Analysis of Internet Multicast Traffic Performance Considering Multicast Routing Protocol

Seiji Ueno[†] Toshihiko Kato[‡] Kenji Suzuki[‡]
[†]*The University of Electro-Communications* [‡]*KDD R&D Laboratories*
ueno@sowa.is.uec.ac.jp

Abstract

Recently, audio and video delivery services are widely spread in the Internet. In order to deliver these data to multiple receivers at the same time, the multicast technologies are indispensable. In such a situation, the performance analysis of multicast traffic will be important in order to design and operate the Internet. Multicast traffic is controlled by its own routing protocol and the multicast routing protocol is more complex than conventional unicast routing protocols. In this paper, we propose a new multicast performance analyzer considering DVMRP and PIM-SM routing protocols. It captures messages of a multicast routing protocol as well as multicast IP datagrams. It also analyzes DVMRP and PIM-SM messages in order to investigate the behaviors of multicast IP datagrams and measures the statistics of multicast traffic in detail. This paper describes the detailed design of the multicast performance analyzer and the results of analyzing multicast datagrams transmitted over networks.

1. Introduction

Recently, audio and video data delivery is widely spread in the Internet. In order to deliver these data to multiple receivers at the same time, the multicast technologies are indispensable [1]. MBone (Multicast Backbone) [2] conveys multicast traffic in a worldwide scale. Currently, multicast traffic itself is not a major one. However, as the bandwidth of the Internet becomes wider and the broadband video transmission becomes more popular, multicast traffic will increase and become a major traffic like WWW traffic and email traffic.

In this situation, the performance analysis of multicast traffic will be important in order to analyze the volumes, the sources/destinations and the impacts on other traffic of multicast traffic. Multicast traffic uses its own communication mechanism. For example, the destination IP address of a multicast IP datagram is a class D address representing a multicast group, not a specific computer system. The multicast routing protocols, which are completely different from unicast routing protocols, provide the function to start and stop multicast traffic as well as that to exchange routing information. Therefore, the performance analysis of multicast traffic requires tools designed for multicast traffic itself.

So far, several tools for analyzing multicast traffic have been developed. Most of them are intended to analyze multicast network topologies [3,4]. For example, MantaRay [3] developed by CAIDA (Cooperative Association for Internet Data Analysis) asks multicast routers on the connectivity with other routers, combine connectivity information obtained from multiple multicast routers, and calculates the possible multicast routes. UCLA and Xerox PARC Multicast Route Monitor [4] listens to route update information of multicast routing protocols and creates the multicast routing tables.

However, those tools cannot measure the multicast traffic performance. The other type of tools is proposed, which captures IP datagrams over a communication link and measures the volume of multicast traffic transmitted over the link. An example is MultiMON [5], developed by CRC (Communications Research Centre) in Canada. It collects packet counts and byte counts of multicast traffic captured at one point of Internet.

However, current measurement tools do not consider multicast routing protocols at all, and therefore, they cannot analyze the dynamical behaviors of multicast traffic. For example, a widely used multicast routing protocol, DVMRP (Distance Vector Multicast Routing Protocol) [6,7] uses a *Broadcast & Prune mechanism.* That is, a broadcast tree is build from a source by exchanging routing information, and the per-source-group multicast trees are created dynamically by pruning (removing

branches from) the source's broadcast tree. Therefore, there are three possible reasons for transmitting multicast datagrams over a particular link.

- A source starts the transmission of multicast traffic.
- A prune is timeout in a multicast router and it starts forwarding multicast datagrams.
- A new member has joined the multicast group.

In order to identify why multicast datagrams are transmitted over a link, the examination of the behaviors of DVMRP is indispensable, which the tools proposed so far do not support.

Therefore, we have developed a multicast performance analyzer considering multicast routing protocols. It captures messages of a multicast routing protocol as well as multicast IP datagrams. It also analyzes a multicast routing protocol in order to investigate the behaviors of multicast IP datagrams and measures the statistics of multicast traffic in detail. As a multicast routing protocol which the analyzer supports, we have selected DVMRP and PIM-SM (Protocol Independent Multicast – Sparse Mode) [8,9], both of which are used widely in Japanese MBone (JP-MBone) [10]. This paper describes the detailed design of the multicast performance analyzers for DVMRP or PIM-SM, and the results of analyzing multicast datagrams transmitted over the Internet by using the analyzers. In the rest of the paper, the next section describes the overview of DVMRP and PIM-SM. Section 3 describes the overall design of the multicast performance analyzer. Section 4 describes the detailed design of the analyzer for DVMRP (DVMRP performance analyzer) and results of adapting the analyzer to an actual network. Section 5 described the detailed design of the analyzer for PIM-SM (PIM-SM performance analyzer) and results of adapting the analyzer to experimental network.

