

# Rendezvous Points Based Layered Multicast

Tran Ha NGUYEN<sup>†</sup>, Kiyohide NAKAUCHI<sup>†</sup>, Masato KAWADA<sup>†</sup>, *Student Members*,  
Hiroyuki MORIKAWA<sup>†</sup>, and Tomonori AOYAMA<sup>†</sup>, *Regular Members*

**SUMMARY** Layered multicast approach enables IP multicast to adapt to heterogeneous networks. In layered multicast, each layer of a session is sent to separate multicast groups. These layers will be transmitted on the same route, or on different routes. However, traditional congestion control schemes of layered multicast do not consider the case when layers of a session are transmitted on different routes. In this paper, at first we show that in sparse-mode routing protocols like PIM-SM and CBT, layers of a session can be mapped to different Rendezvous Points or cores due to the bootstrap mechanism. It means that layers of a session can be transmitted on different routes. We then show that traditional congestion control schemes of layered multicast do not work properly in sparse-mode routing regions. At last we introduce Rendezvous Point based Layered Multicast (RPLM), a novel congestion control scheme suitable for sparse-mode routing regions, and show that RPLM works efficiently in regions using sparse mode routing protocols. RPLM uses per-RP packet loss rate instead of the overall one to detect congestion on each route, and can react to congestion quickly by dropping the highest layer on the congested route. In addition, RPLM simultaneously drops all the layers those are useless in quality's improvement to prevent bandwidth waste.

**key words:** *layered multicast, congestion control, sparse mode routing*

## 1. Introduction

In multicast a sender sends data at a single rate for all receivers. This characteristic makes multicast to perform poorly in heterogeneous networks. In heterogeneous networks, there are differences in bandwidth capacity and processing power of receivers. Receivers with low capacity links suffer congestion while high capacity receivers are underutilized.

To cope with this heterogeneity, several approaches are proposed [1]. One approach is feedback-based congestion control in which the sender decides transmission rate based on feedback from receivers. Another approach is *simulcast* in which the sender simultaneously sends multiple streams of same data but encoded at different rates, while receivers with same available bandwidth are gathered into the same group.

The most bandwidth-efficient approach is layered multicast in which the sender encodes data of a session into some layers and sends them on separate multicast groups. Each receiver decides the number of layers to

subscribe based on its available bandwidth [2]. That is, receivers adapt to network conditions by adding and dropping layers (i.e., joining and leaving multicast groups) [3]–[7]. However, these schemes do not consider any multicast routing protocol.

There are two types of routing protocol in multicast. The former is dense-mode routing protocol like DVMRP [8], MOSPF [9], and PIM-DM, in which the distribution tree is Shortest Path Tree rooted at the sender. The latter is sparse-mode routing protocol like PIM-SM [10] and CBT [11], in which the distribution tree is Shared Tree rooted at a Rendezvous Point (RP) or a core, respectively. We will concentrate on the PIM-SM case for simplicity. Note that the analyses for PIM-SM in this paper are also applicable to the CBT case.

When layered multicast is used in a sparse-mode routing region, due to layer-to-RP mapping mechanism (*bootstrap* mechanism) [10], [11], layers can be mapped to different RP. It means that layers are transmitted on different routes. In this paper we call this mapping as *different-RP mapping*.

When different-RP mapping occurs, existing congestion control schemes like Receiver-driven Layered Multicast (RLM) [3] does not work. The first problem is that receivers cannot detect congestion when it occurs in a route which transmits only one layer or a partial set of layers of a session, because the overall packet loss rate used in those schemes can not reflect congestion in each route. The second is that the receivers react to congestion too slowly even if congestion is detected, because dropping the highest subscribing layer cannot cope with the congestion when it occurs on a route which transmits lower layers.

There are several works providing the mechanism of sending all layers of a session on a single multicast group [12], [13]. These schemes do not suffer from different-RP mapping. However, these schemes require some routers to implement additional and relatively complex functions, and are not easy to deploy.

In this paper, at first we show that *different-RP mapping* occurs at a high probability. Second, we show that existing congestion control schemes for layered multicast do not work due to different-RP mapping. At last, we show a novel congestion control scheme called Rendezvous Point based Layered Multicast (RPLM). RPLM uses per-RP packet loss rate instead of the over-

Manuscript received April 16, 2001.

Manuscript revised July 4, 2001.

<sup>†</sup>The authors are with the School of Engineering, The University of Tokyo, Tokyo, 113-8656 Japan.

all one to detect congestion on each route, and can react to congestion quickly by dropping the highest layer on the congested route. In addition, RPLM simultaneously drops all the layers those are useless in quality's improvement in order to prevent bandwidth waste.

The rest of this paper is organized as follows. In Sect. 2, by analyzing the mechanisms of layered multicast and sparse-mode routing protocols, we show that different-RP mapping significantly degrades the performance of layered multicast. In Sect. 3 we describe the algorithm of RPLM. In Sect. 4 we evaluate the performance of RPLM by simulation. Finally we conclude in Sect. 5.

## 2. Layered Multicast in Sparse Mode Regions

### 2.1 Layered Multicast

Layered multicast can cope with network heterogeneity and can provide a finer granularity of congestion control compare to traditional single-rate multicast.

In layered multicast, a sender encodes data into several layers. These layers include a *base layer* and several *enhancement layers*. The base layer contains indispensable data to provide a basic level of quality, and can be independently decoded. The higher enhancement layers<sup>†</sup> provides better quality only when it is decoded together with all the lower layers. Then the sender transmits each layer on separate multicast groups. On the other hand, each receiver adjusts its subscription (the number of received layers) at an appropriate level by adding and dropping layers.

A receiver adds a higher layer if it detects no packet loss in a certain duration, and drops the highest subscribing layer if its packet loss rate is over the threshold. It means that, subscription level is low enough to avoid congestion, and is high enough to let receivers get the best quality according to their available bandwidth.

### 2.2 Sparse Mode Routing Protocols

Sparse-mode multicast routing protocols have been designed to operate efficiently over a sparse region, where the number of domains or networks with session members is significantly smaller than the total number of domains or networks, or receivers are widely distributed. When a router discovers local receivers, it starts sending periodic join-messages toward a group-specific Rendezvous Point (RP) which is selected by the *bootstrap* mechanism described below. Then a shared RP-rooted distribution tree is formed. Data packets are sent from source's local router to the RP by unicast. Packets reaching the RP are forwarded down the RP-rooted distribution tree toward receivers.

RP-based routing protocols like PIM-SM require that every router consistently maps an active multicast address to the same RP. This requirement is satis-

fied by using the following *bootstrap* mechanism [14]. A Bootstrap Router (BSR, a dynamically selected PIM router within a PIM domain) collects the set of potential RPs and distributes the resulting set of RPs (RP-set) to all PIM routers in the PIM domain. When a Designated Router (DR) receives an IGMP Host-Membership-Report from a directly connected receiver for a group for which it has no state, the DR use the following hash function (1) to bind the group to one of RPs in the RP-set.

$$f = (n * ((n * (G \text{ AND } M) + k) \text{ XOR } C(i)) + k) \text{ mod } 2^{31} \quad (1)$$

where  $G$  is multicast address,  $C(i)$  is IP address of the  $i$ -th RP in the RP-set,  $M$  is a mask,  $n$  and  $k$  are constant. The default mask  $M$  is (hex) FFFFFFFC, which makes sets of four consecutive multicast addresses to be mapped to the same RP. The DR of a receiver then sends a join message towards the selected RP. On the other hand, when a DR receives data packet from a directly connected sender for a session, it performs the same algorithmic mapping and sends the data packet to the RP. The same mechanism is applied to CBT.

Unlike CBT, PIM-SM has the option of switching the distribution tree from Shared Tree to Shortest Path Tree (SPT). This switching process can be initiated manually or automatically based on some metrics. However, the recommended policy is to initiate the switch to the SPT after receiving a significant number of data packets during a specified time interval from a particular source [10]. That value (the configured threshold) is not only set for layered multicast traffic, but also for other traffic of the regions. The small threshold will make the Shared Tree to switch quickly to SPT, but the network will face the scalability problem due to large state numbers. The higher threshold prevents this switch process, it also leads to the different-RP mapping as described below.

### 2.3 Different RP Mapping

Layered multicast approach and its congestion control schemes do not consider any type of multicast tree. They implicitly admit that all layers of a multicast session are transmitted on the same route.

Sparse-mode routing protocols use the bootstrap mechanism described in 2.2 to map a multicast address to a RP. However, in layered multicast, each layer is sent to separate multicast addresses. For this reason, when carrying out layered multicast in networks using sparse-mode routing protocols, we cannot guarantee that all layers of a session will be delivered on the same route. Actually layers are transmitted on different routes at a high probability, as shown in Appendix.

<sup>†</sup>The base layer is the lowest layer and has the highest priority. The higher enhancement layer has lower priority.

We call this mapping as different-RP mapping.

According to formulas (A.1) and (A.4) in Appendix, we see that the possibility of different-RP mapping increases when the number of RPs  $r$  or the number of layers  $l$  increases. Higher possibility of different-RP mapping makes RLM to perform more poorly as shown in 2.4. This also leads to the improvement in performance of RPLM compare to RLM, as shown in 4.2.

We believe that layers of a session also could be transmitted on different routes in regions which adopt QoS routing protocols. However, these cases are out of the scope of this paper.

### 2.4 Problem Statement

Existing congestion control schemes of layered multicast will not work when different-RP mapping occurs. There are two problems in those schemes. The first one is that congestions cannot be detected when the congestions occur on links which transmitting only one or a partial set of layers of a layered multicast session, because existing congestion control schemes use the overall packet loss rate of a session as a congestion indicator. The second one is the slow response to congestion when it occurs on route which transmits lower layers, even if congestion can be detected.

Figure 1 is used to illustrate these problems. In Fig. 1, data of a session is encoded to 6 layers. The base layer  $L_1$  and its successive enhancement layers  $L_2$  and  $L_3$  are mapped to a RP ( $RP_1$ ), while layers  $L_4$ ,  $L_5$ , and  $L_6$  are mapped to another RP ( $RP_2$ ). Suppose that all layers are sent at the same rate, and threshold of overall packet loss rate is set to 20%. Let consider the case when congestion occurs somewhere between the RP  $RP_1$  and the multicast router  $Rt_2$ .

Assume that the link between  $RP_1$  and  $Rt_2$  is congested and the packet loss rate of 3 layers  $L_1$ ,  $L_2$ , and  $L_3$  is 30%, the overall packet loss rate of 6 layers is equal to only one half of that (15%). This value is smaller than the threshold for packet loss rate (20%), and the congestion cannot be detected. It means that, receiver  $R$  suffers heavily loss in lower layers (layer  $L_1$ ,  $L_2$ , and  $L_3$ ) but keeps receiving all layers include the

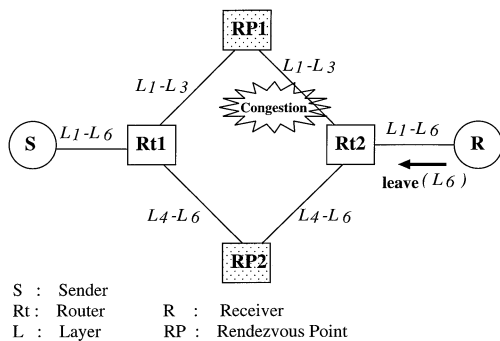


Fig. 1 Existing layered multicast congestion control.

higher one (layer  $L_4$ ,  $L_5$ , and  $L_6$ ). However, receiver  $R$  needs all the lower layers if it want to get higher quality from a higher one. Thus, losses in the lower layers would significantly degrade the received quality.

Now suppose that the link between  $RP_1$  and  $Rt_2$  is congested and the packet loss rate of 3 layers  $L_1$ ,  $L_2$ ,  $L_3$  is 50%, therefore the overall packet loss rate of 6 layers is 25%. This value exceeds the threshold of packet loss rate and the congestion can be detected. Consequently, receiver  $R$  drops the highest subscribing layer  $L_6$  (leave  $L_6$ ). However,  $L_6$  is not transmitted on the congested link, so that congestion cannot be dissolved. It is clearly that the congestion will last until layer  $L_3$  is dropped. In existing congestion control schemes for layered multicast, a receiver conducting drop waits for a certain duration of time after each drop. Obviously, dropping the highest subscribing layer one by one requires a long time. It is not an effective way because of it's slow response to congestion.

### 3. Rendezvous Point Based Layered Multicast

We describe our congestion control scheme called Rendezvous Point based Layered Multicast (RPLM) and show that it can alleviate the problems caused by different RP-mapping in sparse-mode routing protocols.

The overview of RPLM is shown in Fig. 2. Conceptually, RPLM includes two parts, one is the mechanism of congestion detection, and one is mechanism of congestion dissolution.

The congestion detection mechanism considers both the per-RP packet loss rate and the overall packet loss rate of all layers. By using per-RP packet loss rate, RPLM can detect congestion occur in each routes which cannot be detected from the overall packet loss rate.

The congestion dissolution mechanism drops the highest subscription layer among layers passing through the congested link, e.g. layer  $L_3$  in Fig. 2, instead of gradually drops the current highest subscription layer one by one (from layer  $L_6$ ) until layer  $L_3$  is dropped. By adopting this mechanism, RPLM ensures the fast reaction to congestion. In addition, it simultaneously drops all layers which are higher than that layer to prevent bandwidth waste.

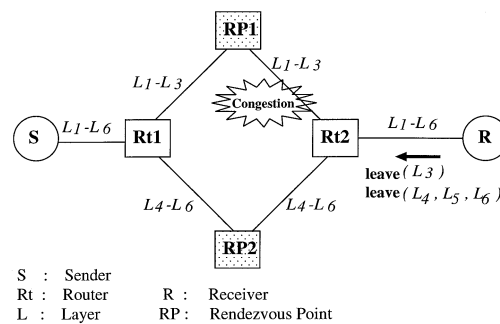


Fig. 2 RPLM congestion control.

