

Performance of IP Multicast in MPLS using PIM-SM (Protocol Independent Multicast-Sparse Mode)

Teerapat Sanguankotchakorn and Tanwa Kongkriengkrai

*Telecommunications Field of Study, School of Engineering and Technology, Asian Institute of Technology,
Pathumthani, Thailand, 12120*

Summary

From a technology perspective, the Internet performs a gigantic role in communication, compelling the Internet Protocol (IP) to be famous. The approach of modern applications requires Quality of Service (QoS) guarantee with multicast support being unavailable in the legacy network.

Multi Protocol Label Switching (MPLS), providing traffic engineering and QoS capability support in IP network, and Protocol Independent Multicast – Sparse Mode (PIM-SM), using two modes of multicast transmission to adapt with network requirement, are the leader in their firm. To combine these two technologies provides an opportunity for service provider to distribute new services from their contemporary network. This work is focusing on the fabrication of multicasting support framework using PIM-SM in MPLS domain. The simulation is used to verify the framework and illustrate its performance competence. Based on the simulation results, Shortest-Path Tree (SPT) multicast transmission offers the best performance, trading-off with the size of state database in each router. RP-rooted shared Tree (RPT) multicast achieves the suitable performance in the case that state database size needs to be concerned. Furthermore, bandwidth sharing problem can be eliminated using Connection-Oriented with QoS guarantee in MPLS network for the premium services.

Key words: *Quality of Service (QoS), MPLS, Protocol Independent Mode-Sparse Mode (PIM-SM)*

1. INTRODUCTION

1.1 Background

By the growth of bandwidth requirement and delay constraint in modern applications, Internet Protocol (IP), which is operated in the Internet with best-effort manner, cannot provide satisfactory services. There are several research challenges in solving these issues by providing Quality of Service (QoS) capability in IP network. Nowadays, Multiprotocol Label Switching (MPLS), Integrated Services (IntServ), and Differentiated Services (DiffServ) play the important role in the development of QoS guarantee provision in IP network.

MPLS has been developed to provide traffic engineering and QoS capability in IP network. MPLS uses switching element with a short, fixed-length, and locally significant identifier called “Label” to speed up packet forwarding process while retain the flexibility of an IP-based network. With connection-oriented scheme, end-to-

end path is established before packets are transmitted so that it is capable to assure resource along the path through the network.

On the other hand, only QoS provision is not enough for some applications such as news feeds, file distributions, interactive games and video conferencing, which are insisting on multicast capability to offer point-to-multipoint and multipoint-to-multipoint communication. Multicast is the capability of a communication network to accept a single message from sender and deliver copies of message to multiple receivers belonging to the same group at different locations. One of the challenges is to minimize the number of network resources employed by multicast.

In order to perform multicast communication, a tree connecting every node in a multicast group must be initiated. This problem of multicast routing in communication network is equivalent to find a tree T in graph G such that T spans all vertices in the multicast group M . Two types of tree are classified based on characteristic: source-specific (or source-rooted) and group-shared (or shared). Based on requirements of an application, many properties are relating to construct an optimized multicast tree based on parameters such as cost, delay, scalability, etc, to support dynamic multicast groups, survivability, and fairness.

In order to acquire an optimized tree, normally, the classical optimization problem in multicast routing to solve a minimum-cost multicast tree called Steiner tree is used. This problem is known as NP-complete problem [7]. Many approximation algorithms in Steiner Tree problem have been proposed. For example, KMB (refer to Kou, Markowsky, and Berman) algorithm [8], shortest-path tree (SPT), core-based tree (CBT), or reverse-shortest-path tree (RSPT) which provide the ease of implementation and efficient computation of the multicast tree.

Along the Internet, multicast is implemented by employing three types of protocols: (1) Internet Group Management Protocol (IGMP)[3] which is used at a host to join and leave from a multicast group, (2) Multicast Interior Gateway Protocol (MIGP) executed at multicast routers to enable multicast communication among an intranet domain. Some examples of MIGPs are Distance-Vector Multicast Routing Protocol (DVMRP) [16], Multicast Extensions for

Open Shortest Path First (MOSPF) [9], Core-Based Tree (CBT) [1], and Protocol Independent Multicast (PIM) [4], (3) Border Gateway Multicast Protocol (BGMP) [14] exploited by border router to allow multicast communication across autonomous system.

In order to confront the above mentioned problems, MPLS with multicast support is a challenging problem because of its QoS guarantee and speed up capabilities. Since the standard architecture of MPLS [12], proposed by MPLS working group in Internet Engineering Task Force (IETF), has an open issue in multicast support, many research works have been proposed in this area. For example, [11] proposed an overview of IP Multicast in a MPLS Environment providing a framework for IP multicast deployment in an MPLS network. In addition, there is another work which was proposed by [5] noticing how to use PIM to distribute MPLS labels for multicast routes.

To construct multicast MPLS framework using the existing multicast routing protocol, is advantageous. PIM is an attractive one adopting two types of tree: shared-tree and source-specific tree, to adapt according to the sender and receiver's requirements. Moreover, PIM does not depend on any unicast routing protocols, it gives more prospect to be implemented in an existing network.

PIM has two modes of operation: PIM Dense Mode (PIM-DM) handling an RSPT, and PIM Sparse Mode (PIM-SM) exerting unidirectional shared trees. PIM-DM is constructed based on an idea that a number of receivers are densely distributed in most branches of network. Since PIM-DM uses broadcasting, a lot of wasteful packets are transmitted to unnecessary branches. On the other hand, PIM-SM is fabricated to avoid the overhead of broadcasting packets when group members are sparsely distributed in the network. By using Rendezvous Point (RP) and explicit join message, PIM-SM leaves some works to receivers to join a shared-tree when they want. Furthermore, PIM-SM can switch over group-shared tree to source-specific tree if data rate exceeds a certain value to cope with sender/receiver requirement.

