Configuring Multicast Offload on the Satellite nV System

This chapter describes the configuration of the Satellite Network Virtualization (Satellite nV) multicast offload on the Cisco ASR 9000 Series Aggregation Services Routers.

Need for Multicast Offload

The Satellite nV System architecture currently requires the Cisco ASR 9000 Series Router host to process all replications for supported multicast traffic and topology profiles. This is in line with the envisioned port extender functionality of the satellite devices where all protocol processing and forwarding decisions happen on the host to fully utilize the IOS-XR functionalities. With the introduction of support for a wider variety of satellite topologies, the satellites are no longer restricted by a need for direct connection to the hosts. The satellites can be connected to a Dual Host system to form a Ring topology or through a hub and spoke topology. The satellites can also be reached over an Layer-2 fabric connection through transport provider EVCs.

Note Refer the *Configuring the Satellite Network Virtualization (nV) System* chapter for more information on Simple Ring and Layer 2 fabric topologies.

This use of satellite fabric connections over EVCs or sharing of the same fabric link by a chain of satellite devices has introduced new bandwidth conservation requirements. One such specific case is the forwarding of multicast traffic over satellite ports where the Host currently does all the replication even if all the copies are destined to the ports on the same satellite device. The satellite device is unaware of any multicast group or membership and if multiple receivers are present on the ports of the same satellite, multiple copies of the same packet is sent to the satellite for each of the receivers which can eventually oversubscribe the fabric bandwidth.

With the increasing scale of satellites over the new topologies, it is evident that the model of Host side replication is not very efficient or scalable. The nV Satellite multicast offload feature is introduced to solve this problem. This feature allows the Host to forward just the pre-replicated multicast streams per multicast route (S,G) to the satellites and offload the per satellite access port replication to the satellite device itself. The protocols still run on the Cisco IOS-XR Software modules of the Host but the final replication happens locally on the satellite device based on selective download of routes to the satellites.

Figure 1: Satellite nV System Multicast Offload Illustration
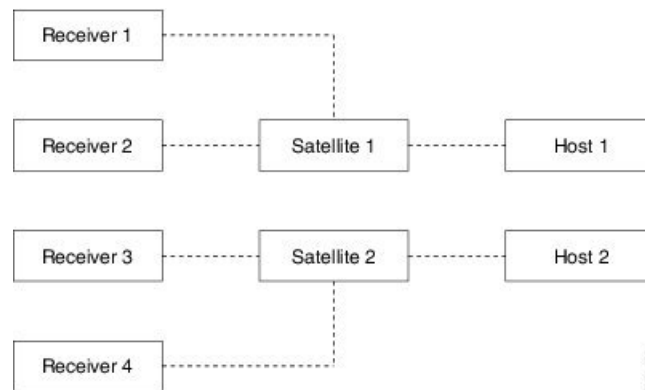
Without offload, Host 1 (assuming that it is the active for both Satellite 1 and Satellite 2) sends 4 copies to receivers 1,2,3 and 4 even if they join the same multicast group, (S1,G1). With offload, Host 1 sends 1 copy and Satellite 1 replicates it twice for receivers 1 and 2 and Satellite 2 replicates it twice for receivers 3 and 4.

The multicast offload feature provides a way for a multicast route which has one or more satellite extended interfaces as its route members known as outgoing interface element lists (OLEs) on the host data plane to be represented locally on the satellite, so that the satellite receives a master copy of a multicast packet from the host and also is capable of replicating that packet into the extended interfaces, that is, the route members (OLEs) of a multicast route.

Whenever a multicast route is offloaded to the satellite, the multicast data received by the satellite on the IC link is tagged with an identifier identifying the multicast route the data belongs to, the satellite hardware reads this tag and through a table lookup finds out the extended interfaces to which the data has to be replicated for a particular route and sends the multicast data on the route-members (OLEs) of that multicast route. The multicast feature channel allows for the host to specify to the satellite which extended interfaces are mapped to which multicast route (tag or multicast isid).

Scope and Prerequisites

These are the supported requirements and prerequisites:

- **Protocol**: IGMP Snooping protocol.

> **Note** There is no support for Satellite nV offload for Layer 3 multicast, IPv6 multicast traffic or MLD snooping protocol.

- **Satellite topologies**: Simple Ring topology with Dual Host, which is a variant of a simple chain with single host, and Hub and spoke topology with a single nV host.

