Application Note

IGMP Proxy Model for IPTV

Laboratory Testing of the IGMP Proxy Model Including Setup, Methodology, JUNOSe Commands and Test Report Excerpts

Juniper Networks, Inc.
1194 North Mathilda Avenue
Sunnyvale, California 94089
USA
408.745.2000
1.888 JUNIPER
www.juniper.net

Part Number: 350109-001 October 2007

Table of Contents

Introduction....3
Scope....3
Description and Deployment Scenario....3
C-VLAN Subscriber Management with M-VLAN....4
Edge Multicast Replication....5
Edge Multicast Admission Control....6
Edge QoS and Multicast....7
Summary....12
About Juniper Networks....12

Introduction

This document describes the methodology for and results from lab tests of an IGMP proxy edge model for IPTV. It discusses IGMP and quality of service (QoS) capabilities provided by the broadband services router (BSR)—in this case, the Juniper Networks E320 Broadband Services Router (E320 BSR)—and the access network (AN). The AN in this model provides per-consumer replication in the data plane and the proxy feature provides IGMP proxy reporting and report suppression in the control plane. The document is specific to DSL networks using a DSLAM for access.

The IGMP proxy model optimizes multicast bandwidth in the access network at the cost of QoS requirements in the digital subscriber line access multiplexer (DSLAM). Originally intended for high-usage multicast TV service, this model is simple and well defined in the industry. One potential drawback is the continued growth of unicast TV services, which create bandwidth management challenges for this model.

Scope

This application note is specific to DSL networks using a digital subscriber line access multiplexer (DSLAM) for access. This test case uses DHCP as the IP assignment protocol. PPPoE is not covered in this test plan because many ANs do not support IGMP processing within a transit Point-to-Point Protocol (PPP) session.

Juniper Networks has also published application notes describing the testing of IPTV models for IGMP passthrough and IGMP transparent snooping.

Description and Deployment Scenario

Figure 1 shows the high-level topology in which the BSR is connected directly to the DSLAMs that serve as access nodes. Multicast groups are replicated to the BSR over the packet network. The BSR is the last branch in each multicast group tree.

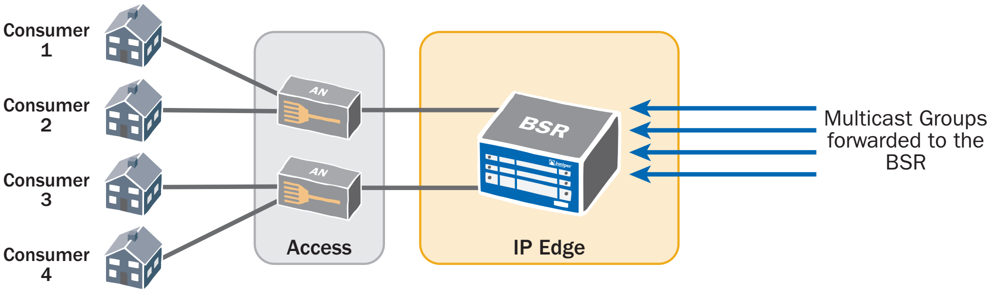

Figure 1. High-Level Topology for Testing

In this model, the BSR processes IGMP messages to determine which DSLAMs should be leaf nodes for each specific multicast group and, when required, replicates a unique copy of a multicast group to the DSLAM. Along with IGMP processing, the BSR allows only a predefined amount of multicast traffic to be forwarded per DSLAM to control the bandwidth used for multicasting.

Because the AN is the last replication point prior to the consumer domain, it must support IGMP proxy (or snooping with report suppression) to handle IGMP requests for location replication.

The BSR also prioritizes traffic using a QoS scheduler with per-subscriber and per-application traffic management for unicast traffic only. The BSR uses a VLAN shaper equalized to the DSL loop speed to remove any requirements for QoS features in the DSLAM. This configuration may oversubscribe the DSL port speed when multicast is inserted at the DSLAM unless the DSLAM supports basic QoS.

C-VLAN Subscriber Management with M-VLAN

The DSLAM maps each consumer DSL port to a VLAN identifier known as the customer VLAN (C-VLAN). This C-VLAN identifier is local to the DSLAM and BSR, so the numbering scheme can be reused for other DSLAM-BSR connections. The C-VLAN tag value combined with the BSR port creates a unique identifier for each consumer in the BSR.