2. Overview of DVMRP and PIM-SM

2.1. Overview of DVMRP

2.1.1. Neighbor Discovery. DVMRP router sends *Neighbor Probe message* periodically to all of its capable network interfaces and tunnel pseudo interfaces. It contains a list of neighbor DVMRP routers from which Neighbor Probe messages have been received. Once a router has received a Probe message that contains its address in the neighbor list, the router has established a two-way neighbor relationship with the neighbor sending this Probe message.

2.1.2. Building Broadcast Tree. The forwarding of multicast datagrams is performed based on the *Reverse Path Multicasting scheme*. That is, when a DVMRP router receives a datagram from a network interface, it forward this datagram only if the network interface is in the shortest path for the source of the datagram. This defines the ordering of upstream and downstream routers. In order to realize the Reverse Path Multicasting, the DVMRP routing table includes tupples with the source network address (prefix and subnet mask), the address of upstream router and the metric to the source network. After a DVMRP router has established a neighbor relationship, it sends information included in the DVMRP routing table to neighbor routers in a *DVMRP Report message*.

Moreover, DVMRP uses the route information exchanges as a mechanism for an upstream router to determine if any downstream routers depend on it for forwarding multicast IP datagrams. If a downstream router selects an upstream router as the best from-gateway for a particular source network, it sends the route information including the original metric plus infinity (32 is used as infinity). This mechanism is called *Poison Reverse*.

2.1.3. Building Multicast Tree. As described above, DVMRP uses a Broadcast & Prune mechanism. If a DVMRP router does not have any members of a particular multicast group and any dependent downstream routers about the multicast group, it sends a *Prune message* to the upstream router to ask for stopping the forwarding of multicast IP datagrams of the multicast group. A Prune message includes source network address, multicast group address and lifetime of this message. The upstream router receiving this message stops the forwarding during the indicated lifetime, and after the lifetime has elapsed, it will start forwarding again. If a new member for a particular multicast group is added, a corresponding multicast router sends a *Graft message*, including source network address and multicast group address, to the upstream router.

2.2. Overview of PIM-SM

2.2.1. General Features. PIM-SM is designed for the environment where multicast group members are distributed across a wide area sparsely. In contrast with DVMRP, a PIM-SM multicast router forwards datagrams only when it receives an explicit *Join message* from downstream routers.

Another feature of PIM-SM is that it uses a shared multicast delivery tree among multicast traffic of the same group. A multicast router called *RP (Rendezvous Point)* is set up for an individual multicast group and multicast datagrams of the group are transmitted via this RP. This delivery tree is called an *RPT (Rendezvous*

Point Tree). When multicast traffic from a specific sender becomes large, this specific traffic is transmitted through the shortest path from the sender. This delivery tree using shortest path is called an *SPT (Shortest Path Tree)*.

2.2.2. Sending Datagrams to Multicast Group. Between an RP and a multicast router that accommodates a sender, a multicast datagram is transmitted by encapsulating it in a *Register message* transmitted as a unicast datagram. When an RP receives a Register message, it decapsulates the original multicast datagram and forwards it to receivers. If the RP does not have any members of multicast group, then it sends a *Register-stop message* to the sender of the Register message. When a router receives a Register-stop message, it stops sending multicast datagrams and starts the *Register-Suppression-Timer*. If it expired, the sender starts transmitting Register messages again.

2.2.3. Joining to / Removing from Multicast Group. When a receiver wants to join to a multicast group, a multicast router accommodating it sends a *Join message* to the corresponding RP. A Join message includes source address, multicast group address, upstream neighbor address, wildcard bit and RPT bit. In this case, the multicast group address in this message is the group address to join, and the source address is the address of corresponding RP. The wildcard and RPT bits are set. This Join message is forwarded in a hop-by-hop manner toward the corresponding RP using the multicast to neighbors. In order to determine an upstream neighbor multicast router, a multicast router will look up a unicast routing table. Multicast datagrams are forwarded in the reverse direction via the path through which the Join message traversed.

When a receiver wants to leave from a multicast group, a *Prune message* are sent in a similar way with a Join message.

