

CHAPTER 5

MULTICAST PRINCIPLES AND PROTOCOLS

This chapter introduces the basic of multicast principle and the protocol to find the appropriate path and forward a multicast packet to the recipients. These theories have been deployed to provide multicast communication between v4 and v6 world. This chapter is organized as follows:

The multicast discipline is introduced in section 5.1. The structure and usage of IP multicast are presented in section 5.2. Multicast routing protocols and forwarding algorithm are presented in section 5.3. Session management protocols for multicast application service are explained in section 5.4. The overview of transport protocol for multicast service is introduced in section 5.5. The summary of the chapter is in section 5.6.

5.1 Introduction to multicast discipline

The multicasting has several disciplines defined for this communication. We focus on the relevant protocols used in our enhancement and cooperate with NAT-PT to provide multicast application between v4 and v6 worlds. They are basic components used to provide normal multicast application. Each protocol is presented in the overview there are as follows:

1. Multicast IP address
2. Routing and forwarding protocol
3. Session management protocol
4. Transport protocol

5.2 Multicast IP address

Multicast is a communication between a single sender and several receivers. Multicast IP address is used to be a group number of multicast application and it was defined as a kind of address in both IPv4 and IPv6 protocols. The group identifies the set of recipients. Any nodes inside or outside network of the sender can join that group and use application at any time. Multicast application and group address are dynamic using. A multicast application can be

identified with an arbitrary group address. A used group address can be assigned to the other multicast applications at the same time. Several disciplines were created to define IPv4 and IPv6 multicast addresses assignment and usage.

a) **Multicast IPv4 address**

Multicast IPv4 addresses are the block of addresses in Class D. The addresses are in the range of 224.0.0.0 to 239.255.255.255. The first four bits of address are 1110 following with multicast group identifier in binary number as in Figure 5.1.

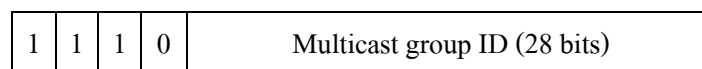


Figure 5.1 IPv4 multicast address

Multicast address assignment is defined in several RFCs in order to classify these addresses for using in different ways. The addresses were classified in two main types: well-known and transient addresses. Well-known address is used to be group address for a particular protocol, network devices to communicate to each other, and, communication between network systems in multicasting. Transient address is used to be group address for any multicast application which assigned by user.

Scope boundaries threshold only TTL value. A multicast packet with TTL value $\geq N$ allowed onto outgoing link. Each router decreases the TTL field when forwarding a datagram. When the TTL value reaches zero the datagram is dropped.

b) **Multicast IPv6 address**

Multicast IPv6 addresses are defined with multicast IPv6 prefix `ff00::/8`. They have the flags (4 bits) and scope (4 bits) fields to identify operation on the address. The following 112 bits can be an arbitrary number which identified group ID. The bit remainder can provide very huge number of IPv6 groups. The structure of multicast IPv6 address defined in RFC2373 [36] is presented in Figure 5.2. Like multicast IPv4 addresses, multicast IPv6 addresses were defined in two types: well-known and transient addresses. The standards such as RFC2373, RFC3306 [37], RFC3307 [38] were defined multicast IPv6 address assignment for address usage in different ways.

FP (8 bits)	Flags (4 bits)	Scope (4 bits)	Group ID (112 bits)
----------------	-------------------	-------------------	------------------------

Figure 5.2 Structure of IPv6 multicast address

Multicast IPv6 address has a new format that includes unicast prefix information as a part of the multicast address. It is unicast-prefix-based IPv6 multicast defined in RFC3306. The structure of this address is in Figure 5.3. The use of this address format allows any network to assign its own multicast address which avoids address collision from the other.

FP (8 bits)	Flags (4 bits)	Scope (4 bits)	Reserved (8 bits)	Plen (8 bits)	Network prefix (64 bits)	Group ID (32 bits)
----------------	-------------------	-------------------	----------------------	------------------	-----------------------------	-----------------------

Figure 5.3 Structure of unicast-prefix-based IPv6 multicast

FP(Format Prefix): specifies type of IPv6 address. The value for multicast IPv6 address is represented by number 1111 1111.