A single multicast VLAN (M-VLAN) configured between the BSR and DSLAM is shared by all customers on the DSLAM. This M-VLAN ID can also be reused across ports but is locally significant when coupled to the BSR port. There could also be multiple M-VLANs per port in cases where S-VLAN (stacked VLANs) is used with multicast replicated per VLAN groups using a common outer tag.

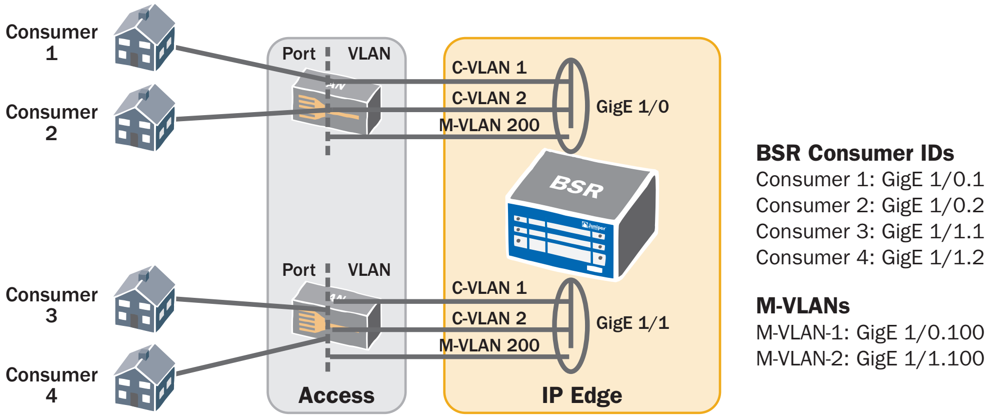

Figure 2. L2 Topology – C-VLAN per Consumer with M-VLAN

A consumer device that requests an IP address from the service provider domain obtains the address and related parameters from a DHCP server. The DHCP server may be contained within the BSR (local server) or may be external. In the case of the external server, the BSR relays DHCP information between the DHCP client and server. Successful completion of the Dynamic Host Configuration Protocol (DHCP) negotiation results in a host route being added into the local routing table, mapping the assigned host address to its respective C-VLAN.

Edge Multicast Replication

In the proxy model, the AN replicates multicast by subscriber and provides report suppression so that subscriber join and leave events are not forwarded to the BSR. As shown in Figure 3, the STB sends the IGMP join/leave information. The DSLAM captures and handles all consumer IGMP events locally. The DSLAM then uses IGMP proxy reporting to notify the BSR of multicast groups that need to be replicated and sent to the DSLAM over the M-VLAN. From the BSR's perspective, the DSLAM appears as a single multicast receiver that is capable of receiving multiple multicast groups.