2.2.4. Switching Delivery Tree from RPT to SPT. When a multicast router detects that multicast traffic from a particular source becomes large, then it tries to switch the delivery tree from RPT to SPT. The detecting router sends a Join message toward the source. In this case, the multicast group address in the Join message is that of the multicast traffic and the source address is the IP address of the source, and the wildcard and RPT bits are unset.

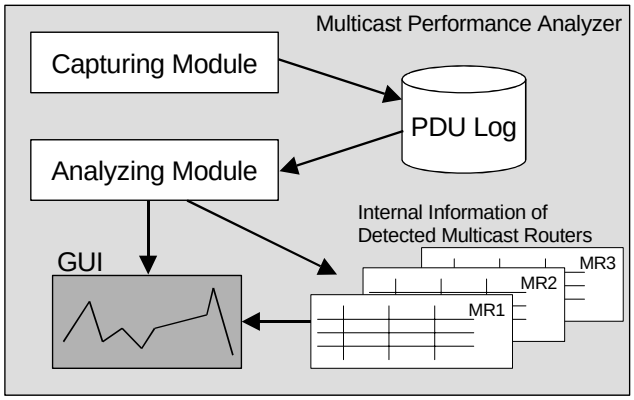


Figure 1. Software structure of multicast performance analyzer

3. Overall Design of Multicast Performance Analyzer

3.1. Overview

The multicast performance analyzer is attached to a link through which multicast datagrams are transmitted. It captures IP datagrams and analyzes multicast datagrams and multicast routing protocol messages. The analyzer is implemented as software running over UNIX workstations. Figure 1 shows the software structure of analyzer. It consists of the capturing module and the analyzing module.

3.2. Capturing Module

The capturing Module monitors an attached link, captures datagrams related multicast traffic and saves the following data in the PDU (Protocol Data Unit) log together with the timestamp of capturing:

- the header information of IP datagram whose destination IP address is a class D address,
- in DVMRP, the header information of a unicast IP datagram encapsulating a multicast IP datagram and that of the encapsulated multicast datagram, and
- the header and body information of a DVMRP message and a PIM-SM message.

3.3. Analyzing Module

The analyzing module is invoked in an offline manner. It reads information stored in the PDU log and performs the following functions.

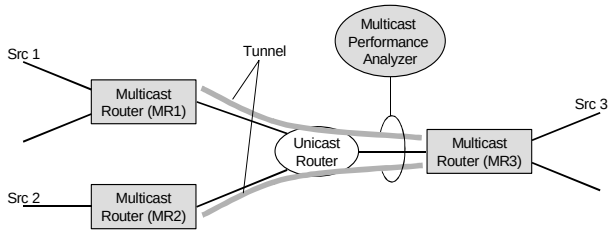


Figure 2. Example network structure with DVMRP performance analyzer

(1) Separation of multicast traffic and measurement its statistics

For information on the multicast datagrams, the analyzing module separates them according to the pair of source and multicast group. In DVMRP, multicast routers may be connected by a tunnel, and it will separate multicast datagrams using encapsuled IP header information. It then measures the statistics, such as transmitted bytes and packet counts in every fixed interval, for separated multicast traffic.

(2) Estimation of router's internal information

Multicast routers have internal information for building trees to deliver multicast traffic. They also have information for dynamically removing and adding branches from and to the multicast delivery tree according to dynamical leave and join of receivers. The analyzing module estimates such internal information by examining messages of multicast routing protocol stored in the PDU log.

(3) Mapping between multicast traffic and estimated internal information

The analyzing module maps the statistics of multicast traffic and the estaimated internal information of multicast routers. This mapping helps to identify reasons that the transmission of multicast datagrams is started or stopped.

4. DVMRP Performance Analyzer

4.1. Procedures to Analyze Multicast Datagrams and DVMRP Messages

4.1.1. Measuring Multicast Datagrams. The DVMRP performance analyzer measures the statistics in the following procedures.

- The analyzer identifies the communication between multicast routers over a link to which it is attached. In an example shown in Fig. 2, it will identify two tunnels, one is between MR1 and MR3 and the other is between MR2 and MR3.
- Since DVMRP controls the multicast traffic on a source-group basis, the analyzer discriminates an individual pair of source IP address and destination multicast group over an individual link or tunnel.
- The analyzer counts datagrams stored in the PDU log on a souce-group basis. The datagrams are counted using the unit of bytes per second. It displays traffic-vs-time graph through a graphical user interface.