Note Only these topology variants are supported:

- L2 access and L2 core with Satellite Host providing only Layer 2 bridging functionality.

- L2 access and L3 core connected over BVI on the Satellite Host.

- ICL bundles and access bundles on satellites connected in a Hub and Spoke topology.

- **Transport types**:

  - MVPN-MLDP Inband

  - MVPN-GRE

  - MVPN-P2MP-TE

  - L2 bridging, VPLS

  - VPLS LSM

  - Native multicast under BVI for PIM, MLDP, and P2MP-TE

  - MVPN IRB with PIM, MLDP, and P2MP-TE core tree type

Prerequisites:

All existing hardware and software prerequisites for multicast support on the Host are applicable. Similarly, all existing hardware and software prerequisites for the simple ring and dual host topologies are applicable.

For better convergence and redundancy for connectivity failures to the Satellite Hosts or the core, the IGMP snoop sync feature needs to be mandatorily turned on for the redundancy group of the Dual Host System. Otherwise, there can be traffic drops when the Designated multicast forwarder picked by satellite does not align to the Unicast Active Host.

Note For information on bidirectional sync on dual-homed satellites using IGMP Snooping, see the *Bidirectional Internet Group Management Protocol Snoop Synchronization for Satellite Dual-Homed System* section in *Multicast Configuration Guide for Cisco ASR 9000 Series Routers*

Multicast Offload Terminology

These are the different terms associated with the Multicast offload solution.

- **Unicast ISID**: It is the ISID value set in 802.1ah header (802.1ad vlan headers for Cisco NCS 5000 series satellites) for unicast data packet to satellite and for all data packets from satellite to represent satellite access port. Unicast ISID is used to identify the Slot/Subslot/Port of a satellite for unicast packets. It is scoped over a satellite.

- **Multicast ISID**: It is the ISID value set in 802.1ah header (802.1ad vlan headers for Cisco NCS 5000 series satellites) for multicast data packet from host to satellite to represent offloaded Multicast Route (S, G) per Ring.

- **Designated Multicast Forwarder**: Each satellite selects a host as the designated multicast forwarder and replicates multicast packets for all receivers (satellite access ports) for route (S, G) from the designated multicast forwarder. This term applies to the simple ring Satellite nV System topology.

- **Backup Multicast Forwarder**: Multicast packets received from Backup Multicast forwarder are not replicated on the satellite for access ports. This term applies to the simple ring Satellite nV System topology.

- **Active Unicast Host**: The unicast customer data traffic is switched through Active Unicast Host while traversing through the satellite.

- **Standby Unicast Host**: In case of a lost connection to Active Unicast Host due to failures such as cut cables or connection interface failure, the Standby unicast host shall become the new Active Unicast Host.

- **Primary Host**: Specifies the Host with the lowest chassis MAC in a Dual head topology.

- **Secondary Host**: Specifies the Host with a higher chassis MAC than the Primary Host in Dual head topology.

- **Master OLE**: Specifies an OLE selected amongst the active OLEs of an offloaded route to represent the receiver of the single instance/ replication of offloaded multicast traffic sent from the host to the satellite. For ring topology, one master OLE is selected per offloaded route (S,G), for each BD and each ring. For Hub and Spoke topology, the master OLE is scoped for each offloaded route, BD and satellite.

Overview of Multicast Offload

As in the case of existing Satellite nV System architecture, the IGMP snooping protocol runs on the satellite hosts with full Cisco IOS XR Software feature parity even with offload. The joins received on the satellite access ports reach the active host and gets processed by the IGMP snoop module.

- The IGMP snoop sync feature running over ICCP protocol synchronizes the joined multicast group membership information to the other host. This keeps the IGMP protocol state for receivers joining over satellite access ports in sync across both hosts in the case of dual host topology.

- In the case of dual host topology, the hosts independently offload eligible routes to the relevant satellite devices through SDAC. This includes the list of local ports that have expressed interest in this route.

- The joins also get propagated to the core by each of the hosts (dual host topology), if they have an active link to the core. This allows better convergence in case of a redundancy switchover on the satellites.

- As the same traffic stream comes in from both the hosts (in the case of dual host topology), the satellites pick a designated multicast forwarder at discovery and continue to replicate from that host until that host goes away or loses its core connectivity. Thus the offload topology is a full active / active system.

Figure 2: IGMP Join Programming (Access L2 and Core L2/BVI)
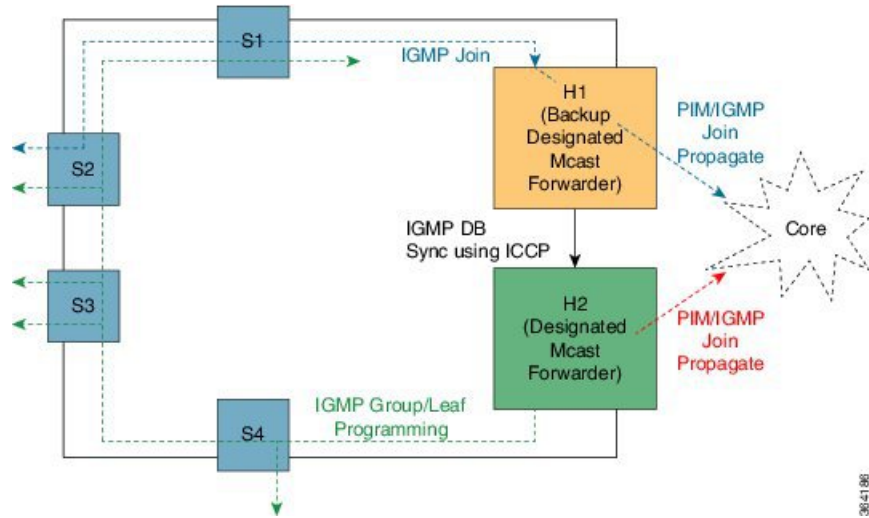

Figure 3: Multicast Offload Data Flow
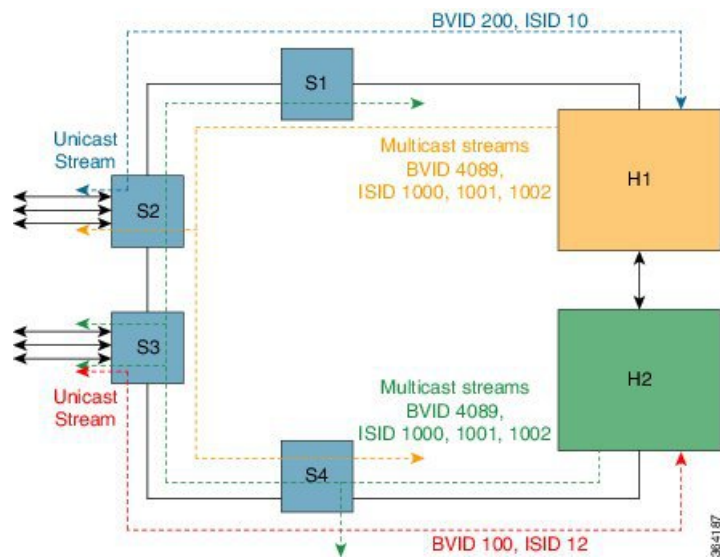
These illustrations show the multicast offload functionality on Satellite nV System that use hub and spoke topologies.

Figure 4: Multicast Offload Data Flow with ICL bundles to different satellites in Hub and Spoke topology
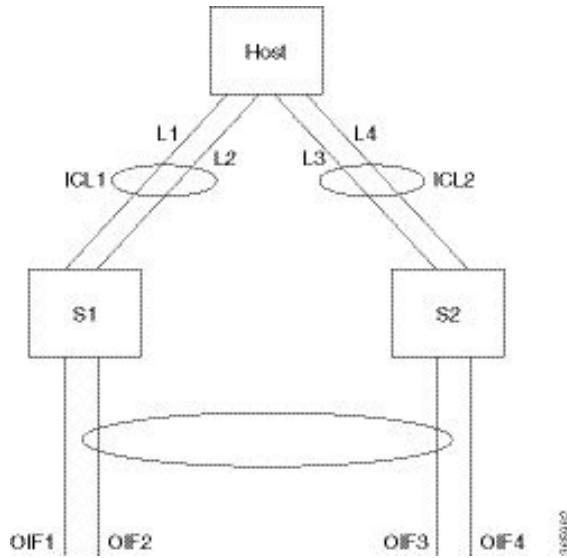
Figure 5: Multicast Offload with ICL bundle from different line cards to a satellite
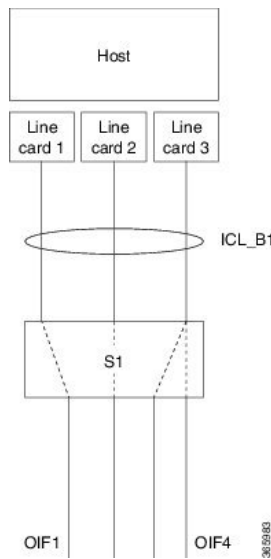
Behavior of Multicast Offload During Satellite Topology Events

These are the behavior of the Multicast offload during various satellite topology events:

Note These are applicable in the case of dual host ring topology setup.

- **Split brain**: During a split brain event, the hosts function independently and the system behaves as two separate chains with IGMP snoop sync broken between satellites and each host offloading to its own set

of satellites. On recovery, there is reconciliation and some reprogramming that can result in traffic loss as IGMP snoop protocol states sync up.

- **RSP failover**: This is expected to be hitless for traffic going over routes that have been offloaded already. New joins during the failover may experience the same set of events and traffic impact as the split brain condition.

- **Ring break and Recovery**: On ring breakage, any satellites aligned to other host as designated multicast forwarder has to realign to the remaining host on the ring. This can have a minimal (sub-second for L2 core) traffic impact on a stable system where the routes are already synchronized. On recovery, there is no revertive switchover and so there must be no traffic impact.

Note For more information, refer *Features of Dual Home Network Architecture* section in the *Configuring the Satellite Network Virtualization (nV) System* chapter.

Difference between Layer 2 and Layer 3 (over BVI) core support

For the Dual host topology systems connecting to an Layer 3 core through BVI, the current recommendation is to align all satellites to a single host as Unicast active. This is required as BVI protocol states are kept down on the standby host to avoid ECMP being triggered from the core for an active/standby Satellite nV system that could lead to traffic drop of half the packets on the standby host.

Currently, multicast offload which is implemented as an inherently active/active system needs to track the BVI redundancy states as well to avoid picking the Designated multicast forwarder on the host having the standby BVI. While this ensures a common feature model with the Layer 2 core, the benefits of an active/active system, especially, the sub-second switchover convergence is not valid anymore. This can be improved by at least allowing IGMP snoop sync to continue by configuring an internal querier with system IP address lower than the BVI IP address and a query max timeout of 1s.

Uplink/Core Isolation Tracker

Under redundancy group configuration, there is an option to configure a backbone interface. Multiple backbone interfaces can be configured depending on the number of links from the Host to the core.ICCP protocol keeps a watch over the link states of these interfaces and if all of these backbone interfaces go down, then a core isolation event is notified to the client.

As part of the multicast offload feature, IGMP Snoop registers as a client to ICCP protocol and propagates these notifications to the satellite devices. Any satellite which still expect this host to be the designated multicast forwarder can then switchover to the other host as core connectivity is lost. In order to avoid traffic downtimes during flaps, this event is triggered only for core link up to down condition and is non-revertive. A satellite device stays with a host as a designated multicast forwarder until it goes down and does not switch back if the previous host that triggered a switchover comes up.

For more information, see *Configuring Interchassis Communication Protocol* section in the *Interface and Hardware Component Configuration Guide for Cisco ASR 9000 Series Routers*

Multicast Offload Scalability

These are the maximum scale capabilities of Multicast Offload Solution:

- The maximum number of rings over which the same offloaded route can be downloaded is 32.

- The maximum number of offloaded routes for each satellite or ring is 2000. The maximum number of multicast routes for each satellite in the hub and spoke topology is 1000.

- The maximum number of offloaded routes for each system is equal to the maximum number of rings multiplied by 2000 up to the Host limit (128k).

- The maximum bandwidth of ICL to ICL traffic with offload enabled for each satellite is 6Gbps in each direction.

Multicast Offload Use Case Scenario

The Satellite nV multicast offload feature has been designed and optimized specifically for an IPTV scenario where the Cisco ASR 9000 Series Satellite nV System replaces the traditional edge or aggregate router and switch solution. The Satellites feed multicast traffic from the Layer 2 or Layer 3 (through BVI) core to the Layer 2 access trunk links connecting DSLAMs which in turn terminate residential gateways and take care of subscriber aware processing, if any, as shown below.

Residential Gateway ---- DSLAM ---- Cisco ASR 9000 Satellite nV System---- Core ---- IPTV source

Therefore, the Satellite nV multicast offload solution does not need to support VLAN rewrite operations on individual offloaded replications as they all go over the same trunk video VLAN. This reduces the offload processing overhead on the satellite devices to achieve line rate replication. Similarly, the solution is optimized to an use case that requires minimal egress feature processing. QoS, ACL and other egress features act on the pre-replicated multicast stream and the configuration needs to be replicated to all participating OLEs in the offload, if at all feature processing is specifically required.