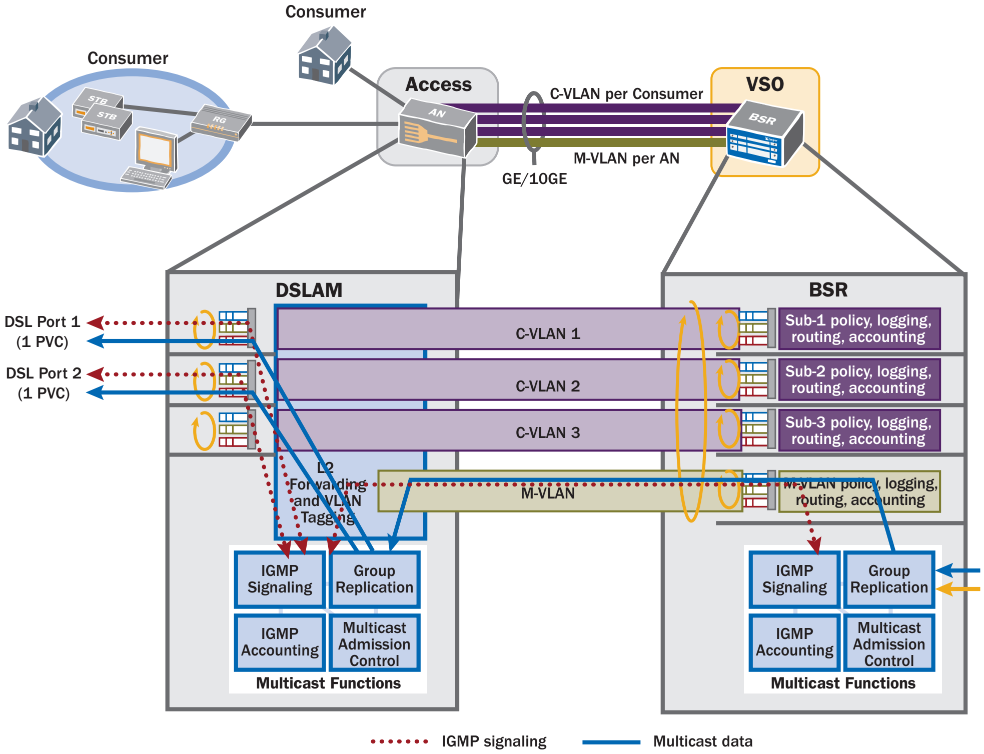

Figure 3. Multicast Signaling and Replication Using IGMP Proxy

If two subscribers on the same DSLAM join a multicast group, the DSLAM handles the multicast replication locally. The BSR receives a single join instance over the M-VLAN interface and need only forward a single copy of the multicast group to the DSLAM using the M-VLAN interface.

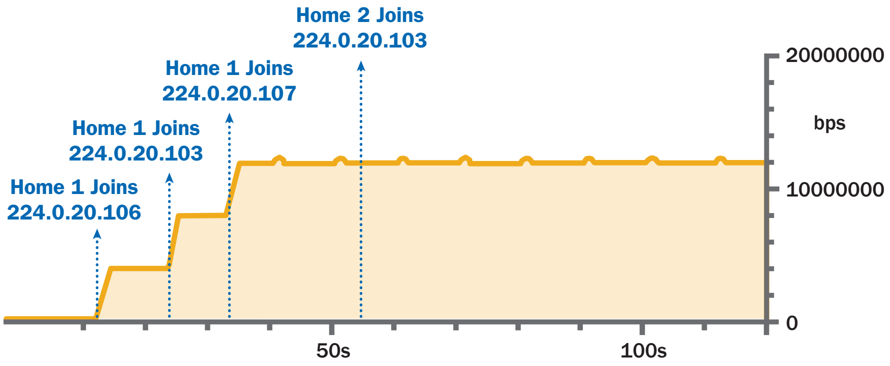

Figure 4. Only Unique Group Joins Increase BSR-DSLAM Link Utilization

Figure 4 shows how this model uses multicast bandwidth efficiently. Multiple subscribers joined to the same multicast group share the same multicast stream forwarded from the BSR. This design results in a higher number of projected subscribers per BSR interface for deployments where multicasting is the primary driver for network traffic into the home. For applications in which unicast applications drive a higher percentage of traffic compared to multicast, the IGMP pass-through model (described in a companion Application Note) may be more cost effective than the proxy model.

Edge Multicast Admission Control

The Juniper E320 BSR can limit the amount of bandwidth consumed by multicast traffic using its multicast admission control capability. Multicast admission control evaluates each IGMP join against a maximum bandwidth limit to determine if replication is allowed. This mechanism augments QoS in situations where it is possible to congest the interface by oversubscribing subscriber bandwidth or by endpoints issuing joins outside a predefined service profile.

Each multicast group is assigned an admission bandwidth. The multicast admission control engine uses this value to determine whether the next join would cause the total bandwidth requested to exceed the admission bandwidth limit. Each multicast group can be statically configured with an admission bandwidth, or the system can dynamically measure each group bandwidth when received.

For the IGMP proxy model, the M-VLAN multicast admission control is configured with the maximum amount of multicast bandwidth allowed to be replicated across that interface using the following Juniper Networks JUNOSe command:

```
ip multicast admission-bandwidth-limit 15000000
```