The main idea of this work is to construct a framework for multicast MPLS based on PIM supporting dynamically change between group-shared tree to source-specific tree.

2. General Concept and Preliminary

2.1 Multicast-supporting MPLS using PIM-SM framework

2.1.1 Framework specification

Prior to construct a framework for multicast-supporting MPLS using PIM-SM, some issues relating to the operation have to be described.

a) *Uncompatible issues between MPLS and PIM-SM*

While PIM-SM is developed to obtain the advantage over any types of unicast routing protocol with dynamically adapting to the alteration of network topology, however MPLS is concerning with QoS support over IP network, requiring more stable network topology and path for end-to-end resource reservation. Some modifications are needed for merging these two technologies.

The differences between PIM-SM and MPLS are summarized in table 2.1.

Table 2.1: The differences between MPLS and PIM-SM.

	MPLS [15]	PIM-SM [4]
Maintenance path mechanism	Hard state maintenance	Soft state maintenance
State refreshment	No periodic setup signaling used. Path status is refreshed by label request, label release, and label withdraw messages to create and destroy an LSP.	Periodic join and prune signal are used to maintain path status.
Unused path decontamination	No garbage collector required to clear the unused path.	Using periodic signal and garbage collector mechanism to clean up unused state.

To combine these two technologies, some specifications have to be defined in the proposed framework.

1. The hard state path maintenance is used to maintain LSP along the route of multicast tree. In MPLS, the construction of LSP is costly. But, to use soft state as in PIM-SM makes LSP frequently constructed and destroyed.

2. Periodic signaling is eliminated. Edge routers have a responsibility to convert PIM-SM signal to LDP message with multicast support. On edge routers, a set of mechanism is implemented to support the conversion as follows:

(a) Join message will be converted to Label-request message to construct an LSP from the edge router to the nearest branch of multicast tree of the group, where it wants to join.

(b) Prune message will be converted to Label-withdraw message to prune off unused path from the multicast group.

(c) Periodical join message, sent to refresh path status, is used to reset timer for LSP ageing maintenance.

(d) When the timer expired, path demolition will be initiated by the edge router.

(e) The RP has a responsibility to take care of path status from any sources to the RP and destroy them when they are not required.

3. To avoid the garbage collector mechanism requirement, some modifications are applied to PIM-SM in the transition procedure from RPT (RP-rooted Shared Tree) to SPT (Shortest Path Tree) to support hard state path maintenance. The details of the transition procedure are manifested in Section 2.3.

b) Layer 2 Characteristics

Although MPLS can be implemented over many types of layer 2 and layer 3 technologies, however, practically, IP is considered as only layer 3 protocol. In layer 2, a lot of technologies are supported with some limitations, such as Limited label space, Merging and TTL [10].

To avoid these problems, the native MPLS using shim header is considered in this framework.

c) IP Multicast LSP Triggers

In general MPLS framework, three types of triggers can be implemented [14].

- Request-driven: To intercept an explicit signal for LSP trigger.
- Topology-driven: To create LSP corresponding to the unicast routing topology.
- Traffic-driven: To create LSP when the ingress LSR receives a data packet.

With PIM-SM, Request-driven is more suitable than the others. PIM-SM has explicit signaling for join/prune multicast tree which can be intercepted by LSR for the LSP triggering.

d) Label Distribution and LSP Control

The proposed framework uses MPLS Downstream on Demand procedures with independent LSP control.

After intercepting a join message, an upstream LSR has to decide whether it requests for a label or not. The downstream on demand or upstream unsolicited can be used in this situation. The downstream allocation is the same mode as in unicast LDP which eliminates the need to develop upstream label distribution procedures. Consequently, downstream on demand procedures is exercised in this framework.

Independent LSP control is applied so that different downstream branches of a multicast distribution tree can independently join the tree. This matter prevents some branch having slow response to control message and affect the construction process of the others.

2.2 Proposed framework overview

An overview of the proposed framework is shown in Fig. 2.1. Only MPLS domain is considered in this

framework under the assumption that all LSRs know where RP is. By default, the receiver will firstly join a shared tree.

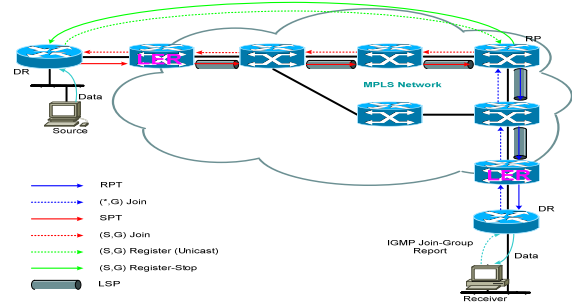


Fig. 2.1: Proposed framework overview of Multicast-supporting MPLS using PIM-SM

When the edge router received a Join-group message from outside MPLS domain, it sends a (*,G) Join message to the RP, where * means any address. All LSRs along the path to the RP create (*,G) state indicating that a shared-tree of group G is constructed. This tree is called RPT (RP-rooted shared tree), since the tree has a root at the RP. Simultaneously, LSP creation is trigger. An LSP is constructed from the ingress LSR, which intercepts the Join message, to the RP using LDP.

When a source S desires to send data to some receivers in a multicast group G, its DR (Designate Router) receives data packets from source and sends them in the RP-register message to the RP for that group. This action uses the same method as native IP network, since the RP-register message mechanism is used in a short time and a construction of LSP is costly. When RP receives the RP-register message, it decapsulates this message and sends the data along the RPT to all receivers. Utilizing packet transmission via sourced-specific tree is more advantageous than employing the RP-register messages, since encapsulation and decapsulation process are expensive for a router to perform. The RP sends back a (S,G) Join message to construct a source-specific tree rooted at S. This tree is called SPT (Shortest Path Tree), since it uses shortest path tree from source to the RP (RSPT in the case of asymmetric link). It will construct an LSP from the nearest edge router of S to the RP. All routers along the path will keep (S,G) state to indicate that the SPT rooted at S for a multicast group G is created. After the SPT is constructed, data packet will be transmitted along two paths: by applying the RP-register message and the SPT. As soon as data packet reaches the RP via the SPT, the RP will send a Register-stop message back to source S. This indicates source-designated router (DR), a router with the highest IP address in a subnet and is responsible to send Join/Prune messages to the RP, to stop sending data by the RP-register messages to the RP. At this time, data packet will be transmitted through the SPT from source S to the RP and to all receivers via RPT. Moreover, exerting the SPT of sender has more advantage in transiting state from the RPT to the SPT of receivers which will be described afterward.

