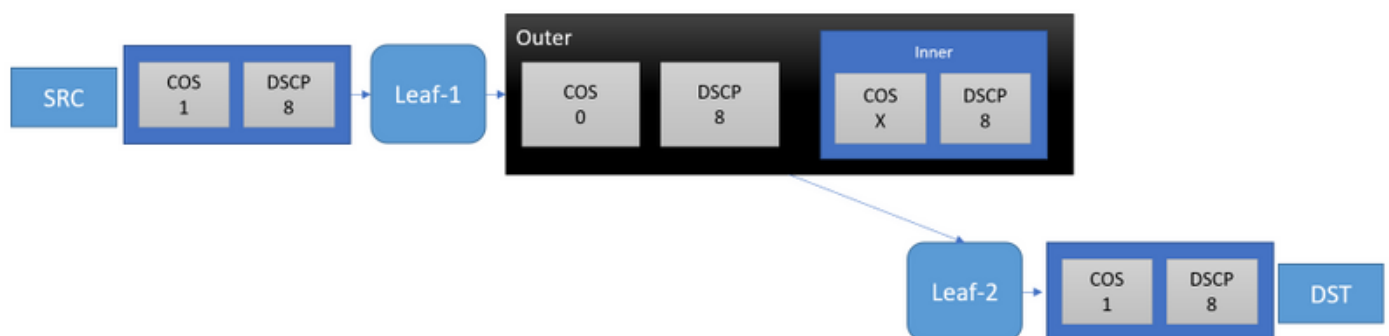
Scenario 3: QoS levels set on EPG

EPG traffic can be marked with certain QOS levels. The default marking is Level 3. Pre 4.0 there were only three user configurable levels - Level 1 to 3. Post 4.0 there are 6 Levels.

The Level is represented on the other header (iVXLAN header) COS as follows:

Pre 4.0:

Level 1	Cos 2
Level 2	Cos 1
Level 3	Cos 0

Post 4.0:

The COS + DEI combinations not mentioned below are reserved for internal use.

Level 1	Cos 2	Dei 0
Level 2	Cos 1	Dei 0
Level 3	Cos 0	Dei 0
Level 4	Cos 2	Dei 1
Level 5	Cos 3	Dei 1
Level 6	Cos 5	Dei 1

Note even though the DEI bit is used , classes 4 , 5 and 6 are not automatically drop eligible duing congestion. The field is just used to because as a convenient way of increasing classes (being adjacent to PCP)

EXAMPLE 4

Here we send traffic from E1D1 to E2D2. Traffic is marked at source with CoS = 1 and DSCP = 8 and EPG-1 is using QOS Class 'Level 1'.

- Level 1 reflects on outer header as CoS 2.

- Since original CoS is 1 and Level is 1 , outer header DSCP is 001010 = 10

- Caveat = if preserve CoS is NOT enabled while using a Level on EPG, the original CoS of data frame is discarded and the one corresponding to the Level is placed on egress frame (this was tested in 3.2.x)
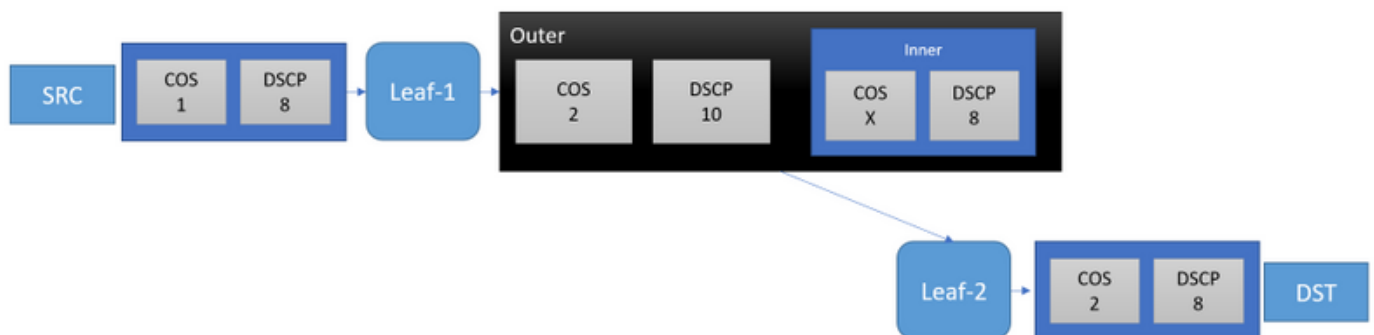
Scenario 4: QoS Class with Dot1P preserve

In this scenario we will also enable Dot1P preserve along with using a QoS Class assignment on EPG-1.

EXAMPLE 5

This will be the same setup as EXAMPLE 4 with the addition of Dot1P Preserve option enabled. With Dot1P Preserve enabled we don't see any unexpected value on egress frame CoS..

Scenario 5: Custom QoS Classes

In this Scenario we will define a custom QoS Class and apply it on our source EPG , EPG-1. If QoS Class and Custom QoS are both used, then Custom QoS takes precedence.

Also within Custom QoS policies, if both "Dot1P Classifiers" and "DSCP to Priority Map" are used, the DSCP Map takes precedence.

The Custom Class will be defined as follows:

- CoS value of 4 should match. If it does, traffic is Classified into Level 2 with CoS of 3 and DSCP CS3 (24)

EXAMPLE 6

Here we will send traffic from E1D1 to E1D2. The traffic is marked at E1D1 with CoS 4 and DSCP 0. The EPG-1 uses the above mentioned Custom QoS policy.

- The class (Level 2) is expressed as CoS 1 in Outer Header

- The rewritten CoS (3) along with the Class is encoded in DSCP = 011001 = 25

Here we observe the same caveat again - without Dot1P Preserve enabled we see the CoS value corresponding to 'Level 2' reflect on the egress data frame. I.E, on E1D2 we will see the frame has CoS 1 and DSCP 24.

The actual expected CoS (3) can be obtained by using Dot1P Preserve: