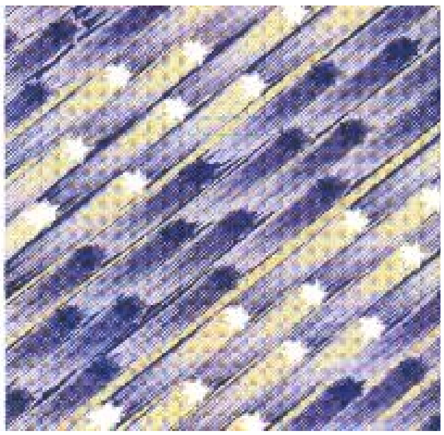


Enabling Multicast IP with MSS

September 10, 1999

Renuke Mendis

IBM Networking Hardware Division
Product Installation Support
Cary, North Carolina

Date Last Changed: September 10, 1999

Abstract

Multicast protocols are used to efficiently distribute information concurrently among multiple hosts. The use of multicast protocols is becoming popular with the increasing availability of multicast enabled applications such as multimedia, database and software distribution, and video conferencing. This paper discusses the architecture and flow of IP multicast traffic, together with information on designing and operating multicast enabled ATM networks, deployed using the IBM Multiprotocol Switch Services (MSS) Server.

(number of pages 9)

Table of Contents

Preface

Enabling multicast protocols on any network requires careful planning, design and management. The purpose of this whitepaper is to give the reader an understanding of the architecture and flow of IP multicast traffic, together with ideas to deploy multicast IP in ATM LANe and ATM CIP environments with minimal impact to the existing networks. Screen captures from the MSS server show detailed information and explanations of statistics related to multicast traffic and multicast routing protocols.

Keywords

MSS, Multicasting, Multicast IP, TCPIP, OSPF, DVMRP, MOSPF

Product List

IBM 8210 MSS, IBM 8265, IBM 8274, IBM 8271

For questions related to this document, please send an e-mail to:

Renuke Mendis
rhmendis@us.ibm.com
(919) 301-4161

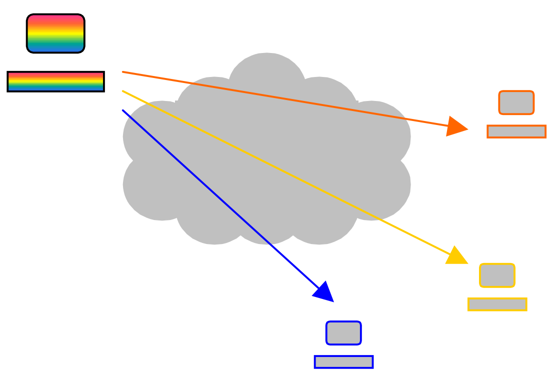
Chapter 1: Overview

Introduction

Multicast protocols are used to efficiently distribute information concurrently amongst multiple hosts. The use of multicast protocols is becoming popular with the increasing availability of multicast enabled applications such as multimedia and software distribution and video conferencing. This paper discusses the flow of multicast traffic, and provides customer proven solutions to enable multicasting in ATM networks, deployed using the IBM Multiprotocol Switch Services (MSS) Server.

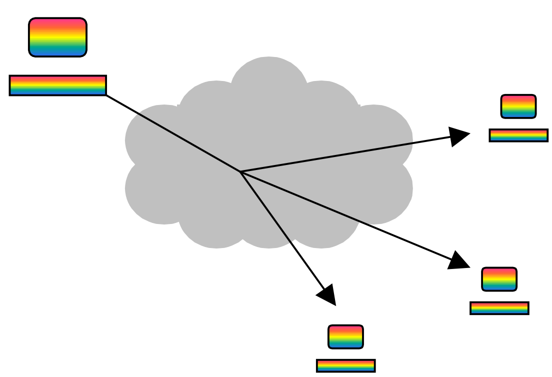
IP multicast traffic performs two functions -- delivery of information to multiple recipients and solicitation of network server resources by clients. Applications that deliver information to multiple clients include software, news and mail distribution systems, video distribution and interactive conferencing. Routers may utilize multicast addresses to distribute routing tables over routing protocols amongst a group of routers.