The limit of 15 Mbps allows up to three unique groups of 3.5 to 4.0 Mbps SDTV video streams to be joined. However, multiple subscribers can watch the same streams, so the model extends to a large number of views. The show ip mroute output shows that the first three groups (224.0.20.103, 224.0.20.106, 224.0.20.107) were joined successfully.

```
E320_MRA_ER_0#show ip mroute
…
(11.1.5.66, 224.0.20.103) uptime 0 00:00:38
   Admission bandwidth: 4005000 bps (adaptive)
   QoS bandwidth: 4005000 bps (adaptive)
   RPF route: 11.1.5.0/24, incoming interface GigabitEthernet3/1/2
      neighbor 11.1.6.2, owner IsIs (ECMP route)
   Incoming interface list:
      GigabitEthernet2/1/3 (11.1.6.1/30), Discard/Pim
      GigabitEthernet3/1/2 (11.1.6.5/30), Accept/Pim (RPF IIF)
   Outgoing interface list:
      GigabitEthernet15/0/0.222 (192.168.13.1/24), Forward/Pim, 0 00:00:38/
never

(11.1.5.66, 224.0.20.106) uptime 0 00:00:41
   Admission bandwidth: 3864000 bps (adaptive)
   QoS bandwidth: 3864000 bps (adaptive)
   RPF route: 11.1.5.0/24, incoming interface GigabitEthernet2/1/3
      neighbor 11.1.6.2, owner IsIs (ECMP route)
   Incoming interface list:
     GigabitEthernet2/1/3 (11.1.6.1/30), Accept/Pim (RPF IIF)
   Outgoing interface list:
      GigabitEthernet15/0/0.222 (192.168.13.1/24), Forward/Pim, 0 00:00:41/
never
```

```
(11.1.5.66, 224.0.20.107) uptime 0 00:00:21
   Admission bandwidth: 4252000 bps (adaptive)
   QoS bandwidth: 4252000 bps (adaptive)
   RPF route: 11.1.5.0/24, incoming interface GigabitEthernet3/1/2
      neighbor 11.1.6.2, owner IsIs (ECMP route)
   Incoming interface list:
     GigabitEthernet2/1/3 (11.1.6.1/30), Discard/Pim
      GigabitEthernet3/1/2 (11.1.6.5/30), Accept/Pim (RPF IIF)
   Outgoing interface list:
     GigabitEthernet15/0/0.222 (192.168.13.1/24), Forward/Pim, 0 00:00:21/
never
```

However, the fourth group (224.0.20.108) join received by the BSR exceeds the admission bandwidth limit, resulting in a non-replicated group:

```
(11.1.5.66, 224.0.20.108) uptime 0 00:00:04
   Admission bandwidth: 3752000 (adaptive)
   QoS bandwidth: 3752000 bps (adaptive)
   RPF route: 11.1.5.0/24, incoming interface GigabitEthernet2/1/3
      neighbor 11.1.6.2, owner IsIs (ECMP route)
   Incoming interface list:
      GigabitEthernet2/1/3 (11.1.6.1/30), Accept/Pim (RPF IIF)
   Outgoing interface list:
      GigabitEthernet15/0/0.222 (192.168.13.1/24), Blocked (intf-adm-limit)/
Pim, 0 00:00:05/never
```

This group is shown as ***blocked***, denoting that group replication is not occurring due to an interface administrative limit (intf-adm-limit). This type of blocking impacts any subscriber issuing a join for a new multicast group that must be delivered over the M-VLAN to the DSLAM. If the DSLAM proxy join is beyond the 15 Mbps M-VLAN bandwidth limit, then DSLAM replication cannot occur.

If per-subscriber admission control limits are required, the DSLAM must provide these controls. This aspect of the model was not tested since the DSLAM used in the test scenario did not support this feature. The BSR cannot provide per-subscriber multicast control since per-subscriber replication is handled by the DSLAM, not the BSR.

Edge QoS and Multicast

The BSR and the DSLAM perform Edge QoS with only unicast packets passing through the C-VLAN scheduler node in the BSR. The BSR marks the packets so that the DSLAM can provide simple Layer 2 (L2) priority class of service (CoS), ensuring that low-priority traffic is dropped when the combination of unicast and multicast traffic congests the DSL port.