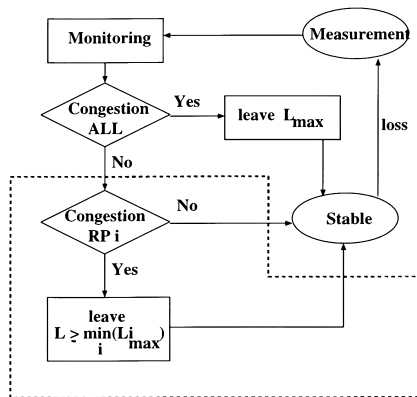


Fig. 3 The RPLM algorithm.

To measure the per-RP packet loss rate, every receiver must know the layer-to-RP map. Within a PIM domain, all receivers can obtain that information through their DRs. In the multi-domain case, we assume that they can obtain that information by using the mechanism implemented in *mtrace* [16], a multicast version of *traceroute* [17].

Figure 3 shows the flowchart of the congestion control algorithm. The part in area bounded by the dotted lines is RPLM's new function comparing to conventional schemes.

There are two states for each receiver, namely, *Stable* state and *Measurement* state. The initial state of a receiver is set to *Stable*. In *Stable* state, if any packet loss is detected, the receiver changes its state to *Measurement* state. In *Measurement* state, the receiver monitors both the overall packet loss rate of all subscribing layers ( $loss_{all}$ ), and the packet loss rate of layers transmitted by each  $RP_i$ , called per-RP loss rate ( $loss_i$ ) (for  $1 \leq i \leq r$ , where  $r$  is the number of RPs).

Firstly, the  $loss_{all}$  is used to detect congestion in common routes which transmit all layers. This step also ensures the back compatibility of RPLM to existing schemes. Then, if congestion in common routes is not detected, per-RP loss rate  $loss_i$  is used to detect congestion in routes passing through each RP. In RPLM, every receiver compares  $loss_{all}$  and  $loss_i$  to the threshold of packet loss rate  $th_{all}$  and  $th_i$ , respectively. The detailed algorithm for each receiver is as follows:

1. if  $loss_{all} \geq th_{all}$ , it means there are congestion somewhere in common routes which transmit all layers (CongestionALL). Then, the receiver drops its highest subscription layer (*leave*  $L_{max}$ ).
2. if  $loss_{all} < th_{all}$  and  $\exists i (loss_i > th_i)$ , it means congestion occurs in a route via  $RP_i$ . Then,
  - a. if only one value of  $i$  exists, the receiver drops the highest layer passing through  $RP_i$  (layer  $L_{i_{max}}$ ). At the same time, it drops all layers which higher than  $L_{i_{max}}$  (*leave*  $L \geq L_{i_{max}}$ ).
  - b. if there are several values of  $i$  exist, the re-

ceiver compares the  $L_{i_{max}}$  together, and identifies a  $j$  for which layer  $L_{j_{max}}$  is the lowest among them, i.e.  $L_{j_{max}} = \min_i L_{i_{max}}$ . Then it drops layer  $L_{j_{max}}$ . At the same time, it drops all layers higher than  $L_{j_{max}}$  (*leave*  $L \geq L_{j_{max}}$ ).

We divide case 2 ( $loss_{all} < th_{all}$  and  $\exists i (loss_i > th_i)$ ) into two cases depending on the number of  $i$ , to clarify the algorithm. Actually, the 2a is a special case of the 2b when  $i = 1$ , as shown in Fig. 3.

Note that RPLM works properly even if all layers are mapped to one RP. This is a special case in which performance of RPLM becomes same as of RLM.

RPLM also works after the distribution tree of any group is switched from Shared Tree to SPT. We consider that SPT as a tree rooted by a virtual RP ( $RP_{virtual}$ ). Note that there is only one SPT that connects a source S and a receiver R independent of group G. If there is any distribution tree of a group is switched from Shared Tree to SPT, the receiver deletes the corresponding layer from set of layers that transmitted by the previous RP. At the same time, it adds that layer to the set of layers which transmitted by  $RP_{virtual}$ . Receivers can know the change in distribution tree by mechanism like *mtrace* [16].

## 4. Simulation

In this section, we show the results of a set of simulations. We use *ns* (network simulator [18]) to evaluate RPLM, by comparing RPLM with RLM.

### 4.1 Simulation Model

Comparison metrics are response time to congestion and *goodput*. In this paper, we define *goodput* as total throughput of layers which are contributing in received quality. For example, when layer  $L_4$  is lost, layers  $L_5$  and  $L_6$  cannot enhance the quality even if they are received without any loss. Therefore, *goodput* is the total throughput of layers  $L_1$ ,  $L_2$ , and  $L_3$ . For simplicity, we only consider the case when layers are assigned continuous multicast addresses. However, RPLM works more efficiently when layers are assigned discontinuous multicast address, because the possibility of different-RP mapping increases (see Appendix).