IP Multicast and Bandwidth Savings



The use of multicast protocols enables efficient bandwidth utilization in data networks. For example, a video distribution server may transmits three video streams of the same movie to each of the recipients using unicast protocols (TCP).

Conversely a single video stream carried over multicast protocols (UDP), are received simultaneously by each of the workstation recipients. As evident in this simple example, the bandwidth savings are exponential in relation to the number of destination workstations, receiving the multicast stream.



IP Multicast Addressing

A *multicast group address* is the combination of the four high order bits 1110 and the *multicast group id* which fall into the Class D IP address range from 224.0.0.0 to 239.255.255.255 in dotted decimal format. Certain multicast addresses have special meaning, indicating that these are well known addresses as defined by the Internet Assigned Numbers Authority (IANA). 224.0.0.1 indicates *all systems on this subnet*, while 224.0.1.1 is for *Network Time Protocol*, and 224.0.0.9 is used by RIP-2 for route information propagation.

Multicast addresses, like other IP addresses need to be resolved to Media Access Control (MAC) addresses for transmission over network media. The IANA specifies ethernet MAC addresses in the range of 01:00:5E:*00:00:00* to 01:00:5E:*7F:FF:FF* which are used for this purpose. The *Multicast group ID*, the 23 bit low order in the dotted decimal IP multicast address, is mapped over the *low order 23 bits* in the ethernet address. Since the upper 5 bits of the dotted decimal format are ignored (as they would otherwise overlap), 32 unique multicast addresses maybe mapped to the same ethernet address. Therefore, devices receiving multicast traffic have to filter these frames based on the IP address before accepting them.

Class D IP Address (31 bits)

```
                         Unused 5 bits in Multicast Group ID
                                     __|__
                                    |     |
                                              1                   2                   3
                            0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

                            1 1 1 0 x x x x  x x x x x x x x x x x x x x x x x x x x x x x x
                                             |_____________________________________________|
                                                                    |
                                              low order 23 bits of multicast group id
                                                     copied to ethernet address
                                              _______________________|_____________________
Ethernet Address (48 bits)                   |                                             |
                    1                   2                   3                   4
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8

0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 1 1 1 1 0 0 x x x x x x x x x x x x x x x x x x x x x x x
```

Flow of Multicast Traffic with LAN Emulation

In ATM forum LAN emulation Version 1, multicast IP frames are forwarded by the Broadcast and Unknown Server (BUS) residing on the MSS. For example, suppose an ATM attached multicast source sends a multicast frame over the multicast send VCC to the BUS, if Broadcast Manager for IP (BCM-IP) is not enabled, upon receiving these frames at the BUS they are forwarded over the multicast forward VCC, created and maintained by the BUS for traffic distribution to all LECs and proxy LECs simultaneously. If BCM-IP is enabled these broadcasts are forwarded to each known IP host within the eLAN as unicasts, thereby negating the benefits of multicasting.

In newer releases of MSS server, the intelligent LES/BUS option creates two separate multicast forward VCCs to distribute traffic to LECs and proxy LECs separately. In environments with multicast traffic, this increases the use of bandwidth of the ATM fabric as two copies of the same multicast (broadcast) stream are forwarded to all LECs, and all proxy LECs on the same eLAN simultaneously.