Network design needs to ensure that there is no congestion on the satellites post replication. This is critical and different from a non-offload solution as the Host takes care of correct priority aware traffic shaping through Auto QoS or a user specified MQC policy for the non-offload case. For offloaded traffic, as the replication happens locally on the satellite devices, the Host QoS is unaware of the total traffic volume post replication, and therefore cannot include it in its bandwidth computations without statically reserving bandwidth on all offload participating OLEs.

Such a permanent reservation might be sub-optimal in most cases and a rigid reservation may not cater to user needs. However, a solution based on intelligent network design is generally straightforward for the IPTV roll outs, as they have well defined bandwidth planning, given the load generally remains constant over time. For residential triple and quadruple play cases, with possibility of other priority and internet traffic causing oversubscriptions, a simple QoS port-shaper policy on the egress video traffic DSCP/CoS markers or VLAN can be used to classify and allocate a reserved bandwidth for multicast offloaded video streams by shaping the remaining traffic to 60-70% of the access port interface bandwidth as required.

Due to various scale bottlenecks and for practical reasons, the need for a simple service QoS model as above the Satellite nV multicast offload solution is recommended to be deployed on a -SE version of the ethernet line cards.

Restrictions for Multicast Offload

These are the restrictions and considerations of the Multicast offload feature:

- Only Layer-2 multicast offload through IGMP snooping for IPv4 multicast traffic is supported. IPv6 multicast traffic / MLD snoop and Layer-3 multicast traffic replication offload is not supported.

- Offloaded OLE interface statistics (except for the master OLE) are only incremented on the main interface, even if the OLE is actually on a sub-interface. L2VPN statistics reflect updates against the correct OLE and multicast route.

- Endian type mismatching RSPs across the two hosts of the Dual Host system is not supported. IGMP snoop sync, a pre-requisite for this feature does not function on such hardware.

- Multicast offload does not support broadcast, multicast router port, or TCN flood event offloading. All such traffic is replicated on the host.

- Only normal Satellite EFPs are considered as OLEs eligible for offload. An OLE consisting of a pseudo wires and BNG subscribers over satellite ports are not supported as eligible OLEs for offload. These might fall back to legacy host side replication.

- VLAN rewrite support for offloaded replications is not supported. All offloaded replications carry the same VLAN tag as the original copy sent by the Host over the dynamically picked master OLE. The same VLAN rewrite configuration, if required, needs to be added on all the EFPs of a route as any OLE can be dynamically selected as the master OLE on host. The same applies for any other egress feature configuration on the OLEs.

- Split horizon does not work with offloaded satellite OLEs on a bridge domain. So, multicast source or multicast routers (mrouter port) must not be over satellite EFPs within the same bridge domain where receivers through offloaded satellite EFPs are also present.

- Offload functionality is disabled by default. If it is enabled, there will be a transient downtime including impact to traffic for all IGMP snoop routes on the system as the offload mode switches. Similarly, all routes (including non-offloaded routes) in the system are impacted if offload is disabled on any of the bridge domains as IGMP snoop moves to non-offload mode.

  Note Local replication can co-exist with offloaded replication for same route/bridge domain/satellite combinations.

- While sub-second convergence and fast reroutes are possible for an Layer-2 core failover scenarios (ring break, split brain, core link down and so on), L3 core (over BVI) requires the BVI redundancy state toggles post failover before protocol states can be synchronized between Hosts. This cold failover mode for L3 core over BVI is slower and in the order of seconds compared to the Layer-2 core case.

- VLAN ranges cannot be used for IGMP ports involved in the sync between the Dual Hosts of a satellite system. Unambiguous VLAN id configuration is required.

- Cisco ASR 9000v and Cisco ASR9000v-V2 only support 12 Gbps of ICL to ICL bidirectional multicast replication and a maximum of 6Gbps for each direction while processing multicast offloaded traffic even if it does not actually participate in local replications. With respect to Hub and Spoke topology, support for multicast offload traffic processing on the satellite is restricted to 12 Gbps or the ICL bandwidth, whichever is lower, in the case of single host topology and 6 Gbps for each host in the case of dual host topology.