Simulation topology and simulation parameters are shown in Fig. 4 and Table 1, respectively.

The simulation starts at  $t_0 = 0$  [s] and lasts for 250 [s]. In this simulation, congestion is caused by an unadaptive cross traffic sent at rate of 600 [kbps] between  $RP_1$  and  $Rt_2$  from  $t_1 = 70$  [s] to  $t_2 = 160$  [s].

### 4.2 Simulation Results

The first simulation is carried out with  $B = 1.0$  [Mbps].

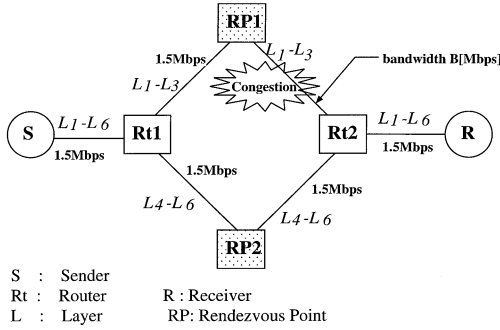


Fig. 4 Simulation topology.

Table 1 Simulation parameters.

Parameter	Value
B (RP1-Rt2 link)	0.8/1.0 Mbps
Number of layers	6
Rate of each layer	200 kbps
Threshold $th_{all}$	0.2
Threshold $th_i$	0.2
Number of RP	2
Layer-to-RP map	L1,2,3 → RP1 ; L4,5,6 → RP2

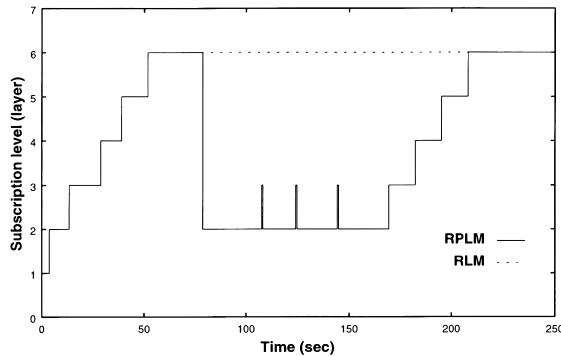


Fig. 5 Number of subscription layers.

Traffic which exceeds  $B$  (200 [kbps]) will be dropped. In this case, the packet loss rate of 3 layers  $L_1$ ,  $L_2$ , and  $L_3$  ( $loss_i$ ) is higher than  $th_i$ , but the overall packet loss rate is equal to only one half of  $loss_i$ , because 3 layers  $L_4$ ,  $L_5$ , and  $L_6$  are transmitted on another route and do not suffer any congestion. In RLM, receiver  $R$  cannot detect congestion because  $loss_{all} < th_{all}$ . Hence, receiver  $R$  continues to subscribe all 6 layers. The dotted line in Fig. 5 shows the number of subscription layers. The receiver  $R$  suffers high packet loss rate in lower layers (layer  $L_1$ ,  $L_2$ , and  $L_3$ ), then goodput heavily oscillates as shown in Fig. 6. Furthermore, it wastes bandwidth for receiving higher layers which passing though  $RP_2$ , because those layers do not enhance the quality.

On the other hand, in RPLM's case, receiver  $R$  can detect congestion in route via  $RP_1$ . It drops 4 layers ( $L_3$ ,  $L_4$ ,  $L_5$ , and  $L_6$ ) simultaneously as shown by the solid line in Fig. 5. As a result, RPLM can react to congestion quickly and goodput is stable at an appropriate level (see Fig. 7). Figure 5 also shows that, receiver  $R$

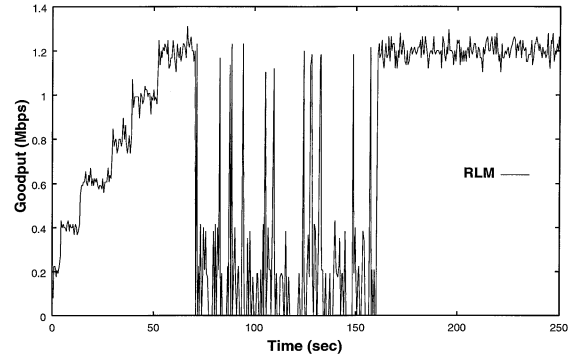


Fig. 6 Goodput (RLM).