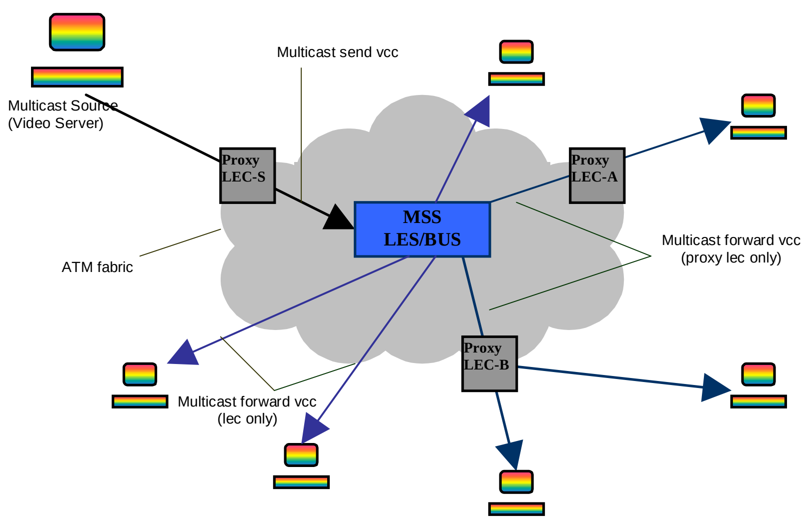


Figure 3.
An
example
of IP Multicast traffic - Video Distribution

In the above example, the source multicast stream from the video distribution server arrives at the BUS via the proxy-LEC's multicast send vcc. Subsequently, the BUS forwards this data stream to each LEC on the eLAN over the multicast forward vcc(s). Suppose workstations A and B were attached to a proxy-LEC-A and workstation C was attached to a proxy-LEC-B, and both proxy-LEC-A and B were attached to the same ATM switch only a single stream of this traffic would be forwarded by the BUS to the ATM switch, which in turn would be sent to each of the proxy LECs A and B over the multicast forward VCCs.

Proxy lecs with Layer 2 Switches

Proxy LECs based on layer 2 switching technology, forward frames based on destination media access control (MAC) addresses. Since the multicast stream arriving at the proxy LEC would have a destination MAC address denoting a multicast functional address, a layer two switch would ordinarily copy this frame to all ports. In the event that layer two switches are deployed the above example, if the multicast source sends data, all destination workstations in the eLAN receive the video traffic stream, even if they are not interested in receiving this data.

Proxy lecs with Layer 3 Switches

Layer 3 switches further improve multicast frame forwarding over the layer 2 technology by enabling the edge switches to differentiate between the destination workstations requesting to receive the multicast frames and those that don't. This is accomplished by the use of IGMP snooping functions that monitor for multicast requests from destinations seeking to join the multicast tree. Most layer 3 switches also support multicast routing protocols which enable multicast distribution to be moved to edge devices further optimizing the bandwidth use of the ATM fabric.

Flow of Multicast IP with Classical IP

In Classical IP environments, the absence of a Broadcast and Unknown Server (BUS) is overcome by the Multicast Address Resolution Server (MARS), a function ably handled by the MSS. When MARS functions are present the multicast source (in our example the video distribution server) sends a multicast registration to the MARS server. When a client workstation attempts to receive the multicast traffic, it queries the MARS server for the availability of that stream. If the multicast stream is available the client contacts the MARS server which in turn sends an add party request to have the ATM switch enabling the client to join the multicast tree built specifically for distributing that multicast stream. This is the most efficient distribution of multicast traffic, however the server and all clients are required to support Classical IP as well as MARS capabilities.

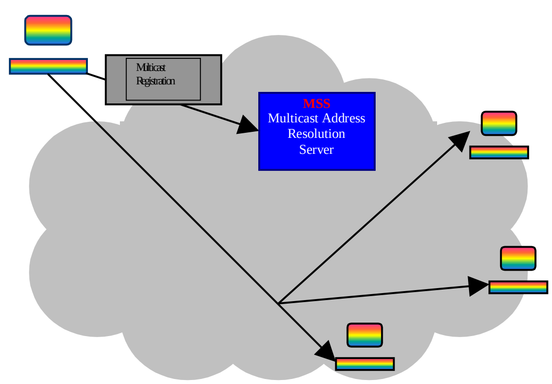
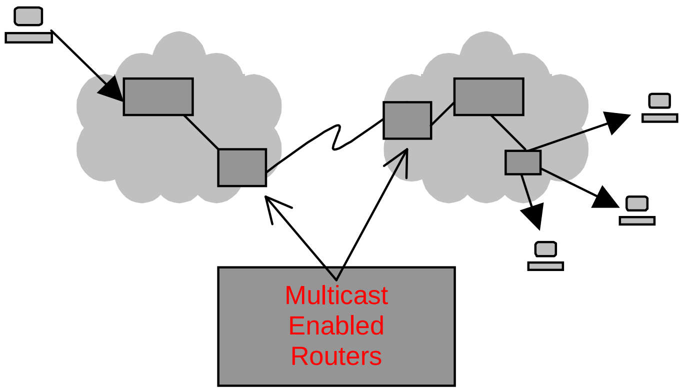