Flags: used to represent that whether network prefix enable. The flags field is a set of 4 flags: | 0 | 0 | P | T |

P = 0 indicates a multicast address that is not assigned based on the network prefix.

P = 1 indicates a multicast address that is assigned based on the network prefix.

If P = 1, T = 1 dynamically allocated (a transient address). Otherwise, T = 0, it is assigned by IANA.

Scope: it is used to limit the scope of multicast group. The values defined in RFC2373 are:

- | | |
|--------------------|----------------------------|
| 0 reserved | 8 organization-local scope |
| 1 node-local scope | E global scope |
| 2 link-local scope | F reserved |
| 5 site-local scope | |

Other values currently unassigned.

Reserved: reserved for the other works. It is always zero.

Plen: indicates the actual number of bits in the network prefix field which identify the

subnet when $P=1$.

Network prefix: indicates the network prefix of the unicast subnet creating the multicast address.

Group ID: is can be an arbitrary number like the original multicast IPv6 address format. It can be set based on the guideline in RFC3306 as a permanent address for network connection.

An example of unicast-prefix-based IPv6 multicast address usage is presented as follow: A network with unicast prefix `3ffe:1234:5678:9000::/64`. By using unicast-prefix-based, the network has multicast prefix `ff3X:0040:3ffe:1234:5678:9000::/96` (where X is any valid scope).

5.3 Routing and forwarding protocol

In unicast routing, traffic is routed to a single path from a source to destination. A unicast router does not need to care about source address – it cares about destination address and how to forward packet to destination network.

In multicast routing, when a sender provides a multicast application, it must send the data packet to every member node wherever they are by group number represented to that group. An original packet is not sent to a destination network but it must be duplicated in order to be distributed to every network that belongs to the members. Multicast router must determine and prevent that the forwarded multicast packet must not be sent back to the source. Otherwise multicast packet forwarding makes the packet looping problem which overloading bandwidth of networks.

In order to avoid the risk, multicast routing protocol was defined in multicast discipline. It is used to arrange for multicast packets to pass through all router needed to reach recipients. Multicast routing protocol uses source address in addition to destination address to find appropriate paths for sending.

The following section presents overview of multicast routing protocols that have been deployed in this work.

5.3.1 Reverse Path Forwarding (RPF)

RPF [39] is multicast forwarding algorithm used by a multicast router in order to forward multicast packets. With RPF, a multicast router must check the direction to forward multicast packets to network of member hosts (downstream). And it checks the direction toward the source (upstream) in order to avoid packet looping problem.

The router uses source address to check which direction is upstream (toward to source) and which direction is downstream (toward to destination). RPF uses the existing unicast routing table with source address of a multicast packet to determine the upstream and downstream neighbors. The router forwards a multicast packet only if it receives on the upstream interface. If there are multiple downstream paths, the router duplicates the packet and forwards the traffic down the suitable downstream paths – which is not essential to be all paths.

A multicast router performs RPF check on a packet, when the packet arrives at the router. If the RPF checking is successful, the packet is processed (which includes forwarding and anything else that is required). Otherwise, it is dropped. The RPF check procedure for traffic flowing down works as follows:

1. Router looks up the source address in the unicast routing table to check whether it has arrived on the interface that is on the reverse path back to the source.
2. If packet has arrived on the interface leading back to the source, the RPF check is successful and the packet is forwarded.
3. If the RPF check in step 2 fails, the packet is dropped.

Examples of RPF check fails and succeeds are presented in

Figure 5.1 and Figure 5.2 respectively.