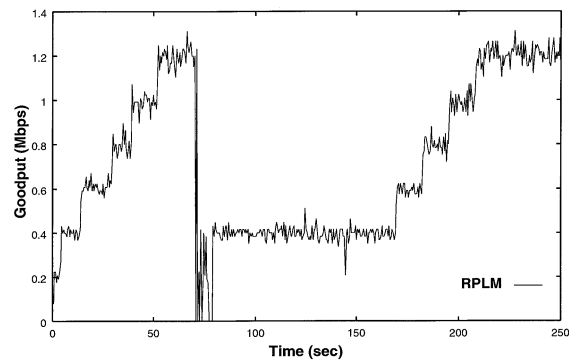


Fig. 7 Goodput (RPLM).

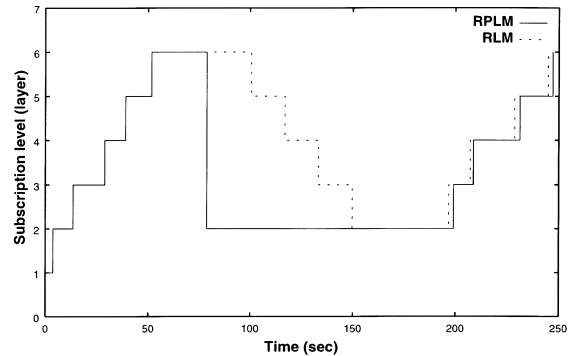


Fig. 8 Number of subscription layers.

have probed for available bandwidth by subscribing to layer  $L_3$  during the congestion, and it gradually subscribes to higher layers when congestion is dissolved.

The second simulation is carried out with  $B = 0.8$  [Mbps]. The exceed traffic is 400 [kbps]. In this case, the overall packet loss rate is greater than the threshold  $th_{all}$ , and the packet loss rate of layers passing through  $RP_1$  ( $loss_i$ ) is greater than threshold  $th_i$ . The dotted line in Fig. 8 shows the number of subscription layers in RLM case. Although receiver  $R$  can detect congestion, it drops only the current highest layer one by one. It takes a long time of about 70 [s] to dissolve congestion. During this period, goodput heavily oscillates due to packet losses in lower layers, as shown in Fig. 9.

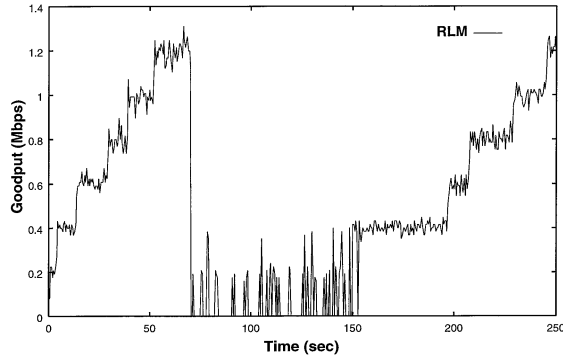


Fig. 9 Goodput (RLM).

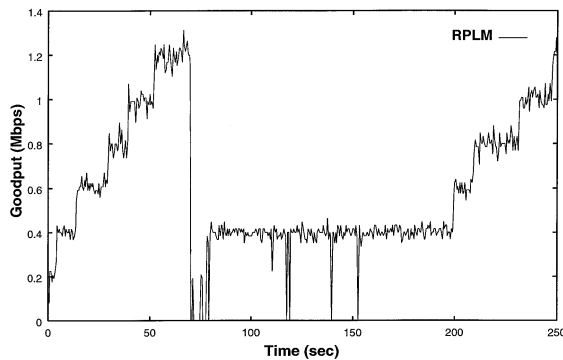


Fig. 10 Goodput (RPLM).

On the other hand, the solid line in Fig. 8 shows that in RPLM case, receiver  $R$  detects congestion and it drops 4 layers ( $L_3$ ,  $L_4$ ,  $L_5$ , and  $L_6$ ) simultaneously. It takes only about 5 [s] to dissolve the congestion. Thus, goodput quickly becomes stable as shown in Fig. 10.

## 5. Conclusions

In this paper we have shown that in regions using sparse-mode routing protocols, layers of a session can be sent on different routes (the different-RP mapping), and existing layered multicast congestion control schemes do not work properly due to problems caused by different-RP mapping. We have described the Rendezvous Point based Layered Multicast (RPLM) congestion control scheme and used simulation to show that those problems are solved by adopting RPLM.