When a receiver desires to leave from a multicast group G, it sends a leave-group report to its DR, and then the DR will remove that receiver from the multicast group. If there is no other member of the group in downstream direction from DR, it will send a (*,G) Prune message to the RP. In MPLS domain, this action, starting at the nearest edge router, will be repeated along the path to the RP until the Prune message reaches the RP or the router that has the other members within responsibility. Consequently, the LSP along the path which has been pruned is removed.

PIM-SM applies periodic Register-stop message and a Register-Suppression-timer to control the (S,G) state of each source. Since hard state LSP maintenance is applied, the RP will take a responsibility to take care of this timer. The LSP from source to the RP will be eradicated by the RP when timeout occurs. Outside the MPLS domain, (S,G) Join message transmission is periodically stopped, so it makes all (S,G) state along the path from source to edge router removed. If the source S desires to send data, it has to re-initiate the process employing RP-register message.

2.3 Transit procedure from RPT to SPT

In the native IP network, a receiver can switch from the RPT to the SPT after receiving packets from source over the RPT. The recommended policy is to initiate the switch-over to the SPT as soon as a source data rate is higher than a specified threshold. The overall procedure is illustrated in Fig. 2.2.

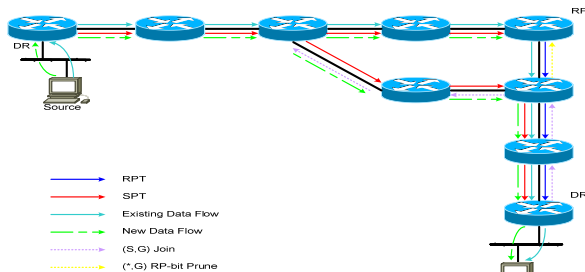


Fig.2.2: PIM-SM Switchover from RPT to SPT in IP network

Fig.2.2 delineates the switchover procedure from the RPT to the SPT in the proposed procedure. The transition is started from a receiver's DR which detects an over threshold of data rate from source S. The DR creates (S,G) state with SPT-bit reset to indicate that an SPT branch from S has not been completely set up. After that, the DR sends (S,G) Join message upstream to S along the SPT using the shortest path from unicast routing. This Join message has been sent until a router possessing (S,G) state received. Along the path to the ingress router, (S,G) state with SPT-bit cleared is set up to maintain the SPT branch. At the ingress router, LSP is set up along the SPT by triggering (S,G) Join message and SPT-bit of the multicast LSP entry is reset.

When the SPT branch is completely set up, data flow will be transmitted through the new path, SPT to receiver.

When routers, having (S,G) entry with the SPT-bit reset, start to receive packets from the source S along the new path, they set the SPT-bit to indicate that the SPT branch was completely set up. The route having (*,G) entry will send a (S,G) RP-bit Prune message toward the RP, indicating that it no longer desires to receive packets from S via the RPT. In MPLS domain, this mechanism is initiated at the nearest ingress router of the receiver where SPT-bit is reset. The (S,G) RP-bit Prune message only informs the RP to stop sending packet from source S along the RPT but the LSP along the RPT is still alive. It is done by keeping source S which has been pruned off into an entry called negative cache corresponding to its outgoing interface and label. This will inform LSRs to stop sending data along this path if the source of packet matches with one in the negative cache entry. Because the RPT maintained by (*,G) entry is still used by the other sources, it will be efficient to maintain the LSP along the RPT until that (*,G) Prune message is received.

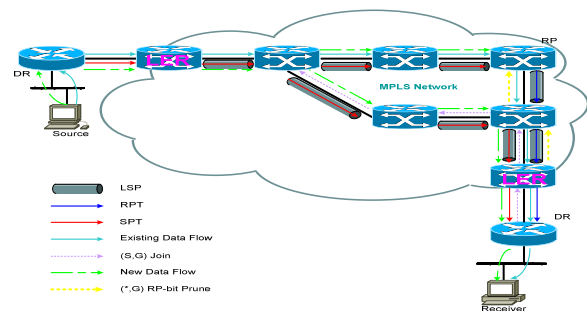


Fig.2.3: PIM-SM Switchover from RPT to SPT in MPLS network

2.4 Modification of MPLS Network Simulator

MPLS Network Simulator (MNS) is modified to simulate and verify the proposed framework. In this work, Network Simulator (NS) version 2.1b8a with MNS version 2.0 is used.

To make MNS supporting multicast feature with QoS guarantee, a set of functions have to be added and modified as follows:

- To support one-to-many label mapping, label mapping and corresponding table are customized.
- Join and Prune functions are added to construct the multicast tree in both RPT and SPT modes, including resource reservation support.
- The mechanism exploited to classify between multicast and unicast traffic is adapted in each MPLS supporting node.
- The capability to maintain the RP and switching over two modes of multicast is included, following PIM-SM architecture.

Some ideas in the simulation program are adopted and further extended from [2] which proposed a multicast simulator over MPLS enhancing the MNS on source-specific tree multicast transmission.

To manage information relating to LSP, three tables are operated: Partial Forwarding Table (PFT) and Label Information Base (LIB) table are used to perform label mapping, Explicit Routing information Based (ERB) is used to maintain the explicit routing information and priority of each flow in case that QoS guarantee is invoked. The native MNS uses these tables to carry out the MPLS packet switching mechanism [6].

In multicast-supporting MNS, a new table called LSG table (the Label for Source and Group table) is defined. This table maintains source and group addresses of multicast group corresponding to incoming label and incoming interface of data packet. Mapping multicast destination address with label has to be consulted with this table. Moreover, to support RPT multicast mode, SPT-bit entry is kept in this table to indicate the constructive status of multicast path. Negative cache entry is inserted into LIB table to inform that RPT path does not desire to receive data from the indicated sources. The insertion and deletion mechanism in ERB table are modified to contain the priority of multicast path. The modified module in the simulation program is highlighted in Fig.2.4.

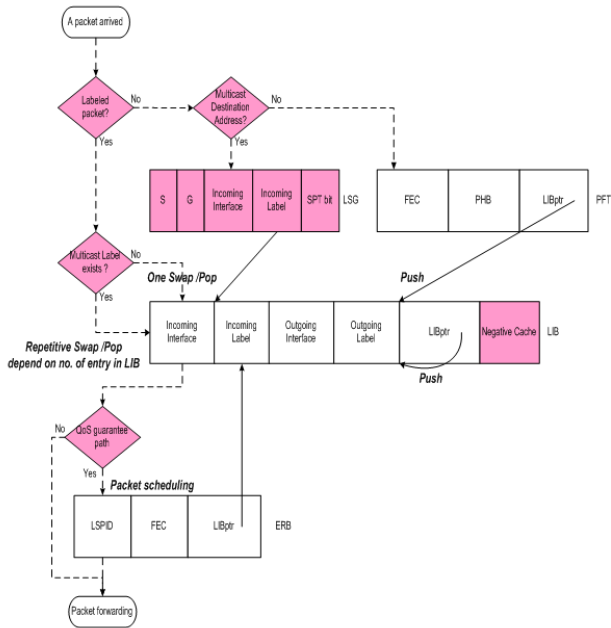


Fig. 2.4: Structure of table for multicast MPLS packet switching with QoS guarantee (Adapted from [6]).

In order to forward data packet in one-to-many manner, a replicator is inserted into MPLS address classifier in MPLS node to replicate new data packet and send it to the exact destination.

3. Simulation Model and Performance Evaluation

The objective of this simulation is to demonstrate the advantages of using multicast over unicast by comparing the utilization of PIM protocol in multicast among RPT, SPT and unicast.

3.1 Network Topology

The network topology used is created from program named BRITE - Boston University Representative Internet Topology generator [8]. A link, which is used to connect each pair of nodes, is randomly established by Waxman model [15]. Every link is constructed as duplex whose bandwidth and delay are assumed to be constant values, which equal 1 (unit) for both directions.

3.2 Traffic type and Parameters

Table 3.1: Traffic type and Parameters

Traffic type	Transmission type	Link speed (kbps)	Remarks
CBR regular traffic	Best-effort, unicast	120	CBR sources are generated and applied at both sides of every links to model the regular traffic in the network.
RPT and SPT	multicast	vary	RPT mode is applied from the beginning of transmission until the 6 th second of simulation. Then, it is altered to SPT mode until the end of transmission.

Note: All links in the network have bandwidth fixed at 150 kbps.

3.3 Performance Evaluation Matrices

3.3.1 Average Bandwidth Requirement per User

Assume that one hop count in the path requires a unit of bandwidth per connection in case of unicast and a unit of bandwidth per source in case of multicast.

Hence, in case of unicast, the average bandwidth requirement per user, $b_{avg}^{unicast}$ can be defined as

$$b_{avg}^{unicast} = \frac{\sum_{i=1}^S \sum_{j=1}^{R_i^*} b_{i,j}^{unicast}}{\sum_{i=1}^S R_i^*} \quad (3.1)$$

Where $b_{i,j}^{unicast}$ = Bandwidth required for transmission from the j^{th} receiver to the i^{th} source, R_i^* = Total number of receivers receiving data from the i^{th} source, S = Total number of sources.

For SPT multicast, the average bandwidth requirement per user, b_{avg}^{SPT} , can be defined as

$$b_{avg}^{SPT} = \frac{\sum_{i=1}^G \sum_{j=1}^{S_i} \sum_{k=1}^{R_i} b_{i,j,k}^{SPT}}{\sum_{i=1}^G (R_i \times S_i)} \quad (3.2)$$

Where $b_{i,j,k}^{SPT}$ = Bandwidth required for transmission from the j^{th} receiver to the nearest branch of SPT rooted at the k^{th} source, sending data to the i^{th} group, R_i = Total number of receivers in the i^{th} group, S_i = Total number of sources in the i^{th} group, G = Total number of groups.

Whereas the average bandwidth requirement per user of RPT multicast transmission, b_{avg}^{RPT} , is defined as

$$b_{avg}^{RPT} = \frac{\sum_{i=1}^G \left(\sum_{j=1}^{R_i} b_{i,j}^{RPT} + \sum_{k=1}^{S_i} b_{i,k}^{RP} \right)}{\sum_{i=1}^G (R_i \times S_i)} \quad (3.3)$$

Where $b_{i,j}^{RPT}$ = Bandwidth required for transmission from the j^{th} receiver to the nearest branch of RPT of the i^{th} group, $b_{i,k}^{RP}$ = Bandwidth required for transmission from RP to the k^{th} source, belonging to the i^{th} group, R_i = Total number of receivers in the i^{th} group, S_i = Total number of sources in the i^{th} group, G = Total number of groups.

3.3.2 Minimum Bandwidth Requirement in a Link

Assume that one traffic flow through a link creates a unit of bandwidth requirement in that link. Therefore, minimum bandwidth requirement in a link, B_{min} is the maximum number of traffic flow in the existing link.

$$B_{min} = \max \{B_{i,j} \mid i, j \in \{1, 2, \dots, N\}\} \quad (3.4)$$

Where $B_{i,j} = \begin{cases} N_{i,j}; & (i \neq j) \text{ and link } (i, j) \text{ exists} \\ 0; & \text{otherwise} \end{cases}$

$N_{i,j}$ = The number of traffic flows, transmitting through link (i, j) , N = Total number of nodes.

3.3.3 Average End-to-End Delay per Receiver

Assume that one hop count in the transmission link causes a unit of delay. The average end-to-end delay per receiver in case of unicast and SPT multicast transmission are equivalent, since they use the same path to transmit data to the receiver.

The average end-to-end delay per receiver of the unicast, $D_{avg}^{unicast}$ and SPT multicast, D_{avg}^{SPT} , is defined as

$$D_{avg}^{unicast} = D_{avg}^{SPT} = \frac{\sum_{i=1}^S \sum_{j=1}^{R_i} d_{i,j}^{unicast}}{\sum_{i=1}^S R_i^*} \quad (3.5)$$

Where $d_{i,j}^{unicast}$ = Delay from the j^{th} receiver to the i^{th} source, R_i^* = Total number of receivers receiving data from the i^{th} source, S = Total number of sources (unicast transmission) = Summation of total number of sources in each group (multicast transmission).

For RPT multicast, the average end-to-end delay per receiver, D_{avg}^{RPT} , can be defined as

$$D_{avg}^{RPT} = \frac{\sum_{i=1}^G \sum_{k=1}^{S_i} \sum_{j=1}^{R_i} (d_{i,j}^{RPT} + d_{i,k}^{RP})}{\sum_{i=1}^G (S_i \times G_i)} \quad (3.6)$$

Where $d_{i,j}^{RPT}$ = Delay from the j^{th} receiver belonging to the i^{th} group to the RP, $d_{i,k}^{RP}$ = Delay from the RP to the k^{th} source sending data to i^{th} group, R_i = Total number of receivers in the i^{th} group, S_i = Total number of sources in the i^{th} group, G = Total number of groups.

4. Simulation Results and Discussions

4.1 Average Bandwidth Requirement per receiver

4.1.1 Analysis

Let $X_{avg}^{unicast}$, X_{avg}^{RPT} and X_{avg}^{SPT} = Average distance from source to receiver in unicast route, RPT, and SPT multicast route, respectively. $B_{avg}^{unicast}$, B_{avg}^{RPT} and B_{avg}^{SPT} = Average bandwidth requirement per receiver of unicast route, RPT, and SPT multicast route, respectively.

The relationship between average distance from source to receiver of unicast, RPT, and SPT multicast route can be illustrated as

$$X_{avg}^{unicast} = X_{avg}^{SPT} \leq X_{avg}^{RPT} \quad (4.1)$$

The unicast transmission requires bandwidth in each link to be equal to the number of traffic flows traversing through the link, while the SPT multicast transmission, using the same route, consumes only one unit as shown in Fig.4.1.

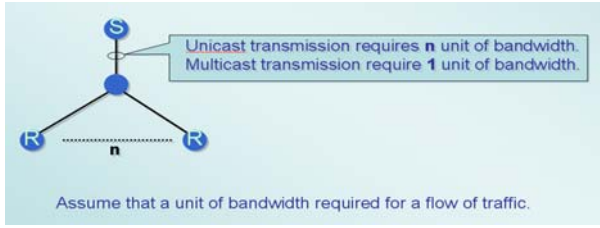


Fig.4.1: Bandwidth consumption in a link: unicast and SPT multicast

Since $X_{avg}^{unicast} = X_{avg}^{SPT}$, the relationship between average bandwidth requirement per receiver of the unicast and the SPT multicast can be written as:

$$B_{avg}^{unicast} \geq B_{avg}^{SPT} \quad (4.2)$$

The bandwidth requirement per receiver of RPT multicast is slightly larger than that of SPT and converges to the same point when the number of receivers in each group increases. Since, the average distance between each node is almost the same in this simulation model, as shown in Fig. 4.2. When the number of receivers is larger than the number of sources in a group, the bandwidth requirement per receiver in the case of RPT multicast mostly depends on the amount of traffic transmitted from the RP to receivers.

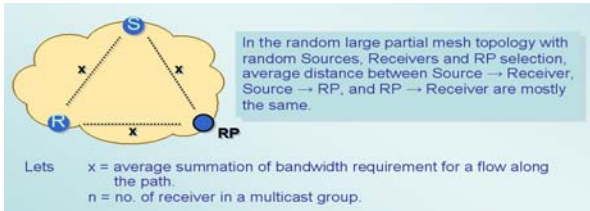


Fig.4.2: The relationship of distance and bandwidth requirement for a flow along the path between RPT and SPT multicast

From Fig.4.2, the relationship between average bandwidth requirement per receiver of the SPT and RPT multicast can be simply illustrated as follows:

$$B_{avg}^{SPT} \approx \frac{nX}{n} = X \quad (4.3)$$

$$B_{avg}^{RPT} \approx \frac{(nX + X)}{n} = X + \frac{X}{n} \quad (4.4)$$

$$\therefore B_{avg}^{SPT} \leq B_{avg}^{RPT} \quad (4.5)$$

$$\text{For large } n; B_{avg}^{SPT} \approx B_{avg}^{RPT} \quad (4.6)$$

4.1.2 Simulation Results and Discussion

The simulation is performed according to the parameters in each case as shown in table 4.1

Table 4.1: Simulation Cases and Parameters

Case	Parameters	
	Number of source nodes	Number of receivers
(a)	vary	20
(b)	100	vary

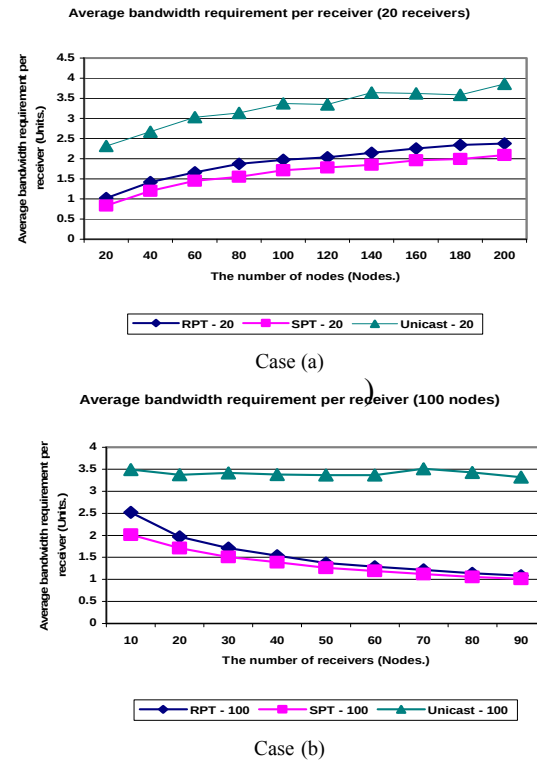


Fig.4.3: The simulation results of average bandwidth requirement per receiver

It is obvious from Fig.4.3 (a) that RPT multicast requires a little more bandwidth than SPT multicast. This is because of the increment of distance between nodes. In unicast, the bandwidth requirement per receiver is proportional to the average distance between two nodes, but irrelevant to the number of receivers. It can be concluded that the more the number of source nodes increases, the higher the average bandwidth requirement per receiver is.

As illustrated in Fig.4.3 (b), the RPT multicast uses slightly more bandwidth than SPT Multicast. According to Eq.4.3 to 4.6, bandwidth requirement of RPT and SPT multicast decreases until they converges to the certain level when the number of receivers increases. On the other hand, unicast needs more bandwidth per receiver than multicast, due to the inability of bandwidth sharing.

4.2 Minimum Bandwidth Requirement in a Link

4.2.1 Analysis

Let $B_{min}^{unicast}$, B_{min}^{RPT} and B_{min}^{SPT} = Minimum bandwidth requirement for a unicast, RPT and SPT multicast, respectively.

For the RPT multicast, the bottleneck occurs when many traffic from every source using the same RP are

transmitted through a link downstream through RP, as illustrated in Fig.4.4.

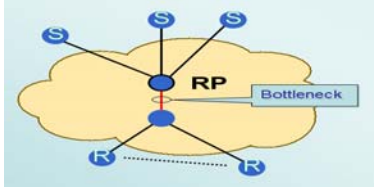


Fig.4.4: Bottleneck in case of using RPT

$B_{\min}^{RPT}$ can be defined as

$$B_{\min}^{RPT} = \sum_{i=1}^{g^{RPT}} S_i^{RPT} \quad (4.7)$$

where S_i^{RPT} = Number of sources in the i^{th} group sending data via RPT, g^{RPT} = Number of group using the same RP.

For the SPT multicast, the bottleneck occurs due to the shortest path trees rooted at each sources overlap to each other in the network, as illustrated in Fig.4.5.

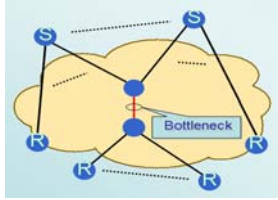


Fig.4.5: Bottleneck in case of using SPT

$B_{\min}^{SPT}$ can be defined as

$$B_{\min}^{SPT} = \sum_{i=1}^{g^{SPT}} S_i^{SPT} \quad (4.8)$$

where S_i^{SPT} = Number of source in the i^{th} group using an overlap link to receiver, g^{SPT} = Number of group using an overlap link to receiver.

According to $S_i^{SPT} \leq S_i^{RPT}$ and $g^{SPT} \leq g^{RPT}$, then

$$\therefore B_{\min}^{SPT} \leq B_{\min}^{RPT} \quad (4.9)$$

In case of unicast, minimum bandwidth requirement in a link cannot be simplified due to undistinguishable topology of real network including randomly positioning of sources and receivers. Therefore, only estimation of minimum bandwidth requirement in a link can be done, as shown in Fig.4.6.

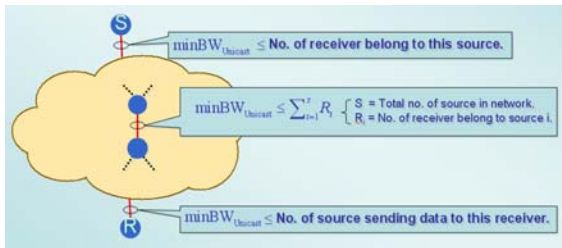


Fig.4.6: Estimation of unicast's minimum bandwidth requirement in a link

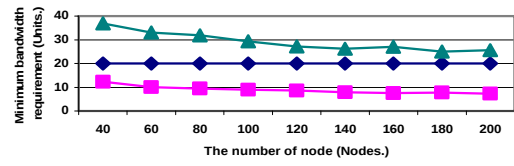
4.2.2 Simulation Results and Discussion

The simulation is performed according to the parameters in each case as shown in table 4.2.

Table 4.2 Simulation Cases and Parameters

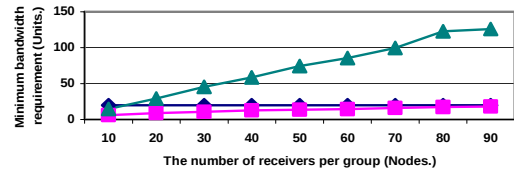
Case	Parameters		
	Number of source nodes	Number of receivers per group	Number of groups
(a)	vary	20	20
(b)	100	vary	20
(c)	200	40	vary

Minimum bandwidth requirement in a link (20 groups, 20 receiver)



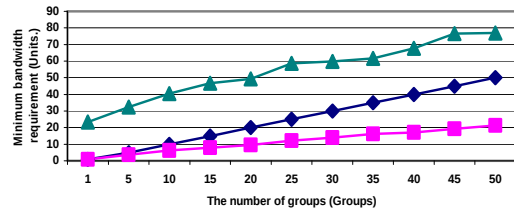
Case (a)

Minimum bandwidth requirement in a link (20 groups, 100 nodes)



Case (b)

Minimum bandwidth requirement in a link (200 nodes, 40 receivers/group)



Case (c)

Fig.4.7: The simulation results of minimum bandwidth requirement in a link.