Multicast routing table	
Network	Interface
66.33.0.0/16	S0
100.50.25.0/24	S1
191.10.20.0/24	E0

Packet arrived on wrong interface.
Discard packet

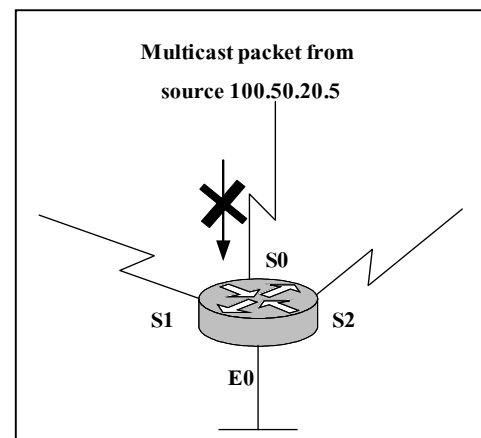


Figure 5.1 RPF check fails

A multicast packet from source 100.50.25.5 is received on interface S0. A check of the unicast route table shows that the interface that this router would use to forward unicast data to 100.50.25.5 is S1. Since the packet has arrived on S0, the packet is discarded.

Multicast routing table	
Network	Interface
66.33.0.0/16	S0
100.50.25.0/24	S1
191.10.20.0/24	E0

Packet arrived on S0
correct interface.

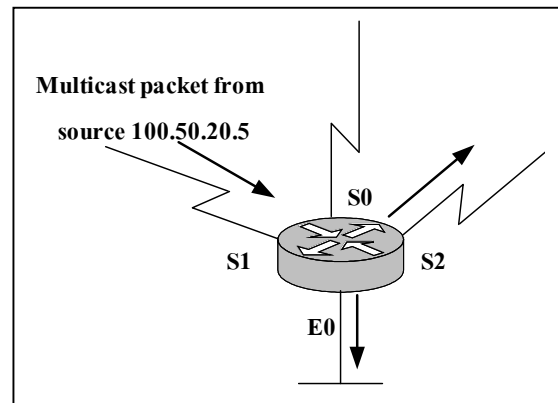


Figure 5.2 RPF check succeeds

This time the multicast packet has arrived on S1. The router checks the unicast routing table and finds that S1 is the correct interface. The RPF check passes and the packet is forwarded.

5.3.2 Distance Vector Multicast (DVMRP)

DVMRP [40] is designed to be used as an interior gateway protocol (IGP) within a multicast domain. It relies on source –based multicast delivery trees using RPF algorithm. The original specifications were derived from the Routing Information Protocol (RIP) and deployed the Truncated Reverse-Path Broadcasting (TRPB) technique. However, a major difference between RIP and DVMRP is that RIP calculates the next hop toward a destination, while DVMRP computes the previous hop back toward a source.

When a DVMRP router gets a multicast packet, RPF is employed to check the appropriate paths on the router’s interfaces for sending multicast packet to downstream member’s networks. If no group members on its directly attached leaf subnets, leaf routers transmit prune messages back toward the source. These prune messages will cause upstream routers to remove all branches that do not lead to group members from the tree. The resulting in this way is a source-based, shortest-path tree.

In order to create paths to the member networks to join a new group, DVMRP also implements a mechanism to graft back a previously pruned branch of a group’s delivery tree.

5.3.3 Protocol Independent Multicast (PIM)

PIM is a multicast routing protocol that uses unicast routing table to forward multicast packets. PIM does not depend on a particular unicast routing protocol so it is called protocol independent mode. It uses unicast routing information to perform multicast forwarding function by using RPF (Reverse Path Forwarding) to check function instead of building up a completely independent multicast routing table. PIM has two types of services: PIM-Dense Mode (PIM-DM) [41] and PIM-Sparse Mode (PIM-SM) [42]-[43].

PIM-DM: uses a push model to flood multicast traffic to network. This method is designed on the assumption that recipients for any particular multicast group will be densely distributed throughout the network. It is assumed that most (or at least many) subnets in the network will want any available multicast packet. PIM-DM initially floods multicast traffic throughout the network. Routers that do not have any downstream neighbors prune back the unwanted traffic.

The flood and prune is the way that the routers accumulate their state information – by receiving the data stream. These data streams contain the source and group information so that downstream routers can create their multicast forwarding table. PIM-DM can support only source trees – source and group (S, G) entries. It cannot be used to create a shared distribution tree. PIM-DM has been applied in the following scenarios:

1. Senders and receivers are not far from to one another.
2. There are not many senders and many receivers.
3. The density of multicast traffic is high.
4. The stream of multicast traffic is constant.
5. It relies on push mode to flood multicast traffic.