## References

- [1] T. Turletti and J.C. Bolot, "Issues with multicast video distribution in heterogeneous networks," Proc. 6th International Workshop on Packet Video, Oct. 1994.
- [2] N. Shacham, "Multipoint communication by hierarchically encoded data," Proc. IEEE INFOCOM'92, May 1992.
- [3] S. McCanne, V. Jacobson, and M. Vetterli, "Receiver-driven layered multicast," Proc. ACM SIGCOMM'96, Aug. 1996.
- [4] X. Li, S. Paul, and M. Ammar, "Layered video multicast with retransmissions (LVMR): Evaluation of hierarchical rate control," Proc. IEEE INFOCOM'98, March 1998.
- [5] L. Vicisano, L. Rizzo, and J. Crowcroft, "TCP-like congestion control for layered multicast data transfer," Proc. IEEE INFOCOM'98, March 1998.
- [6] L. Wu, R. Sharma, and B. Smith, "Thin streams: An architecture for multicasting layered video," Proc. NOSS-DAV'97, May 1997.
- [7] T. Turletti, S.F. Parisi, and J.-C. Bolot, "Experiments with a layered transmission scheme over the internet," INRIA Research Report, no.3296, Nov. 1997.
- [8] D. Waitzman, C. Partridge, and S. Deering, "Distance vector multicast routing protocol," IETF RFC 1075, Nov. 1988.
- [9] J. Moy, "Multicast extensions to OSPF," IETF RFC 1584, March 1994.
- [10] D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, M. Handley, V. Jacobson, C. Liu, P. Sharma, and L. Wei, "Protocol independent multicast-sparse mode (PIM-SM): Protocol specification," IETF RFC 2362, June 1998.
- [11] A. Ballardie, "Core based trees (CBT) multicast routing architecture," IETF RFC 2201, Sept. 1997.
- [12] S. Bhattacharjee, K.L. Calvert, and E.W. Zegura, "Network support for multicast video distribution," Technical Report, GIT-CC-98-16, 1998.
- [13] K. Nakauchi, H. Morikawa, and T. Aoyama, "A network-supported approach to layered multicast," Proc. IEEE ICC 2001, June 2001.
- [14] D.G. Thaler and C.V. Ravishankar, "Using name-based mappings to increase hit rates," IEEE/ACM Trans. Networking, vol.6, no.1, pp.1-14, Feb. 1998.
- [15] D. Estrin, M. Handley, A. Helmy, P. Huang, and D. Thaler, "A dynamic bootstrap mechanism for rendezvous-based multicast routing," Proc. IEEE INFOCOM'99, March 1999.
- [16] W. Fenner and S. Casner, "A "traceroute" facility for IP Multicast," draft-ietf-idmr-traceroute-ipm-06.txt, March 2000.
- [17] W. Fenner, D. Meyer, and R. Roberts, "MSDP traceroute," draft-ietf-msdp-traceroute-05.txt, March 2001.
- [18] UCB/LBNL/VINT, The Network Simulator—ns-2, <http://www.isi.edu/nsnam/ns/>

## Appendix: Probability of Different RP Mapping

We show that *different-RP mapping* happens at a high probability. This probability is calculated on two cases, depending on the continuity of multicast addresses of layers.

At first, consider the case when layers are assigned discontinuous multicast addresses. Assume that  $l$  is the number of layers and  $r$  is the number of RPs. There are totally  $m_{dis} = r^l$  ways of mapping layers to RPs. Among these  $m_{dis}$  cases there are  $r$  cases that all of layers are mapped to same RP. Therefore, the probability of different-RP mapping is calculated as follows:

$$P_{dis} = \frac{m_{dis} - r}{m_{dis}} = 1 - \frac{1}{r^{l-1}} \quad (\text{A.1})$$

For example, with a session contains 6 layers ( $l = 6$ ), and a RP-set has 2 RPs ( $r = 2$ ), the probability of different-RP mapping is  $P_{dis} = 97\%$ .

Then consider the case when layers are assigned

**Table A.1** Relationship among  $N$ ,  $l$ , and  $c_i$ .

$N \bmod 4$	$l_1$	$c_1$	$l_2$	$c_2$
0	$\lfloor \frac{N}{4} \rfloor$	1	$\lfloor \frac{N}{4} \rfloor + 1$	3
1	$\lfloor \frac{N}{4} \rfloor + 1$	4	0	0
2	$\lfloor \frac{N}{4} \rfloor + 1$	3	$\lfloor \frac{N}{4} \rfloor + 2$	1
3	$\lfloor \frac{N}{4} \rfloor + 1$	2	$\lfloor \frac{N}{4} \rfloor + 2$	2

continuous multicast addresses. With the default value of mask  $M$ , sets of 4 consecutive layers are mapped to the same RP. By considering 4 consecutive layers as one *virtual layer*, Eq.(A.1) can be applied here. Suppose that  $m_{con}$  is the number of possible maps from virtual layers to RPs. Using the same way of calculation with discontinuous address case, we can obtain the probability of different-RP mapping as follows:

$$P_{con} = \frac{m_{con} - r}{m_{con}} = 1 - \frac{r}{m_{con}} \tag{A.2}$$

We need to find  $m_{con}$  for calculating  $P_{con}$ . Therefore, it is crucial to know the number of virtual layers. Let  $N$  ( $N \geq 4$ ) is the number of layers, and  $l$  is the number of virtual layers. The value of  $m_{con}$  depends on the way of partitioning  $N$  into virtual layers. More accurately, the number of layers of a session in the first virtual layer decides the value of  $l$  and therefore affects to  $m_{con}$ .