Figure 4.
Multicast Address Resolution Server with CIP

Routing Multicast Traffic

Within the same physical or logical ATM network, multicast traffic distribution is easy. However in some instances, a multicast stream originating in one physical or logical network may need to be received by destination workstations located on another physical or logical network. In such cases multicast streams must be routed via wide area ATM links or by external routers. Specific routing protocols are used to minimize the bandwidth used to distribute such multicast streams. MSS supports MOSPF and DVMRP protocols for multicast routing.

Figure 5.
Multicast Enabled Routers distribute the multicast streams between logical networks



MSS Implementation Options

MSS configuration options can make a significant difference in the performance of multicast traffic as well existing unicast traffic in LAN emulation and CIP/MARS environments. By definition, MARS offers the most efficient distribution of multicast frames in ATM environments. However, it may not be possible to implement MARS due to LANe and legacy network requirements.

As explained in the previous sections, the flow of Multicast traffic dramatically affects the use of the BUS. Therefore, enabling the BUS to operate in **VCC splice mode** will reduce the latency of the broadcast and multicast traffic distribution. This will automatically disable BCM functions, which as discussed above, causes multicast traffic to be distributed as non-unicast frames.

In environments with multiple MSS servers **offload BUS functions to at least one MSS running in VCC splice mode dedicated for a Multicast Elan**, with **no LAN emulation clients** enabled. If network management is required, a classical IP client may be configured, however LAN emulation clients or Host IP should not be configured on this MSS.

If multiple IP subnets are required, they may be setup on the same eLAN or on separate eLANs with an MSS server to route IP using DVMRP. Multicast OSPF may be used in place of DVMRP, but this may not be as efficient in terms of bandwidth use, as the broadcast and prune algorithm of DVMRP.

Setting up a separate eLAN to support Multicast traffic is also beneficial. Coupled with multicast capable edge devices with multicast VLAN support, this design option enables the most efficient use of bandwidth on the ATM backbone and the legacy edge networks.

Chapter 2: Detailed Design and Implementation

Installation Experience

Three IP Multicast implementation examples are presented in this section. The first example, from a large educational institution was

Example 1

Design Overview : LAN emulation, Routed IP/IPX, Super eLAN Bridging SNA/Netbios

Customer Profile : Large Educational Institution (3000 desktops)
Multicast Source : MBone, Video on Demand (VOD) Server
Hardware : IBM 8210 MSS, IBM 8265, IBM 8260, IBM 8274, Cisco 75xx, Proteon, SG Cray VOD server
Protocols: IP, IPX, SNA, Netbios
Environment : 8265/8260/8274 PNNI network (50 nodes), 8274 edge devices
Implementation Details: One MSS was setup as central LES/BUS in VCC splice mode with backup and redundancy on other MSSs, for five eLANs. This LES/BUS MSS was managed via a classical IP interface and no lecs were defined to minimize the traffic to this MSS's processor. Two MSSs were used as IP/IPX routers with DVMRP support to enable IP multicast routing between eLANs and to the outside world via Cisco and Proteon routers. One was MSS used as a Super eLAN Bridge to setup shortcuts for non-routed protocols (Netbios/SNA). IP/IPX filters & BBCM-NB was enabled on MSSs used for Super eLAN bridging, since those protocols were routed.
Installation experience: The key to this installation was the single LES/BUS MSSs setup in VCC splice mode, able to distribute a large number of multicast streams within the elans. Further, isolating the IP and IPX routing functions (including multicasr IP) to separate MSSs as well as bridging the non-routed protocols (netbios, SNA) helped the BUS handle the large number of packets for multicast.
Configuration: On the IBM 8210 MSS, VCC splice mode is configured on the LES/BUS panels by ELAN. To enable multicast routing is required, DVMRP is configured on the command line interface under *protocol DVMRP*. Alternatively, Multicast extensions to OSPF maybe configured on a box level as well as on the interfaces that need multicast capability.