- The offload solution has been optimized and specifically characterized for the scale numbers mentioned. There might be a fallback to legacy Host side replication if any of those numbers are exceeded or if any of the limitations stated above make an OLE ineligible for offload. However, these fallbacks may be delayed or may not work gracefully, especially in scale exceed cases. Hence, network design must ensure that the scale numbers are complied for optimal performance.

- On reload of one of the hosts in the dual host system, the hosts might go out of sync because offload failures are not synchronized between hosts. A restart of the IGMP snoop process or a receiver leave/join can resolve this. The impact is only for scale exceed cases, which is not recommended in general.

- In the case of Layer 2 Multicast offload, the satellite does not support any egress features on the replicated packet on satellite. Egress features such as ACL, QoS shaping/ policing, VLAN rewrite, SPAN, netflow will work on the host on the Master OLE copy sent to the satellite. If any egress feature need to applied to offloaded multicast streams, it needs to be configured on all the EFP as any OLE can be selected as Master OLE on host.

- ICLs (physical interfaces or bundle interfaces) must be main interfaces. This means that sub-interfaces on ICLs are not supported.

- You cannot configure satellite rings and satellite hub and spoke topologies on the same host if any of the satellites are participating in Layer 2 multicast offload. The satellites can be configured in either ring or hub and spoke topologies on a per-router basis.

Configuring Satellite nV Multicast Offload

Prerequisites

All the existing configuration for implementing multicast and Satellite nV System have to be configured and the multicast offload is only an incremental function that allows for the final routes to be offloaded to the satellite instead of the local Cisco ASR 9000 Series Router line card hardware.

By default, multicast offload is disabled on the host.

Note Refer the *Multicast Configuration Guide for Cisco ASR 9000 Series Routers* for information on Multicast routing and IGMP snooping. All the configuration has to be manually synchronized between the two hosts for the functionality to perform correctly.

Enabling Multicast Offload on a Bridge Domain

To enable multicast offload use the **nv satellite offload ipv4 multicast enable** command:

```
RP/0/0/CPU0:(config-l2vpn-bg-bd-ac)#show run l2vpn

l2vpn
 bridge group bg1
  bridge-domain bd1
    nv
     offload ipv4 multicast enable
```

By default, multicast offloading is disabled on all bridge domains. To enable IPv4 Multicast offloading (including IGMP), run this command:

```
(config)#l2vpn
(config-l2vpn)#bridge group <bg>
(config-l2vpn-bg)#bridge-domain <bd>
(config-l2vpn-bg-bd)#nv
(config-l2vpn-bg-bd-nv)#offload ipv4 multicast enable
```

Note This configuration needs to be enabled on both hosts participating in the Dual Host system of the Ring topology.