PIM-SM: uses a poll model, which contrasts with PIM-DM, to deliver multicast traffic. This solution is designed on the assumption that recipients for any particular multicast group will be sparsely distributed throughout network. Only networks that have active recipients that have explicitly requested a data will be forwarded the traffic. PIM-SM uses a shared tree to distribute the information about active sources. The traffic can remain on the shared tree to optimize source distribution tree. The traffic starts to flow down the shared tree then routers along the path

determine whether there is a better path to the source. If a better, more direct path exists, the designated router will send a join message toward the source and then reroute the traffic along this path.

PIM-SM relies on the concept of the Rendezvous Point or RP, since it uses shared trees. The RP must be administratively configured in the network. Sources register with the RP then data is forwarded down the shared tree to the receivers. If the shared tree is not an optimal path between the source and the receiver, the router dynamically create a source tree and stop traffic from flowing down the shared tree. PIM-SM has been deployed in the following scenario:

1. There are few receivers in a group.
2. Senders and receivers are separated by WAN links.
3. The type of traffic is intermittent.
4. It relies on pull mode to deliver multicast traffic.

5.4 Session management protocol

Since multicast address is dynamic, it is not convenient to register multicast address into DNS server in order to provide address resolution for multicast applications because the server must update DNS records following the changing of group address. Multicast has several ways to present an available group address of application. There are two styles.

Firstly, normal communication of network devices and system is broadcasting message or information for setting and management network pass through physical media to all nodes inside a network. For application, session description or multicast information including group address can be distributed to all nodes in the same way as network system communication by session description and announcement protocol.

For example Session Description Protocol (SDP) [44] is used to contain session information and carried by the announcement protocol such as Session Announcement Protocol (SAP) [45]. Then every node can learn and know which applications are available from an announced message advertised by a router or a server.

Using protocol to announce multicast information is convenient to provide multicast service that is dynamic all time. Network devices can listen the advertisement automatically. A session description can be delivered to the participant networks as long as these networks are in the scope of that session.

Secondly, a general news which is published in the Internet and human communications. The information of services and applications are configured and distributed manually via e-mail, www, TV program, in a magazine or Short Message Service (SMS). A user must configure and enter a group address and multicast information from these announcements into a multicast application to access that multicast service. This style may be limited for several networks from the advertised scope of group.

This following section presents the session management protocol. It is normally used to describe session or an available application for network and user.

5.4.1 Session Description Protocol (SDP)

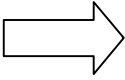
SDP [44] is used to describe information of session and relevant setup information of multimedia and conference. It advertises essential information for participation to recipients. Session description in SDP can be distributed to recipients in two ways. Firstly, it relies on other protocol, such as SAP [46]. This way is disseminated to specific scope or group which assigned in TTL and connection information. An administrator or announcer can advertise SDP message directly to a specific group of participants.

Secondly, SDP is distributed by email or www with the MIME content type “application/sdp”. This way can distribute session description to outside the scope that specific inside SDP packet and wider than SAP message.

The session description is a simple ASCII text message which can be transmitted as part of a UDP packet or as a part of a MIME message. A session description consists of a session-level description and optionally several media-level descriptions. SDP has session name and purpose, schedule to activate application, type of media and information (address and port) to receive the media. Other optional includes such item as the used bandwidth, contact information.

There are three classes of parameters: session description, time description and media description as shown in Figure 5.1. Only the protocol version, owner, session name, session time, and media name/transport address are required in session description; all others are optional.

The connection address in the part of session connection data can take on one of three forms. It can be a simple unicast or multicast IP address in which case it must include a TTL used as a scope indicator (appended to the IP address using slash as the separator). It can be an adjacent range of addresses. For example, a connection description of

`c = IN IP4 224.2.1.1/127/3` 

`c = IN IP4 224.2.1.1/127`

`c = IN IP4 224.2.1.2/127`

`c = IN IP4 224.2.1.3/127`

is similar to multiple “c” lines in a media description.