On the 8274, default broadcast threshold per port is 192K of broadcasts. Since multicasts fall into this category, this flood limit must be disabled or raised to handle multiple multicast streams. This can be done on a per group basis using the *FLC* command and setting the flood limit to zero.

Example 2

Design Overview : Classical IP/MARS
Customer Profile :
Multicast Source : VOD Server
Hardware : IBM 8210 MSS, IBM 8265, classical IP clients with MARS support
Protocols : Classical IP
Environment : 8265 PNNI network
Implementation Details : The classical IP ARP server and MARS server were installed on the same MSS server. In an ATM network with heavy loads of multicast streams, these functions should be distributed amongst many MSSs. MSS also treats it's own CIP client interface as another router interface, thereby supporting MOSPF and DVMRP on this interface.

Example 3

Design Overview : LAN emulation & Classical IP/MARS with Layer 3 switches
Customer Profile : Lab Configuration
Multicast Source : IBM VideoCharger Server
Hardware : IBM 8210 MSS, IBM 8265, IBM 8274 GRS, IBM 8274, IBM 8271-712, IBM 8281, IBM 8239, IBM 8228
Protocols : Classical IP
Environment : 8265 PNNI network
Implementation Details : This hybrid network was setup to test interoperabilty between classical IP MARS and routed multicast IP. In this network all layer two devices handled multicast traffic similar example 1. However, by enabling DVMPRP on the layer 3 edge switch, the forwarding of multicast traffic was limited to the destinations that sent IGMP requests to receive that multicast stream. This helps the performance of the workstations, however this does take a toll on the edge switch.

Special Characteristics and Considerations

Physical and logical limitations of the above were minimized in the above implementations by distributing and isolating the LES/BUS functions and the routing and super eLAN bridging functions on separate MSSs. There are no rules to determine the number of MSSs needed to support a large network. For example, on an 10/100MB ethernet network of 1000 users, with requirements to support 5 multicast MPEG 2 video streams at 1.5Mbits/second per stream, it is possible to determine that with layer two edge devices, 10Mbit workstations on this network would be overwhelmed by the multicast traffic alone, leaving little bandwidth for other applications. It would be wise to isolate the multicast traffic in such a network by enabling a multicast ELANs / VLANs supporting the multicast streams and isolating the devices needing multicast access. This may not always be possible due to other requirements where VLAN technology on the edge device may be used to isolate the devices requiring multicast access. The IBM 8270, 8371, 8274 and 8274 GRS switches support this function today.

Appendix A: Detailed Configuration Information

This section describes MSS configuration functions discussed earlier in this paper to enable multicast operation in an existing LAN emulation environment. Changing the VCC splice mode enables more efficient operation of the broadcast and unknown server (BUS). In a flat network this would be the only change necessary to run multicast IP traffic.

MSS Configuration - Setting the BUS Mode to VCC-Splice Mode

*TALK 6

Config>NETWORK 0
ATM user configuration
ATM Config>LE-SERVICES
LAN Emulation Services user configuration
LE Services config>LES-BUS
  (1) <<< New LES/BUS >>>
  (2) Ethernet Multicast
  (3) Token Ring ELAN 1
  (4) Ethernet ELAN 1
Choice of LES/BUS [0]? 2
LES-BUS configuration
LES-BUS config for ELAN 'Ethernet Multicast'>**SET BUS-MODE**

BUS Mode
    (1) System
    (2) Adapter
    (3) Vcc-Splice
    (4) Exit (no-change)

Enter Selection: [1]? 3
Selection "Set BUS Mode" Complete
LES-BUS config for ELAN 'Ethernet Multicast'>**LIST**
LES-BUS Detailed Configuration
Name: Ethernet Multicast
  LES-BUS Enabled/Disabled: Enabled
  ATM Device number: 0
  End System Identifier (ESI):   40.00.82.10.00.00
  Selector Byte: 0x04
  ELAN Type: (S2)   Ethernet
  :
  :
  -LES-BUS Options-
  BUS Mode: Vcc-Splice   <- Bus mode @ vccsplice
  :
  :