Source Network	Subnet Mask	From Gateway	Metric	Expire Time (sec)
Src 1	255.255.255.0	Unknown	4	Unknown
Src 2	255.255.255.0	MR3	6	80
Src 3	255.255.255.0	MR3	3	60

Figure 3. Example of estimated DVMRP routing table for MR1

Source Network	Subnet Mask	From Gateway	Metric	Expire Time (sec)
Src 1	255.255.255.0	MR3	6	120
Src 2	255.255.255.0	Unknown	4	Unknown
Src 3	255.255.255.0	MR3	3	50

Figure 4. Example of estimated DVMRP routing table for MR2

Source Network	Subnet Mask	From Gateway	Metric	Expire Time (sec)
Src 1	255.255.255.0	MR1	5	70
Src 2	255.255.255.0	MR2	5	90
Src 3	255.255.255.0	Unknown	2	Unknown

Figure 5. Example of estimated DVMRP routing table for MR3

4.1.2. Estimating DVMRP Routing Table. By analyzing Report messages stored in the PDU log, the DVMRP performance analyzer estimates a DVMRP routing table maintained in a detected multicast router. The estimated DVMRP routing table includes source network address, subnet mask, from gateway, metric and expire time. “Source network” and “Subnet Mask” indicate a network that sent multicast datagrams. “From Gateway” indicates the upstream neighbor router for a source network. The “From Gateway” is identified by analyzing Poison Reverse mechanism. In the Fig. 2, networks Src 1, Src 2 and Src 3 are reached from MR1, MR2 and MR3, respectively. In this case, the DVMRP routing tables of

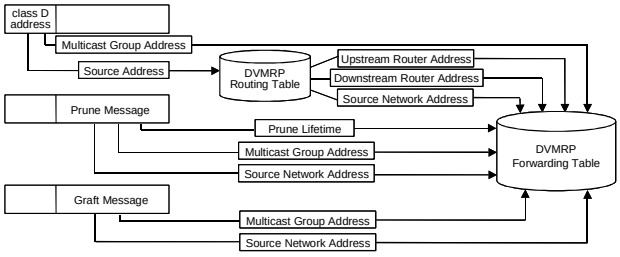


Figure 6. Estimation of forwarding table

Source Network	Multicast Group	Upstream Router		Downstream Routers	
		Address	Prune Lifetime	Addresses	Prune Lifetime
Src 1	224.1.1.1	MR1	200	MR2	200
	224.2.2.2	MR1	-	MR2	-
	224.3.3.3	MR1	-	MR2	-
Src 3	224.2.2.2			MR1, MR2	150, -

Figure 7. Example of estimated DVMRP forwarding table (of MR3)

MR1, MR2 and MR3 can be estimated in shown in Figs. 3, 4 and 5, respectively. It should be noted that, in this example, the upstream router of MR1 for Src 2 is MR3 because and tunnel are not created between MR1 and MR2. "Metric" indicates a cost for a source network. "Expire Time" indicates a remaining time in second until removing this entry from routing table.

4.1.3. Estimating Dynamical Maintenance of Multicast Tree. As described above, a multicast router selects dependent downstream routers using the Poison Reverse mechanism and the forwarding to downstream routers are dynamically controlled by Prune and Graft messages. In order to estimate these DVMRP behaviors, the analyzer introduces a DVMRP forwarding table for individual detected multicast routers. This table consists of

- source network address,
- multicast group for individual source network address,
- upstream router and associated Prune Lifetime, and
- list of dependent downstream routers and associated Prune Lifetime.

The multicast performance analyzer estimates the DVMRP forwarding table in the following procedures (See Fig. 6).

(1) If the analyzer finds a multicast datagram in the PDU log, it detects the source network address corresponding to the source host address of this datagram. This is done by looking up the DVMRP routing table. It then searches the DVMRP forwarding table for an entry corresponding to the source network address and multicast group address. If there are no corresponding entries, then it adds a new entry. In this case, it searches the DVMRP routing table for upstream or dedicated downstream router address. Figure 7 shows an example of estimated forwarding table for multicast router MR3 in Fig. 2. In this example, some datagrams have been detected from Src 1 to multicast groups 224.1.1.1, 224.2.2.2 and 224.3.3.3, and from Src 3 to a multicast group 224.2.2.2. For datagrams from Src 1, the analyzer only understands that MR1 is the upstream router of MR3 from the estimated routing table. It also understands that MR2 is a dependent router of MR3 from the fact that, for Src 1, the from gateway is MR3 in the MR2's routing table. Similarly, for datagrams from Src 3, the analyzer understands that both MR1 and MR2 are dependent downstream routers of MR3. Therefore, four entries, three from Src 1 and one from Src 3, are created for the above mentioned multicast traffic.