In media description parameter, only three transport protocols have been defined. They are RTP/AVP, or RTP using the audio/video profile carried over UDP, VAT (the Video Audio Tool packet format) carried over UDP and UDP alone. Possible items which can be included in the attribute lines of session and media description include: category, keywords, tool, packet time, conference type, character set, frame rate, and format specific parameter. An example SDP description from RFC 2327 is:

```

v=0
o=mhandley 2890844526 2890842807 IN IP4 126.16.64.4
s=SDP Seminar
i=A Seminar on the session description protocol
u=http://www.cs.ucl.ac.uk/staff/M.Handley/sdp.03.ps
e=mjh@isi.edu (Mark Handley)
c=IN IP4 224.2.17.12/127

```

Session description

v = protocol version

o = owner/creator

c = connection information (optional if include in all media)

s = session name

i = session information

u = URI of description

e = email address

p = phone number

c = connection information (optional if include in all session-level)

b = bandwidth information

One or more time descriptions

z = time zone adjustment

k = encryption key

a = attributes (zero or more session attribute lines)

Time description

t = time the session is active

r = repeat time (zero or more repeat time)

Media description

m = media name and transport address

i = media title

c = connection information (optional if include in all session-level)

b = bandwidth information

k = encryption key

a = attributes

Figure 5.1 Session Description Protocol entries

5.4.2 Session Announcement Protocol (SAP)

SAP [45]-[46] is an application protocol that used as a carrier to deliver a session description from an announcer to listeners in network. The structure of SAP header is presented in Figure 5.1. SAP announcer periodically advertises SAP packet to well-known multicast address and port. The time period between repetitive announcements is dependent on the scope of the session and the number of other sessions currently being announced by other session directory instances.

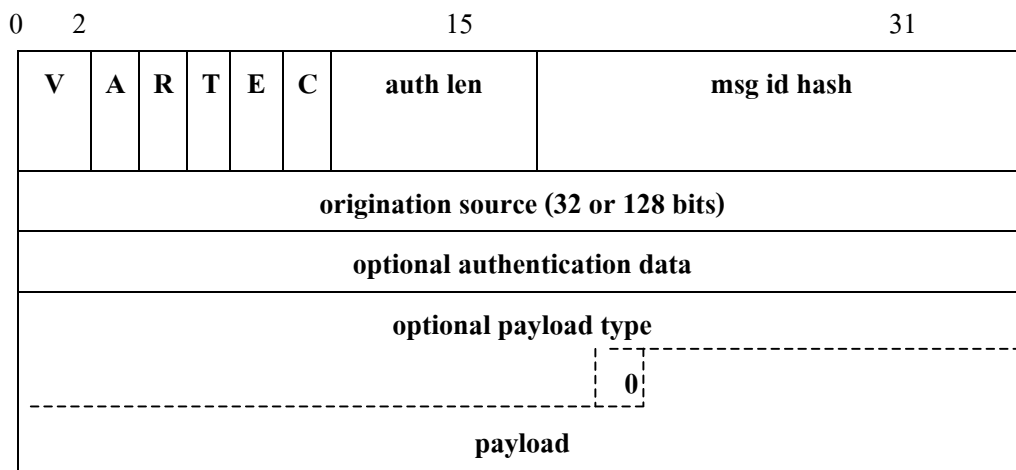


Figure 5.1 SAP packet format

V = version number: must be set to 1.

A = address type

R = reserved

T = message type: 0=announcement packet, 1= description deletion packet.

E = encryption bit: if set, means text payload is encrypted.