In Fig.4.7 (a), when the number of source nodes grows up, the SPT multicast's minimum bandwidth requirement drops a little bit. On the contrary, RPT multicast is quite stable. The reason is that with fixed number of receivers, the more the number of source nodes is, the higher the

probability of the SPT route overlapping becomes. However, in case of RPT, all information has to be firstly transmitted to the RP, hence bottleneck occurs at the downstream link of RP. Consequently, the minimum bandwidth requirement is fixed at that bottleneck. In case of unicast, the minimum bandwidth requirement is greater than that in multicast because each receiver cannot share bandwidth together. This is obviously a disadvantage of unicast.

It is obvious from Fig.4.7 (b) that the minimum bandwidth requirement in RPT multicast is constant, since it depends only on the number of sources using the same RP. In SPT multicast, according to Eq. 4.7 to 4.9, the minimum bandwidth requirement is limited to the RPT. Consequently, it converges to that of RPT. For unicast transmission, the more the number of receiver is, the requirement of minimum bandwidth in a link proliferates. Since each receiver needs an individual connection, therefore the probability of link overlapping increases.

Fig.4.7 (c) indicates that among three cases, the minimum bandwidth requirement in a link of SPT multicast rises at a lowest rate. Since the number of connections from each source overlapping in the same link is equivalent or less than the total number of sources sharing RP in RPT as shown in Eq.4.7 to 4.9.

In case of unicast, the minimum bandwidth requirement in a link grows up when the number of groups increases. Since the increment of the number of groups leads to the increment of the number of receivers, the unicast transmission, which depends on the individual connection between each source and each receiver, is affected directly.

4.3 Average End-to-End Delay per Receiver

4.3.1 Analysis

As illustrated previously, unicast and SPT multicast have the same end-to-end delay, while RPT multicast has larger end-to-end delay per receiver. Since unicast and SPT use the same path to transmit data packet, while data packets need to visit the RP prior to be transmitted to receiver in RPT.

Let D_{E2E}^{SPT} and D_{E2E}^{RPT} =Average end-to-end delay from source to receiver using SPT and RPT route, respectively.

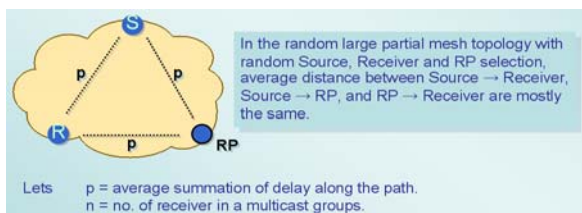


Fig.4.8: The relationship of distance and average summation of delay along the path between RPT and SPT multicast

As shown in Fig.4.8, the average summation of delay along the path between each node in the network is mostly the same. The relationship between D_{E2E}^{SPT} and D_{E2E}^{RPT} can be illustrated as follows:

$$D_{E2E}^{SPT} \approx \frac{np}{p} = p \quad (4.10)$$

$$D_{E2E}^{RPT} \approx \frac{n(p+p)}{n} = 2p \quad (4.11)$$

$$\therefore D_{E2E}^{RPT} \approx 2(D_{E2E}^{SPT}) \quad (4.12)$$

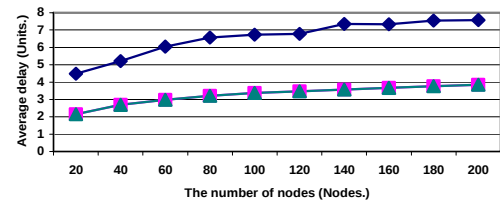
4.3.2 Simulation Results and Discussion

The simulation is performed according to the parameters in each case as shown in table 4.3.

Table 4.3: Simulation Cases and Parameters

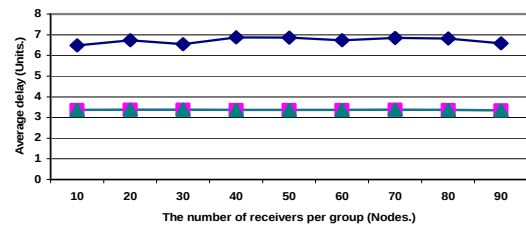
Case	Parameters		
	Number of source nodes	Number of receivers per group	Number of groups
(a)	vary	20	20
(b)	100	vary	20
(c)	100	40	vary

Average end-to-end delay per receiver (20 groups, 20 receivers)



Case (a)

Average end-to-end delay per receiver (100 nodes, 20 groups)



Case (b)

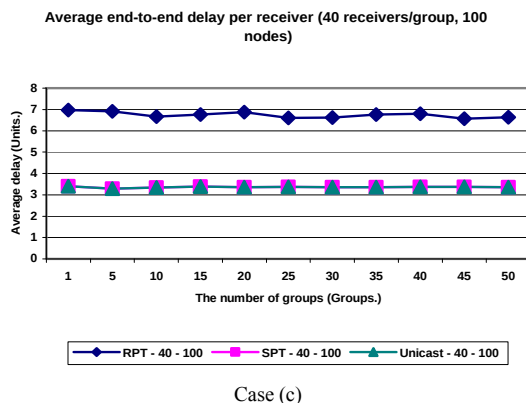


Fig.4.9: The simulation results of average end-to-end delay per receiver

Fig.4.9 (a) illustrates that the average end-to-end delay per receiver of RPT multicast is two times of unicast and SPT multicast. This result coincides with Eq.4.10 to 4.12. Since in RPT, the data has to be firstly transmitted to RP, then the time is wasted.

