

IP Multicast Network Setup

Enfinity Suite 6

IP Multicast Network Setup

Document ID: TP-LIB-65-01-01

Publication date 2011-03-17

These materials are subject to change without notice. These materials are provided by Intershop Communications AG and its affiliated companies ("Intershop Group") for informational purposes only, without representation or warranty of any kind, and Intershop Group shall not be liable for errors or omissions with respect to the materials. The only warranties for Intershop Group products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

This document and all of its parts are protected by copyright. All rights, including those of duplication, reproduction, translation, microfilming, storage on electronic media and processing in electronic form are expressly reserved.

Intershop® and Enfinity™ are trademarks or registered trademarks of Intershop Communications AG. All other company, product and brand names are trademarks or registered trademarks of their respective owners.

Copyright © 2005-2011 Intershop Communications. All Rights Reserved.

Table of Contents

Chapter 1: Introduction	4
About this Guide	5
Typographical Conventions	5
Chapter Overview	5
Preliminary Notes	5
Chapter 2: Network Setup	7
IP Group Management Protocol (IGMP)	8
Multicast Routing With IGMP	8
IGMP Snooping	8
Network Setup	9
Network Interface Selection	9
Reliable (Secure) Multicast	10
Mapping Multicast IP To MAC Addresses	10
Conclusion	10

Chapter 1

Introduction

About this Guide

As Enfinity Suite 6 uses IP multicast groups for some internal purposes on one hand, and for some facets IP multicast behaves much different than "normal" IP unicast traffic on the other, the lack of multicast related knowledge may lead to unintended results. This guide is intended to provide network administrators with some background insight of how multicast works in order for them to successfully use IP multicast traffic in a network with a non-trivial structure (which may be any network with more than one ethernet switch).

Currently only IPv4 traffic is supported, consequently this guide does not cover IPv6 related information. Furthermore, the routing of IP multicast beyond local network borders is outside the main scope of this guide and is only partially discussed, so everything mentioned on that must not be considered as comprehensive.

Typographical Conventions

The following typographical conventions are used throughout the guide:

■ Cross-references

References to other parts of this guide and to other documentation appear in *italics*.

■ Commands

All commands to be typed at command prompts appear in `courier` font.

■ Reserved or Special Words

Names of files, directories, or cartridges appear in *italics*.

This guide uses the `<Placeholder_Name>` syntax to denote the Enfinity Suite 6 installation directory (`<IS.INSTANCE.DIR>`), the location of the Shared File System (`<IS.INSTANCE.SHARE>`) as well as various other directory, group and user names, IDs etc. See the *Enfinity Suite 6 Installation Guide* for a complete list of placeholders.

Example code, attribute names, methods and database table names appear in Courier; for example, `init()`. In addition, the `#` sign refers to the number of an Enfinity Suite 6 instance.

Chapter Overview

The remainder of this chapter gives a general overview of IP multicast as an introduction to the topic.

Chapter 2 outlines the appropriate network setup approaches for IP multicast.

Preliminary Notes

Multicast communication is the network based communication of one or more senders with a given group of receivers. Its main implications, or design principles, include that none of the senders has to know the (current) set of

members of the multicast group(s) (set of receivers) and none of the receivers has to know any (possible) senders (e.g., for subscription purposes).

This seems trivial, but it makes multicast communication much different from the well-known unicast messaging. Therefore, network administrators must keep in mind some aspects when dealing with IP multicast:

- It is always good to clearly separate IP multicast from ethernet multicast because they refer to different levels of network communication. Especially for understanding multicast routing, there is a need to understand the ISO/OSI model and on which levels network devices act (and on which levels they passively work).
- IP Multicast is UDP based (it has to be, as the sender of some multicast datagram(s) is not intended to know the whole set of recipients). This directly implies that neither the order of packets nor the delivery for any client at all will and can be guaranteed. Both disordering and complete loss of packets can occur without further notification.
- No member of an IP multicast group is intended to have any knowledge about the set of members by means of the underlying transport protocol(s). This has the immediate implication that IP multicast routing can never be based on a recipient's address, but has to be done either (a) by some active instance (logically) outside the IP multicast group members or (b) by treating IP multicast traffic as broadcast - this is what low-cost switches (without IGMP snooping) do and what may flood VLANs with unwanted IP multicast traffic.
- As the projection of an IP multicast group address to an ethernet address is not unique, filtering out the multicast traffic for joined IP multicast groups can never be done exclusively by the network interface hardware (NIC), but has to occur (at least partially) at the OS kernel's network protocol stack. So the use of multicast protocols is going to cause the machine workload to increase.
- The selection of an IP multicast group management instance should be done based on the knowledge of both the underlying network's structure and the intended full set of members. That in turn is not and cannot be done at any of that group's members as none of them knows of the other ones - this choice cannot be exercised by automatic means at all.
- For the group management tasks, one additional, but simple protocol, IGMP, is needed. This protocol only defines 8 types of messages: requests and replies for creation, join, leave, and confirmation of a multicast group or group membership, respectively.
- As ethernet switches and hubs act on OSI/ISO layer 2 and IGMP is a layer 3 protocol, either they have to treat multicast datagrams as broadcast (and to replicate those datagrams to all ports in the VLAN, flooding all hosts with possibly unwanted traffic), or some higher level mechanism is needed to avoid the flooding. This is where IGMP snooping (see IGMP Snooping) comes into play.

Chapter 2

Network Setup

IP Group Management Protocol (IGMP)

For trivial network structures (all peers connected to the same switch), IP multicast routing is also trivial – either the switch treats it as a broadcast, or it will use the "IGMP snoop" mechanism (see IGMP Snooping), but without any need for periodic querying as it is still able to know all multicast group members. However, this is not the case if more than one switch is involved. Some management capability is needed for "intelligent" switches. IGMP provides this capability.

Multicast Routing With IGMP

With IGMP, there are 3 levels of conformance defined, and each network device may implement and conform to exactly one level:

- with IGMP v1 (the most commonly used level), a host may be disjoined from a multicast group after a certain period of inactivity (timeout),
- IGMP v2 implements an explicit **Leave**, so that no permanent activity (pinging the group) is needed any more,
- with IGMP v3 (rarely used) enables source based filtering at a multicast router.

For routing multicast traffic between different LANs, however, more protocols may be needed to avoid unintended flooding of network segments with multicast packets and notify the routers about the multicast group management events. For that purpose, Protocol Independent Multicast (PIM) is widely in use. With this respect, the IP multicast group addresses 224.0.0.1 (all-hosts; for the router(s) to periodically query for multicast groups) and 224.0.0.2 (all-routers; for hosts to asynchronously send reports to the router(s)) have a special meaning. Keep in mind that nearly the whole network range 224.0.0.0/24 consists of special addresses and is not recommended for regular multicast use.

NOTE: Be aware that (efficient) multicast routing is one of the topics that is still in active research.

IGMP defines the following message types (both as requests and reports):

- **Join** – connect a host to a multicast group. This may also mean to create the given group, if it does not already exist.
- **Leave** – disjoin a host from a multicast group. As a result, the group may be dissolved, if the leaving host was the last one of that group.
- **create** – have the multicast router create a new IP group and return the group's address in case of success, or zero otherwise.
- **confirm** – periodically confirm the group membership.

IGMP Snooping

For effectively avoiding the above mentioned multicast flooding, most modern switches implement what is called "IGMP snooping". With this technology,

switches are turned from regular (ISO/OSI) level 2 devices to level 2.5 devices. It works as outlined below.

Each network device (switch) maintains lists of known IP multicast group addresses and ports and listens to the IGMP traffic. Whenever an IGMP Join message occurs on any port, it adds that port to the list of ports for that group. Also, if an IGMP `Leave` message occurs, the affected port will be removed from the according list.

Although this works for one switch, an instance known as "IGMP Querier" is required in the LAN. This querier, usually located on a switch, periodically creates IGMP `confirm` messages, which each switch must forward to its connected hosts in the network segment. With the responses each switch reads, it can keep its multicast lists up to date. The "IGMP Querier" is, however, another critical active instance on which depends the entire IP multicast network.