There are maximum two possible values of  $l$  for each  $N$ , depending on the way of partitioning  $N$ . Let  $c_i$  ( $i = 1, 2$ ) is the number of groups that have same value of  $l$  in each case.

For example, we can divide 6 layers into 4 groups, each group has 2 or 3 virtual layers. Those groups are  $(L_1, L_{2345}, L_6)$ ,  $(L_{12}, L_{3456})$ ,  $(L_{123}, L_{456})$ , and  $(L_{1234}, L_{56})$ . There are 3 groups ( $c_1 = 3$ ) that have 2 virtual layers ( $l_1 = 2$ ): group  $(L_{12}, L_{3456})$ ,  $(L_{123}, L_{456})$ , and  $(L_{1234}, L_{56})$ ; and one group ( $c_2 = 1$ ) that has 3 virtual layers ( $l_2 = 3$ ): group  $(L_1, L_{2345}, L_6)$ . It's obvious that  $c_1 + c_2 = 4$  independent of  $N$ , because the number of layers in the first virtual layer ranges from one to four.

Table A.1 shows the relation among the number of layers  $N$ , the number of virtual layers  $l_i$ , and the number of groups  $c_i$ .

Then  $m_{con}$  can be calculated as:

$$m_{con} = c_1 \times r^{l_1} + c_2 \times r^{l_2} - (c_1 + c_2 - 1) \times r \tag{A.3}$$

$$= c_1 \times r^{l_1} + c_2 \times r^{l_2} - 3 \times r \tag{A.4}$$

For example, with a session contains 6 layers ( $N = 6$ ), and a RP-set has 2 RPs ( $r = 2$ ), we get  $l_1 = \lfloor \frac{N}{4} \rfloor + 1 = 2$ ,  $c_1 = 3$ ,  $l_2 = 3$ ,  $c_2 = 1$ , and  $m_{con} = 14$ . Therefore the probability of different-RP mapping is  $P_{con} = 86\%$ .



**Tran Ha Nguyen** received the B.E. degree in Information and Communication Engineering from the University of Tokyo, Tokyo, Japan, in 2000. He is currently an M.E. student in Information and Communication Engineering at the University of Tokyo. His research interests are in the areas of computer networks, multicast routing, congestion control, and differentiated services.



**Kiyohide Nakauchi** received the B.E. and M.E. degrees in Information and Communication Engineering from the University of Tokyo, Tokyo, Japan, in 1998 and 2000 respectively. He is currently a Ph.D student of Information and Communication Engineering at the University of Tokyo. His research interests are in the areas of multicast communications and peer-to-peer networking. He is a member of IEEE and IPSJ.



**Masato Kawada** received the B.E. and M.E. degrees in Information and Communication Engineering from the University of Tokyo, Tokyo, Japan, in 1999 and 2001 respectively. He is currently a Ph.D. student of Information and Communication Engineering at the University of Tokyo. His research interests are in the areas of computer networks, network collaboration, and multicast communications.



**Hiroyuki Morikawa** received the B.E., M.E., and Dr. Eng. degrees in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1987, 1989, and 1992, respectively. He is currently an Associate Professor of the Department of Information and Communication Engineering at the University of Tokyo. From 1997 to 1998, he stayed in Columbia University as a visiting research associate. His research interests are in the areas of computer networks, distributed computing, mobile computing, wireless networks, and network services. He served as Editor of Transactions of Institute of Electronics, Information and Communication Engineers (IEICE) and on the technical program committees of IEEE/ACM conferences and workshops. He is a member of IEEE, ACM, ISOC, IPSJ, and ITE.



**Tomonori Aoyama** received the B.E., M.E. and Dr. Eng. from the Univ. of Tokyo in 1967, 1969, and 1991 respectively. Since he joined NTT in 1969, he was engaged in R & D on various communication networks and systems in the NTT Labs. He stayed in MIT as a visiting scientist in 1973–1974. In 1994 he was appointed to Director of NTT Opto-Electronics Laboratories, and then in 1995 he became Director of NTT Op-

tical Network Systems Labs. In 1997 he left NTT, and joined the University of Tokyo, and he is now Professor in the Department of Information and Communication Engineering there. He is a member of IEEE, and currently a member of IEEE ComSoc Board of Governors. He is the former President of IEICE Communication Society, and is Chairman of IEEE ComSoc Japan Chapter. He is co-author or co-editor of several books for digital signal processing, ATM broadband networks and optical fiber transmission systems.