MSS Configuration - Enabling OSPF to Forward Multicast - Box Level

*TALK 6
Gateway user configuration
Config>
Config>PROTOCOL OSPF
Open SPF-Based Routing Protocol configuration console
OSPF Config>LIST ALL

    --Global configuration--

```
          OSPF Protocol:        Enabled
          # AS ext. routes:     100
          Estimated # routers:    50
          Maximum LSA size:  :    2048
          External comparison:    Type 2
          RFC 1583 compatibility: Enabled
          AS boundary capability: Disabled
          Multicast forwarding:   Disabled          <- Mcast is disabled
          Demand Circuits:        Disabled
          Least Cost Area Ranges: Disabled
          Maximum Random LSA Age: 0

          --Area configuration--
Area ID        Stub? Default-cost Import-summaries?
0.0.0.0        No      N/A         N/A

               --Interface configuration--
IP address       Area          Auth   Cost  Rtrns  Delay  Pri  Hello   Dead
10.1.0.1       0.0.0.0          0      1     5      1      1    10     40
10.1.1.1       0.0.0.0         N/A     1     5      1      1    10     40
10.1.2.1       0.0.0.0          0      1     5      1      1    10     40
10.1.5.254     0.0.0.0          0      1     5      1      1    10     40

               --NBMA configuration--
          Interface Addr     Poll Interval
          10.1.0.1           120

            --Neighbor configuration--
Neighbor Addr     Interface Address  DR Eligible   Cost
10.1.0.16         10.1.0.1           Yes           N/A
OSPF Config>
OSPF Config>ENABLE MULTICAST
Inter-area multicasting enabled? [No]:

OSPF Config>LIST ALL

               --Global configuration--
          OSPF Protocol:        Enabled
          # AS ext. routes:     100
          Estimated # routers:    50
          Maximum LSA size:  :    2048
          External comparison:    Type 2
          RFC 1583 compatibility: Enabled
          AS boundary capability: Disabled
          Multicast forwarding:   Enabled           <- Mcast is enabled
          Inter-area multicast:   Disabled
          Demand Circuits:        Disabled
          Least Cost Area Ranges: Disabled
          Maximum Random LSA Age: 0

          --Area configuration--
Area ID        Stub? Default-cost Import-summaries?
0.0.0.0        No      N/A         N/A

               --Interface configuration--
IP address       Area          Auth   Cost  Rtrns  Delay  Pri  Hello   Dead
10.1.0.1       0.0.0.0          0      1     5      1      1    10     40
10.1.1.1       0.0.0.0         N/A     1     5      1      1    10     40
10.1.2.1       0.0.0.0          0      1     5      1      1    10     40
10.1.5.254     0.0.0.0          0      1     5      1      1    10     40

                 Multicast parameters
IP address      MCForward       DLUnicast
```

```
10.1.0.1        Off          Off                                  <- Mcast is off on
10.1.1.1        Off          Off                                     all interfaces
10.1.2.1        Off          Off
10.1.5.254      Off          Off

            --NBMA configuration--
        Interface Addr     Poll Interval
        10.1.0.1           120

           --Neighbor configuration--
Neighbor Addr    Interface Address  DR Eligible   Cost
10.1.0.16        10.1.0.1           Yes           N/A
OSPF Config>
```

MSS Configuration - Enabling OSPF Interfaces to Forward Multicast

```
OSPF Config>SET INTERFACE 10.1.1.1              <- repeat for each interface
Interface 10.1.1.1 already exists - record will be modified.
Attaches to area [0.0.0.0]?
Retransmission Interval (in seconds) [5]?
Transmission Delay (in seconds) [1]?
Router Priority [1]?
Hello Interval (in seconds) [10]?
Dead Router Interval (in seconds) [40]?
TOS 0 cost [1]?
Demand Circuit (Yes or No)? [No]:
Authentication Type (0 - None, 1 - Simple) [0]?
Forward multicast datagrams? [No]: Yes          <- Enables Multicast Forwarding
Forward as data-link unicasts? [No]:
OSPF Config>
```