As long as the packet TTL is 1 (default), multicast packets will be dropped on each router and thus remain local. Packet TTL counting is not relevant on switches as it is a level 3 mechanism and switches still act on level 2.

Network Setup

Network Interface Selection

In case a machine has only one network interface, it seems obvious that this is also the one for multicast use. Unfortunately, what is very common for desktop computers is not that common for server machines. There are two possibilities for an application: Either it sets a dedicated multicast interface or it has the operating system to decide this. As the operating system probably does not have enough information on which interface multicast traffic has to appear, it has the following options:

- select the first interface,
- select the first active (up) interface,
- select the first configured active interface,
- get an interface from the routing table,
- replicate the traffic to/from every interface (which is secure but cost intensive and produces unneeded network flooding).

There is no global definition of any common behavior for the choice of a multicast network interface when none is set. However, for BSD-like systems the following statement is generally approved:

"Berkeley-derived kernels choose the default interface for an outgoing multicast datagram by searching the normal IP routing table for a route to the destination multicast address, and the corresponding interface is used. This is the same technique used to choose the receiving interface if the process does not specify one when joining the group. The assumption is that if a route exists for a given multicast address (perhaps the default route in the routing table), then the

resulting interface should be used for input and output." (W. Richard Stevens: UNIX Network Programming, Vol. 1, Second Ed., p. 498)

However, the activation of additional network interfaces on a host may still influence the multicast routing if no default interface is set, the default interface also changes, or different routes for different traffic are needed. That is, in nearly all cases it is preferable to explicitly bind multicast traffic to a dedicated interface if more than one are available. Interface binding in Enfinity Suite 6 is done by specifying the interface's IP address in the appropriate property.

Reliable (Secure) Multicast

There are many approaches that intend to circumvent the inherent drawbacks of IP multicast traffic (e.g., no guaranteed delivery or packet order). None of them, however, seems to be widely accepted. Also, all of them seem to need active instances in the network. That in turn means that the reliability of the multicast traffic fully depends on the availability of those actors - hence the decision to take is to either rely on them or to (more or partially) prefer unicast traffic over multicast.

Mapping Multicast IP To MAC Addresses

Mapping the address of an IP multicast group to the according MAC address works as illustrated below. The most significant 25 bits of a multicast MAC address are defined as 01:00:5e and a 0 bit. The following 23 bits are taken from the least significant part of the IP multicast group address.

```

IP multicast address (decimal):      224 . 1 . 2 . 3
IP multicast address (binary):      11100000.00000001.00000010.00000011
                                     ----- | taken from group IP |
                                     | Prefix (25 bits) |
MAC (binary):                       00000001.00000000.01011110.00000001.00000010.00000011
MAC (hexadecimal):                  01 : 00 : 5e : 01 : 02 : 03

```

As the highest order half-byte of each multicast IP Address is hexadecimal "e", there are five relevant bits left out, making this projection a unique, but not bijective one. The consequence of that is that each operating system's IP stack will have to filter out the unwanted multicast traffic for up to 31 unwanted groups when it has set up its NIC(s) to receive only one.

On the other hand, this can be a favorable approach when selecting group addresses for hosts receiving more than one group (as, for example, Enfinity Suite 6 does). In these cases, it would obviously be better to choose IP multicast addresses that only vary in the highest order byte (and/or the most significant bit of the second byte) as these bits are not taken to build the multicast MAC address.

Conclusion

When using IP multicast, always be aware of the following aspects:

- Try to keep the network structure simple (single switch), as this is the easiest and most stable scenario.
- If this is not possible, activate the IGMP Querier functionality on exactly one switch device.

- As the correct function of the IGMP querier is essential for the multicast network, we recommend to monitor IGMP Confirm traffic.
- Bind the multicast traffic to a dedicated network interface.
- When choosing multiple IP multicast addresses for the same host, ensure to vary the first byte only and/or the most significant bit of the second byte of the address and keep the rest identical.