A VLAN policy is configured in the E320 as shown in the following:

```
vlan policy-list DSLAM-QoS
  classifier-group “voice-egress”
    mark-user-priority 4
  classifier-group “vod-egress”
    mark-user-priority 2
  classifier-group “be-egress”
    mark-user-priority 0
```

This VLAN policy is then attached to the C-VLAN interface:

```
interface gigabitEthernet 15/0/0.223
  vlan id 223
    vlan policy output DSLAM-QoS statistics enabled baseline enabled
```

The BSR uses H-QoS capabilities with a set of queues assigned to each subscriber where each queue supports a unique traffic class. The H-QoS scheme provides both application priority and subscriber fairness. The QoS profile used for testing is shown below.

```
E320_MRA_ER_0#show qos-profile IPTV-CVLAN
qos-profile IPTV-CVLAN:
t-class interface rule    traffic    scheduler    queue    drop     statistics
group   type   type       class      profile      profile  profile  profile
------- --------- ----- ----------- --------- ------- ------- ----------
        vlan   node                  14M-CVLAN
        vlan   queue best-effort     be           default  default  iptv-stat
        vlan   queue EF              voice        default  default  iptv-stat
        vlan   queue gaming          gaming       default  default  iptv-stat
        vlan   queue VOD             VOD          default  default  iptv-stat
```

Each C-VLAN has a VLAN node shaper to limit the amount of egress traffic per subscriber and thus per DSL port. In the test case, a value of 14 Mbps is used. The VLAN shaper value can be statically configured, applied during initial consumer network attachment through an Authentication, Authorization and Accounting (AAA) or policy system, or can be obtained from the DSLAM using ANCP.

The test network includes individual queues for video on demand (VOD), gaming, expedited forwarding (EF) voice, and best-effort traffic. In the test case, the Internet service is allowed to burst to the full VLAN node-shaping rate to show that bandwidth carve-out is not required between services. This is highlighted in the diagram in Figure 5 with the corresponding Figure 6 analyzer trace showing the Internet bandwidth utilized in the test C-VLAN.

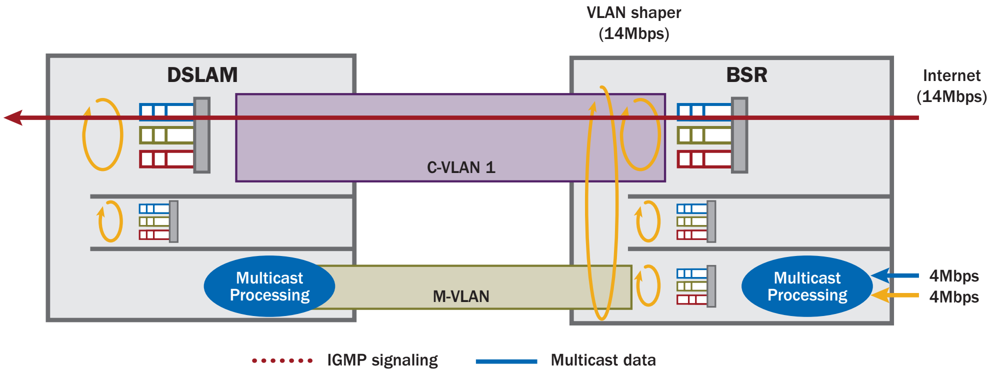

Figure 5. Unicast-Only Data Flow

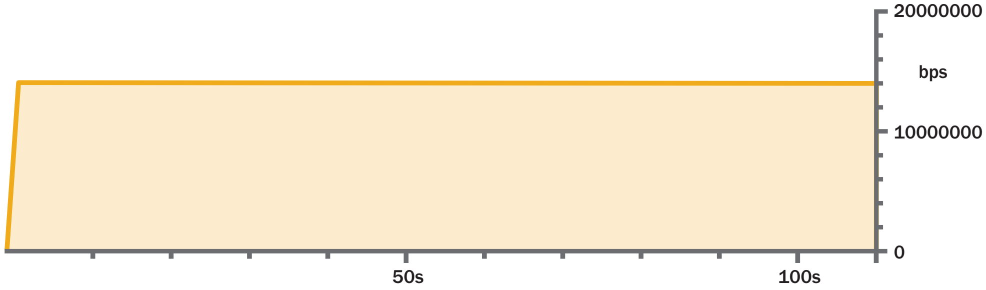

Figure 6. Internet Utilization on the BSR-DSLAM Link for C-VLAN 1

When an IGMP join is sent from the home, the DSLAM only forwards an IGMP join up to the BSR if it needs to receive the requested multicast group. The BSR is unaware of any multicast group replication downstream and continues to forward Internet traffic at the full VLAN shaping rate. Although Internet traffic is shown, this applies to all unicast flows that are prioritized at the BSR but without any per C-VLAN scheduler back pressure due to the multicast packets sent by the DSLAM onto the DSL port.