As shown in Fig.4.9 (b) and (c), an increment of both the number of receivers per group and the number of groups does not affect end-to-end delay, because the increasing delay results from the increasing receivers. Therefore, the average delay per receiver is constant.

5. Conclusions

In this work, PIM-SM in MPLS is proposed and the performance is studied. According to simulation results, it is obvious that unicast, comparing to multicast, consumes a larger bandwidth for point-to-multipoint data transmission. In addition, unicast also accommodates a bottleneck problem, since each receiver cannot share the bandwidth. On the contrary, SPT multicast can save the bandwidth usage as well as reduce end-to-end delay by bandwidth sharing. On the other hand, RPT multicast is developed to allow shared multicast tree for the same receiver group by using a core-based tree where the core node is RP. In RPT multicast, since all data have to be transmitted to the RP firstly before they are distributed to each receiver, the bandwidth used and delay are worse than that of SPT multicast. Sharing of RP is the reason of bottleneck problem.

Acknowledgements

The authors wish to thank Prof. Akinori Nishihara of The Center for Research and Development of Educational Technology, Tokyo Institute of Technology, Japan, for his support and valuable comments as well as Asian Institute of Technology for providing the time to complete this

work. This work was supported partially by the grant from National Institute of Information and Communication Technology (NICT), Japan.

References

- [1] Ballardie, T., "Core Based Trees (CBT version 2) Multicast Routing", *RFC 2189*, September, 1997.
- [2] Boudani, A., Guitton, B., and Cousin, B., "GXcast: Generalized Explicit Multicast Routing Protocol", *the 9th IEEE Symposium on Computers and Communications (ISCC 2004)*, Alexandria, Egypt, June 2004.
- [3] Deering, S. E., "Host extensions for IP multicasting", *RFC 1112*, August, 1998.
- [4] Estrin, D. et al., "Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol specification", *RFC 2362*, June, 1998.
- [5] Farinacci, D., Rekhter, Y., Rosen, C. E., and Ted, Qian, "Using PIM to Distribute MPLS Labels for Multicast Routes", *draft-farinacci-mpls-multicast-03.txt*, November, 2000.
- [6] Gaeil, A. and Woojik, C., (2000), "Design and Implementation of MPLS Network Simulator Supporting LDP and CR-LDP", *Proceedings IEEE International Conference on Network (ICON 2000)*, pp. 441-446, 5-8 September 2000.
- [7] Karp, R.M., "Complexity of Computer Computations", *New York: Plenum*, pp. 85-103, 1972.
- [8] Kou, L., Markowsky, G., and Berman, L., "A Fast Algorithm for Steiner Trees", *Acta Informatica*, vol. 15, no. 2, pp. 141-145, 1981.
- [9] May, J., "Multicast Routing Extensions for OSPF", *Commun. ACM*, vol. 37, pp. 61-66, August, 1994.
- [10] Medina, A., Lakhina, A., Matta, I., Byers, J., "BRITE: Universal Topology Generation from a User's Perspective", Computer Science Department, Boston University, April, 2001.
- [11] Ooms, D., Sales, B., Livens, W., Acharya, A., Griffoul, F., Ansari, F., "Overview of IP Multicast in a Multi-Protocol Label Switching (MPLS) Environment", *RFC 3353*, August, 2002.
- [12] Rosen, E., Viswanathan A., and Callon R., "Multiprotocol Label Switching Architecture", *RFC3031*, January, 2001.
- [13] Roth, R. et al., "IP QoS Across Multiple Management Domains: Practical Experiences from Pan-European Experiments", *IEEE Communications Magazine*, pp. 62-69, January, 2003.
- [14] Thaler, D., Estrin, D., and Meyer, D., "Border Gateway Multicast Protocol (BGMP): Protocol Specification", *Internet draft*, August, 1998.
- [15] Thomas, S. A., "IP Switching and Routing Essentials: understanding RIP, OSPF, BGP, MPLS, CR-LDP, and RSVP-TE", *John Wiley & Sons, Inc.*, pp. 221-281, 2001.
- [16] Waitzman, D. and Partridge, C., "Distance Vector Multicast Routing Protocol", *RFC 1075*, November, 1998.
- [17] Waxman, B.M., "Routing of Multipoint Connections", *IEEE Journal on Selected Areas in Communications*, pp. 1617-1622, December, 1988.

- [18] Tanwa Kronkriengkrai, "IP Multicast in MPLS Environment using Protocol Independent Sparse-Mode (PIM-SM)" M.Eng Thesis, Telecommunications, School of Engineering and Technology, Asian Institute of Technology, Pathumthani, Thailand, 2003.



Teerapat Sanguankotchakorn was born in Bangkok Thailand on 08 December 1965. He received the B.Eng in Electrical Engineering from Chulalongkorn University, Thailand in 1987, M.Eng and D.Eng in Information Processing from Tokyo Institute of Technology, Japan in 1990 and 1993 respectively under Japanese Government Scholarship. In 1993, he joined Telecommunication and Information Systems Research

Laboratory at Sony Corporation, Japan, where he holds two patents on Signal Compression. He has been an Assistant Professor and Associate Professor at Telecommunications Field of Study, School of Engineering and Technology, Asian Institute of Technology since 1998 and 2004 respectively. His research interests include Digital Signal Processing, High Speed Network, IP-based multimedia applications and QoS in IP and Ad Hoc Network. He is the senior member of IEEE.



Tanwa Korngkriengkrai was born in Bangkok Thailand. He received the B.Eng in Electrical Engineering from Chulalongkorn University, Thailand in 2001 and M.Eng in Telecommunication from Telecommunications field of Study, School of Engineering and Technology, Asian Institute of Technology, Thailand in 2003. Currently, he is working at the TOT Corporation Public Company in

Thailand. His major fields of research are High Speed Network and Quality of Service in IP Network.