Configuration examples for Satellite nV multicast offload

These show command outputs provide the necessary details about multicast offload:

```
RP/0/RSP0/CPU0:routershow igmp snooping group summary debug

Bridge Domain bg:bd

                                    #Mem  #Inc  #Exc  Annot                        Master
      #Mem   Master OLE
Group            Source       Ver GM  Ports Ports Ports Key            ISID  Ver  OLE XID
      Annot InterfaceName
-----           ------       --- -- ----- ----- ----- ------------ ----- ---- ----------
      ----- -------------
225.0.0.1      {}            V3  EX  2     -     -     0x2-c8         0x3e8 0x4  0xa0000005
   2     Bundle-Ether50
225.0.0.2      {}            V3  EX  2     -     -     0x2-64         0x3e8 0x4  0xa0000005
   1     Bundle-Ether50

                                                      0x2-c8         0x3ed 0x6  0xa0000004
   1     Bundle-Ether40
225.0.0.3      {}            V3  EX  2     -     -     0x2-c8         0x3e9 0x4  0xa0000005
   2     Bundle-Ether50
225.0.0.4      {}            V3  EX  2     -     -     0x2-c8         0x3ea 0x4  0xa0000005
   2     Bundle-Ether50
225.0.0.5      {}            V3  EX  2     -     -     0x2-64         0x3e9 0x4  0xa0000005
   2     Bundle-Ether50
225.0.0.6      {}            V3  EX  2     -     -     0x2-c8         0x3eb 0x4  0xa0000005
   2     Bundle-Ether50
225.0.0.7      {}            V3  EX  2     -     -     0x2-64         0x3ea 0x4  0xa0000005
   1     Bundle-Ether50

                                                      0x2-c8         0x3ee 0x6  0xa0000004
   1     Bundle-Ether40
225.0.0.8      {}            V3  EX  2     -     -     0x2-64         0x3eb 0x4  0xa0000005
   1     Bundle-Ether50

                                                      0x2-c8         0x3ef 0x5  0xa0000004
   1     Bundle-Ether40
225.0.0.9      {}            V3  EX  2     -     -     0x2-c8         0x3ec 0x4  0xa0000005
   2     Bundle-Ether50
225.0.0.10     {}            V3  EX  2     -     -     0x2-64         0x3ec 0x4  0xa0000005
   2     Bundle-Ether50


RP/0/RSP0/CPU0:routershow igmp snooping port group debug

Key: GM=Group Filter Mode, PM=Port Filter Mode
Flags Key: S=Static, D=Dynamic, E=Explicit Tracking, R=Replicated, I=Inherited OLE
OLE Offl Key: F=False, T=True, d=Offload disabled, H=Offload disabled, not in SHG-0,
              M=Offload disabled, mrouter port, t=Offload disabled, topo not supported

                       Bridge Domain bg:bd


                                                                        OLE
Port                   PM Group          Ver GM Source       Exp  Flgs  Offl
Ul_ifh    InterfaceName
----                   -- -----          --- -- ------       ---  ----  ----
---------- -------------
BE40                   -  225.0.0.1      V2  -  *            92   D     T
```

```
0x000036e0 GigabitEthernet200_0_0_4
BE40                           -  225.0.0.2      V2  -  *              86    D    T
0x00003720 GigabitEthernet200_0_0_3
BE40                           -  225.0.0.3      V2  -  *              87    D    T
0x00003760 GigabitEthernet200_0_0_2
BE40                           -  225.0.0.4      V2  -  *              85    D    T
0x000036e0 GigabitEthernet200_0_0_4
BE40                           -  225.0.0.5      V2  -  *              90    D    T
0x000037e0 GigabitEthernet100_0_0_2
BE40                           -  225.0.0.6      V2  -  *              87    D    T
0x00003760 GigabitEthernet200_0_0_2
BE40                           -  225.0.0.7      V2  -  *              88    D    T
0x00003720 GigabitEthernet200_0_0_3
BE40                           -  225.0.0.8      V2  -  *              89    D    T
0x00003720 GigabitEthernet200_0_0_3
BE40                           -  225.0.0.9      V2  -  *              86    D    T
0x00003760 GigabitEthernet200_0_0_2
BE40                           -  225.0.0.10     V2  -  *              88    D    T
0x000037e0 GigabitEthernet100_0_0_2
BE50                           -  225.0.0.1      V2  -  *              85    D    T
0x000037a0 GigabitEthernet200_0_0_1
BE50                           -  225.0.0.2      V2  -  *              91    D    T
0x000036a0 GigabitEthernet100_0_0_1
BE50                           -  225.0.0.3      V2  -  *              90    D    T
0x000037a0 GigabitEthernet200_0_0_1
BE50                           -  225.0.0.4      V2  -  *              85    D    T
0x000037a0 GigabitEthernet200_0_0_1
BE50                           -  225.0.0.5      V2  -  *              84    D    T
0x000036a0 GigabitEthernet100_0_0_1
BE50                           -  225.0.0.6      V2  -  *              85    D    T
0x000037a0 GigabitEthernet200_0_0_1
BE50                           -  225.0.0.7      V2  -  *              90    D    T
0x000036a0 GigabitEthernet100_0_0_1
BE50                           -  225.0.0.8      V2  -  *              90    D    T
0x000036a0 GigabitEthernet100_0_0_1
BE50                           -  225.0.0.9      V2  -  *              89    D    T
0x000037a0 GigabitEthernet200_0_0_1
BE50                           -  225.0.0.10     V2  -  *              85    D    T
0x000036a0 GigabitEthernet100_0_0_1


RP/0/RSP0/CPU0:router**show nv satellite multicast satellite 210**

---------------
Satellite 210
---------------
Multicast channel state: Open
Host state: Designated Forwarder (DF)
RP/0/RSP0/CPU0:SAT-SCALE-HOST1#
```

Note   For more information on IGMP Snooping commands, see *Multicast Command Reference for Cisco ASR 9000 Series Routers*. For more information on L2VPN forwarding commands, see *VPN and Ethernet Services Command Reference for Cisco ASR 9000 Series Routers*