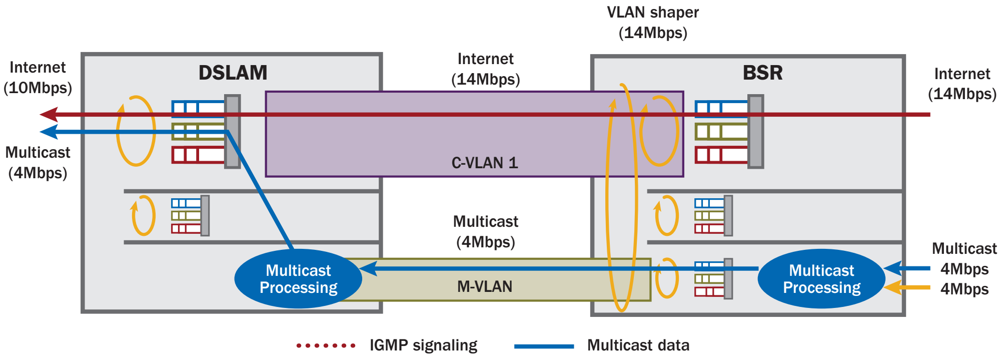

Figure 7. QoS Requirement in the DSLAM

Even when the multicast join happens and a 3.75 Mbps stream is sent into the home, the BSR still forwards unicast to the DSLAM at the 14 Mbps shaping rate, along with the multicast traffic forwarded over the M-VLAN, as shown in Figure 7. The BSR does not react to the multicast replication downstream. The link utilization should decrease for TCP applications where backoff can occur due to packet drops in the DSLAM.

The following report shows E320 M-VLAN and C-VLAN queue statistics for this case. All multicast traffic is contained in the M-VLAN shown as interface Gigabit Ethernet 15/0/0.222. The C-VLAN interface Gigabit Ethernet 15/0/0.22 only forwards unicast applications such as VOD, best-effort, voice (EF) and gaming. The C-VLAN scheduler will provide prioritization so that only the best-effort traffic is dropped during periods of congestion.

```
 E320_MRA_ER_0#show egress-queue rates interface gigabitEthernet 15/0/0.222
                       traffic  forwarded   aggregate   minimum   maximum
     interface          class      rate      drop rate    rate      rate
------------------------------ --------- --------- --------- --------- -------
vlan GigabitEthernet15/0/0.222 MULTICAST 11753784     0 133495000 200000000
```

```
E320_MRA_ER_0#show egress-queue rates interface gigabitEthernet 15/0/0.22
                        traffic  forwarded   aggregate   minimum   maximum
     interface          class      rate      drop rate    rate      rate
---------------------------- ----------- --------- --------- ------- --------
vlan GigabitEthernet15/0/0.22 best-effort 13962576     0  459000    14000000
                              EF                 0     0  459000    500000
                              gaming             0     0  459000    500000
                              VOD                0     0  4000000   4000000
```

If the DSLAM does not provide QoS, then all packets are placed into a common egress queue. All packets – Internet and video – are sent into a common queue that can become congested, resulting in queue tail drops. The DSLAM lacks the ability to prioritize application flows. Congestion can occur due to the lack of back pressure in the E320 unicast scheduler when multicast is replicated downstream out the DSL port. The following report shows a large number of dropped packets due to congestion.

```
officer SEC>> show interface 11.2 queuecount
   --- Egress Queue Statistics ---
                                                                       Max
                  Dropped                                         Queue  Queue
Interface      Q  Packets         Sent Packets     PBit         Depth  Depth
-------------- -  --------------  --------------   ----------   ------ -------
11.2.0         7               0               0            7        1      20
               6               0               0            6        1      40
               5               0               0            5        1      40
               4               0               0            4        1     200
               3               0               0            3        1      80
               2               0               0            2        1      20
               1               0               0            1        1      80
               0           63206         3278336            0      159     160
```