(2) If the analyzer finds a Prune message in the PDU log, it searches the DVMRP forwarding table for an entry corresponding to the source netwrok address and the multicast group address included in the Prune message. Then it updates the value of Prune Lifetime in the forwarding table. In the example shown in Fig. 7, the analyzer has detected Prune messages transmitted by MR2 for the multicast traffic from Src 1 to 224.1.1.1 and the Prune Lifetime value was 200 seconds. Also a Prune message has been detected from MR1 for the traffic from Src3 to 224.2.2.2 with Prune Lifetime 150 seconds.

(3) If the analyzer finds a Graft message in the PDU log, it similarly searches the forwarding table for an entry corresponding to the source network address and the multicast group address included in the Graft message. Then it resets Prune Lifetime in the entry. This means

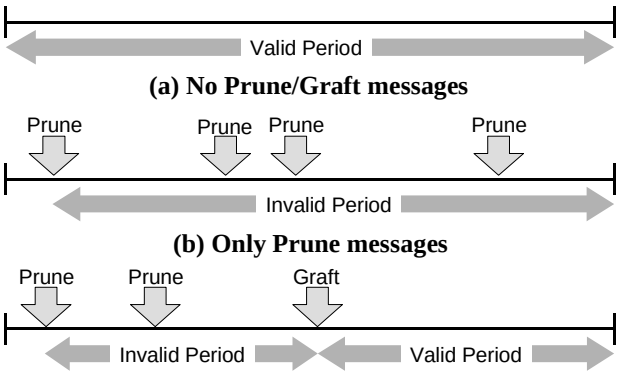


Figure 8. Valid and invalid period

that this multicast traffic needs to be forwarded again.

4.1.4. Mapping between Traffic and Estimated Information. Since DVMRP uses the Broadcast & Prune mechanism, it sometimes occurs that multicast IP datagrams are transmitted due to the Prune Lifetime expires and another Prune message is generated to stop this transmission. This continues until the sender stops the transmission or a receiver comes up. This kind of procedure can be detected by analyzing Prune and Graft messages. If there is a Graft message or there are no Prune messages, the analyzer considers that one or more members exist for this multicast group in the downstream network and handle this time period as a valid period. If Prune messages are transmitted periodically, the analyzer considers that no members exist for this multicast group in the downstream network and handle this time period as a invalid period. Figure 8 shows some exam-

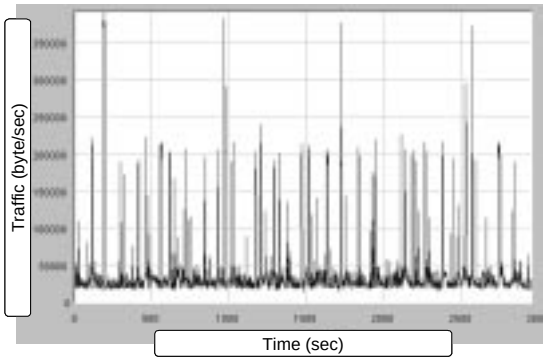
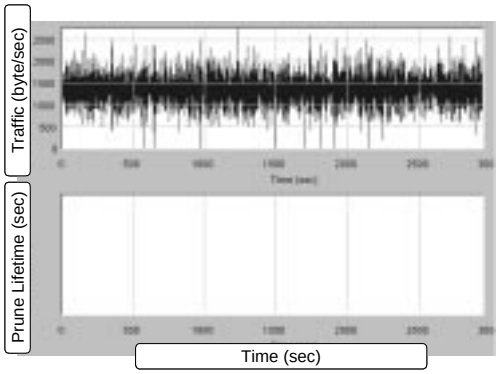
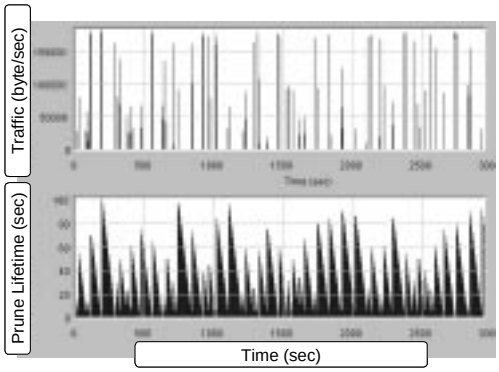