C = compressed bit: if set, payload is compressed using zlib compressed algorithm

auth len = authentication length: number of 32-bit words.

msg id hash = message identifier hash: globally unique ID for this announcement.

originating source = original source address

optional authentication data: digital signature of text payload.

payload: text payload

SAP was created for IPv4 and IPv6 protocols. The multicast well-known addresses 224.2.127.254 and ff0X:0:0:0:0:2:7ffe are assigned for SAP in IPv4 and IPv6 respectively (X in multicast IPv6 address represents valid scope number) with port number 9875. The detail of administrative scope IP multicast for IPv4 and IPv6 multicast address assignments for SAP are presented in RFC2365 [47] and RFC2375 [48].

For each administrative scope zone at a particular site, a session directory which running at that site needs to know following: the multicast address used for the announcement, the UDP port which announcements should be sent, the TTL, the address range used for session in this scope zone and the total bandwidth used by the session directory for session announcements in the administrative scope zone. The announcement is multicast with the same scope as the session it is announcing; the scope is checked by the group address range or the TTL. This checking keeps session announcements restricted to the location where the session is likely to take place. For example, a departmental conference does not have to be announced to the entire corporation.

SAP can be used to modify session description. A previously announced session can be modified by announcing the modified session description. However, in this case, the modification must be made by the authorized session only. It is recommended that the modified announcement contains an authentication header which is signed by the same key as the original session announcement. An alternative solution is that the original and modified announcements do not contain an authentication header but both originate from the same host.

SAP can also used to delete session description by sending a SAP session deletion packet with the verification procedures as used for session modification. Sessions can be deleted via explicit and implicit timeout. In the first case, the session description is deleted by the expired

timestamp which specific start and end time. In the second case, the session description does not include end time. It is expected that a session announcement message will be received periodically. If this periodic message is not received within a certain time, the announcement will be deleted.

SAP includes mechanisms for ensuring the integrity of session announcement (using the msg id hash) and encrypting announcement, it also supports authenticating the origin of an announcement. Session announcement can be encrypted with either a symmetric or an asymmetric algorithm.

If a session modification announcement is received that contains a valid authentication header but which is not signed by the original creator of the session then the session must be treated as a new session. This is in addition to the original session, unless the originator of one of the session descriptions can be authenticated using a digital certificate that has been signed by a trusted third party.

5.5 Transport protocol

Several of multicast applications are real time and interactive between members of that group. These services are usually relied on real time protocol to transport packet to interact with users. The real time protocol is called RTP.

Real-Time Protocol (RTP) and Real-Time Control Protocol (RTCP)

Multimedia applications such as streaming video, IP telephony, videoconference, teleconferencing seem to be announced daily. The service requirements of these application differ from the elastic application such as e-mail, web, FTP and Telnet. Several multimedia applications are highly delay-sensitive but loss-tolerant. There are rules to provide optimization between the two things.

A pair of protocols, RTP and RTCP [34], were designed to assist in the distribution of multimedia data on IP networks. They are transport protocol for unicast or multicast transmission. RTP consists of two closely linked parts: data and control.

1. The real-time transport protocol (RTP) is the data part. It is used to transmit packet with a real time characteristic. Multicast application such as multimedia applications, audio and video rely on RTP on top of UDP to deliver data packet. It is used to providing support applications with real-time properties such as continuous media,

including timing reconstruction, loss detection, security and content identification.

2. The real-time control transport protocol (RTCP) is used to monitor and control the same session of data packet for RTP. It supports source identification and gateway like audio and video bridges. It offers quality-of-service feedback from receivers to the multicast group as well as support for the synchronization of different media streams. RTP uses IP address to identify source and group of a particular application while RTCP contains the description and control message of that application. They work with together to control service and manage the multicast session. For an RTP session, there is a single multicast address and all RTP and RTCP packets belonging to the session use the multicast address. RTP and RTCP packets are distinguished from each other although the use of distinct port numbers. The RTCP port number is set to be equal to the RTP port number plus one.

5.6 Summary

When IP protocol was improved from IPv4 to IPv6 and IP multicast was created to be a kind of address in IPv4 and IPv6, the relevant multicast disciplines such as routing and forwarding protocol were also developed to support the different addresses. However, several parts of the existing and the new approaches of multicast disciplines have similar features and have the same basic tasks for multicast communication. They provide multicast communication independently for each IPv4 and IPv6 applications.