MSS Configuration - Enabling DVMRP

```
* TALK 6
Config>PROTOCOL dvmrp
Distance Vector Multicast Routing Protocol config console
DVMRP Config>LIST all

DVMRP off
MOSPF 1 1
DVMRP Config>ENABLE DVMRP
DVMRP Enabled
DVMRP Config>LIST all

DVMRP on
MOSPF 1 1
DVMRP Config>
```

MSS Monitoring - Protocol DVMRP

```
* TALK 5
+PROTOCOL dvmrp
DVMRP console
DVMRP>?

Possible completions:
    DUMP routing tables
    EXIT
    INTERFACE summary
```

```
      JOIN
      LEAVE
      MCACHE
      MGROUPS
      MSTATS
(you may cycle through these commands by pressing the TAB key)

DVMRP>INTERFACE

Virtual Interface Table
 Vif  Local-Address                  Metric  Thresh  Flags
  0   10.1.0.1       subnet: 10.1.0.0       1       1    querier
  1   10.1.1.1       subnet: 10.1.1.0       1       1    querier
  2   10.1.2.1       subnet: 10.1.2.0       1       1    querier
  3   10.1.5.254     subnet: 10.1.5.0       1       1    querier

DVMRP>MCACHE

      0: ATM/0        1: NHRPL/0       2: TKR/0
      3: Eth/0        4: Eth/1         5: IPPN/0
      6: BDG/0        7: Internal

  Source        Destination     Count   Upst    Downstream
  10.1.2.0      224.100.100.100 3298    3       None
  10.1.2.0      224.200.200.200 2831    3       None

DVMRP>MGROUPS

          Local Group Database
Group           Interface       Lifetime (secs)
224.100.100.100  Eth/0               159
224.200.200.200  Eth/0               181

DVMRP>MSTATS

             MOSPF forwarding:           Disabled
             Inter-area forwarding:      Disabled
             DVMRP forwarding:           Enabled

  Datagrams fwd (unicast):       0  Locally delivered:            0
  Unreachable source:            0  Unallocated cache entries:    0
  Off multicast tree:            0  Unexpected DL multicast:      0
  Buffer alloc failure:          0  TTL scoping:                  0

  # DVMRP routing entries:       4  # DVMRP entries freed:        0
  # fwd cache alloc:             4  # fwd cache freed:            2
  # fwd cache GC:                0  # local group DB alloc:       2
  # local group DB free:         0

DVMRP>DUMP
             Multicast Routing Table
Type   Origin-Subnet  From-Gateway   Metric  Age  In  Out-Vifs
Direct 10.1.0.0       10.1.0.1          1     0   0   1* 2* 3*
Direct 10.1.1.0       10.1.1.1          1     0   1   0* 2* 3*
Direct 10.1.2.0       10.1.2.1          1     0   2   0* 1* 3*
      Group: 224.200.200.200  Prune sent lifetime: 6980   Graft send in progress: N
      Group: 224.100.100.100  Prune sent lifetime: 6980   Graft send in progress: N
Direct 10.1.5.0       10.1.5.254        1     0   3   0* 1* 2*
```

Appendix B: References

IBM Redbooks
TCP/IP Tutorial and Technical Overview, GG24-3376-05, November 1998, http://www.redbooks.com

Other published Books
The Protocols. TCP/IP Illustrated Volumes 1 & 2, W. Richard Stevens, 1994, Addison-Wesley, ISBN 0-201-63346-9 / ISBN 0-201-63354-X

Web Sites
IBM Redbooks : http://www.redbooks.ibm.com
Internet Engineering Task Force : http://www.ietf.org
Internet Assigned Numbers Authority : http://www.iana.org
RTP Information : http://www.cs.columbia.edu
RSVP Information : http://www.isi.edu/div7/rsvp/rsvp.html