Once the DSLAM is configured with L2 CoS to map the E320 priority markings into local priority queues, then the video flows—unicast or multicast—are sent via the DSL port. Internet packets are buffered—and possibly dropped—during periods of congestion. When handling a combination of unicast and multicast traffic, the DSLAM must have a robust scheduler capable of handling high data rates. The following DSLAM queue statistics show the prioritization of video over best-effort, and the DSLAM will drop best-effort first during periods of congestion.

```
officer SEC>> show interface 11.2 queuecount
 --- Egress Queue Statistics ---
                                                                       Max
                   Dropped                                      Queue  Queue
Interface       Q  Packets          Sent Packets    PBit        Depth  Depth
--------------  -  --------------   --------------  ---------- ------ ------
11.2.0          7               0                0           7      1     20
                6               0                0           6      1     40
                5               0                0           5      1     40
                4               0              144           4      1    200
                3               0                0           3      1     80
                2               0             3604           2      1     20
                1               0                0           1      1     80
                0            1235              646           0    160    160
```

With the DSLAM providing bandwidth management per DSL port, each IGMP join from the home reduces the amount of unicast bandwidth delivered into the home, as shown in Figure 8. However, the unicast bandwidth value may remain the same between the BSR and DSLAM if TCP congestion management is not applied.

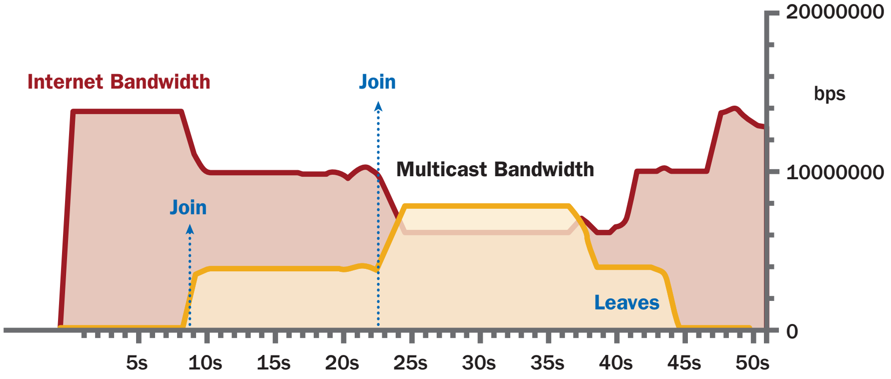

Figure 8. DSL Port Bandwidth Utilization

Summary

The IGMP Proxy model distributes all per-subscriber IGMP processing to the AN and uses report suppression so that the BSR only detects a single join per multicast group to be replicated over the M-VLAN. Since the AN hides the per-subscriber replication state from the BSR, the BSR will continue to forward unicast traffic at the full C-VLAN shaping rate, requiring priority queues in the AN.

Advantages of the IGMP proxy model:

- Multicast data plane optimization with multicast replication in the AN.
- Simplified IGMP control plane scaling requirements in the BSR.

Disadvantages o the IGMP proxy model:

- The AN must be enhanced to include:
  - – IGMP Proxy features with report suppression.
  - – Per-DSL port priority QoS due to local multicast insertion.
- The BSR cannot provide per-subscriber accounting.
- The BSR cannot apply back pressure to the C-VLAN shaper due to multicast join state.

About Juniper Networks

Juniper Networks, Inc. is the leader in high-performance networking. Juniper offers a high-performance network infrastructure that creates a responsive and trusted environment for accelerating the deployment of services and applications over a single network. This fuels high-performance businesses. Additional information can be found at www.juniper.net.

CORPORATE HEADQUARTERS
AND SALES HEADQUARTERS FOR
NORTH AND SOUTH AMERICA
Juniper Networks, Inc.
1194 North Mathilda Avenue
Sunnyvale, CA 94089 USA
Phone: 888.JUNIPER (888.586.4737)
or 408.745.2000
Fax: 408.745.2100
www.juniper.net

EUROPE, MIDDLE EAST, AFRICA
REGIONAL SALES HEADQUARTERS
Juniper Networks (UK) Limited
Building 1
Aviator Park
Station Road
Addlestone
Surrey, KT15 2PG, U.K.
Phone: 44.(0).1372.385500
Fax: 44.(0).1372.385501

EAST COAST OFFICE
Juniper Networks, Inc.
10 Technology Park Drive
Westford, MA 01886-3146 USA
Phone: 978.589.5800
Fax: 978.589.0800

ASIA PACIFIC REGIONAL SALES HEADQUARTERS
Juniper Networks (Hong Kong) Ltd.
Suite 2507-11, 25/F
ICBC Tower
Citibank Plaza, 3 Garden Road
Central, Hong Kong
Phone: 852.2332.3636
Fax: 852.2574.7803