

# SAI attributes for disabling L3 rewrites for IP Multicast forwarding

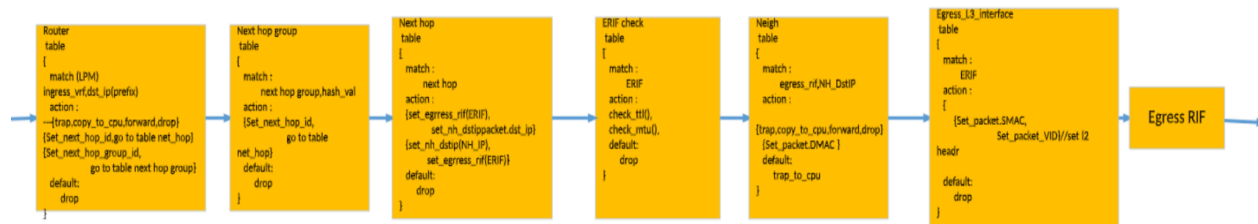
Author: mholankar@

## Overview

This document discusses requirements and the SAI spec proposal for disabling rewriting fields (Source MAC, VLAN) as part of IPMC routing.

## Background

SAI pipeline for IPMC forwarding. After Multicast Replication, the egress pipeline is very similar to Unicast Forwarding. The reference pipeline is from Unicast Forwarding.



- Nexthop sets the egress RIF and NextHop IP
- Neighbor table lookup on NextHop IP to set packet's destination MAC address
- Egress RIF lookup to set packet's source MAC address, VLAN and port.

## Requirements

We require knobs for disabling rewrites to following fields as part of Multicast forwarding flows

- Src MAC disable
- Vlan rewrite disable

## Shared externally

We have scenarios where we need knobs for disabling header field rewrites, for Multicast Replication.

Case 1: For certain flows, the switch does Multicast Replication for a VLAN tagged packet. However, we do not want the VLAN tag to be over-written during replication, we want the replication to retain the VLAN tag. For such a case we would like to disable L3 VLAN rewrite.

Case 2: For some scenarios, we may want to rewrite the SRC MAC. Having Next Hop based knobs for such rewrites would be very helpful.

In summary, with SDN based forwarding, Controller treats the L2 fields like any other header field which can be controlled and requires them to be configured flexibility as part of Multicast Replication.

## Proposal

Since we require capability to disable the rewrites for certain Multicast flows and not for all flows via/to neighbor, the best option is to have these as part of Next Hop object:

- SAI\_NEXT\_HOP\_ATTR\_DISABLE\_SRC\_MAC\_REWRITE
- SAI\_NEXT\_HOP\_ATTR\_DISABLE\_DST\_MAC\_REWRITE
- SAI\_NEXT\_HOP\_ATTR\_DISABLE\_VLAN\_REWRITE

## Example SAI object creation for an IPMC Group:

- Create a **SAI\_OBJECT\_TYPE\_IPMC\_GROUP** for the multicast group with the following attributes.
- Create a **SAI\_OBJECT\_TYPE\_ROUTER\_INTERFACE** with following attributes:
  - SAI\_ROUTER\_INTERFACE\_ATTR\_VIRTUAL\_ROUTER\_ID
  - SAI\_ROUTER\_INTERFACE\_ATTR\_SRC\_MAC\_ADDRESS
  - SAI\_ROUTER\_INTERFACE\_ATTR\_TYPE=SAI\_ROUTER\_INTERFACE\_TYPE\_PORT
  - SAI\_ROUTER\_INTERFACE\_ATTR\_PORT\_ID
- Create a **SAI\_OBJECT\_TYPE\_NEIGHBOR\_ENTRY** for each neighbor with:
  - "ip" = Link local address
  - "rif"
  - "switch id"
  - SAI\_NEIGHBOR\_ENTRY\_ATTR\_DST\_MAC\_ADDRESS (optional)
  - SAI\_NEIGHBOR\_ENTRY\_ATTR\_NO\_HOST\_ROUTE=true
- Create **SAI\_OBJECT\_TYPE\_NEXT\_HOP** with:
  - SAI\_NEXT\_HOP\_ATTR\_TYPE=SAI\_NEXT\_HOP\_TYPE\_IPMC
  - SAI\_NEXT\_HOP\_ATTR\_ROUTER\_INTERFACE\_ID
  - SAI\_NEXT\_HOP\_ATTR\_IP = "ip" of neighbor

## Shared externally

- SAI\_NEXT\_HOP\_ATTR\_DISABLE\_SRC\_MAC\_REWRITE = true
  - SAI\_NEXT\_HOP\_ATTR\_DISABLE\_VLAN\_REWRITE = true
- Create a **SAI\_OBJECT\_TYPE\_IPMC\_GROUP\_MEMBER** with following attributes:
- SAI\_IPMC\_GROUP\_MEMBER\_ATTR\_IPMC\_GROUP\_ID with ipmc\_group\_oid
  - SAI\_IPMC\_GROUP\_MEMBER\_ATTR\_IPMC\_OUTPUT\_ID with rif\_oid
  - SAI\_IPMC\_GROUP\_MEMBER\_ATTR\_IPMC\_NEXT\_HOP with next\_hop\_oid