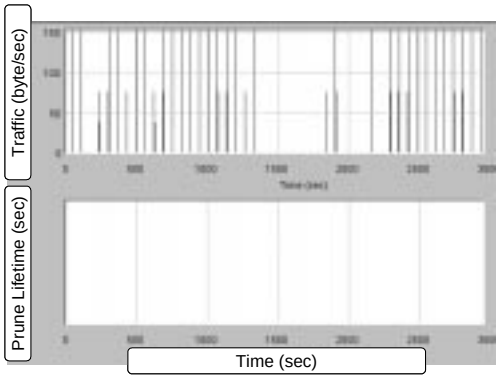
Figure 9. Aggregated multicast traffic over monitored link



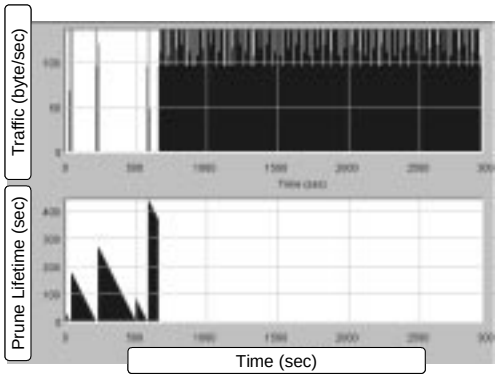
(a) Traffic transmitted continuously



(b) Pulse type traffic with frequent Prune



(c) Pulse type traffic without associated Prune



(d) Traffic with Prune and Graft

Figure 10. Results of per-source-group analysis based on router's internal information

ples of valid and invalid periods. The analyzer measures the statistics of multicast traffic for valid and invalid periods separately.

4.1.5. Graphical User Interface. As described above, the multicast performance analyzer displays bytes of multicast traffic transmitted in one second for any pair of source and multicast group. It also displays the associated Prune Lifetime value graph in order to indicate valid and invalid period explicitly. If Prune message is detected, the Prune Lifetime value in the graph in increased to the value indicated in the Prune message and is decreased to zero as the time elapses. If a Graft message is detected for the pair of source and multicast group while the Prune Lifetime value in the graph is not zero, it is decreased to zero at this moment. As the result, the time period while the Prune Lifetime value graph is not zero corresponds to the invalid period.

4.2. Results of Experiments

We have implemented the DVMRP performance analyzer and have performed some experiments in an actual network about one hour. Figures 9 and 10 show some examples of these experiments. When the analyzer monitors a link, the multicast traffic with class D address shown in Fig. 9 was observed. This is an aggregated traffic over the monitored link. The analyzer then separated this traffic into source and multicast group basis. As the result, we obtained four types of multicast traffic shown in Figs. 10 (a) through (d). The analyzer also shows the associated Prune Lifetime. The first type is a traffic in which multicast datagrams are transmitted continuously. The graph in Fig. 10 (a) shows that there are no Prune messages and therefore it is considered that the source is continuously sending datagrams and there are one or more members for this traffic in downstream networks. The second one is a pulse type traffic with associated Prune messages. The graph in Fig. 10 (b) shows that the Broadcast & Prune is performed in this traffic. That is, this traffic is that in the invalid period. We observed a similar pulse type traffic as shown in Fig. 10 (c). However, there are no Prune messages detected. This means that this is not a Broadcast & Prune type traffic, but that a pulse type traffic is transmitted by the source and there are members for this traffic. The last type consists of a pulse type traffic and continuous data transmission. From the Prune Lifetime value graph in Fig. 10 (d), it is concluded that a Graft message is transmitted for this multicast traffic, and after that, datagrams are transmitted continuously.

Table 1. Statistics of captured multicast traffic

Number of detected groups	52 groups
Transmitted Packet count	112799 packets
Transmitted bytes	52602947 byte
Average throughput	17874 byte/sec
Invalid transmitted packet counts	37185 packets (33.0%)
Invalid transmitted bytes	26781609 byte (50.9%)
Average throughput of invalid traffic	9100 byte/sec

As described in section 4.1.4, the analyzer measures the statistics of multicast traffic for the valid and invalid periods. Table 1 shows the result of this measurement. In this table, the valid traffic is only 49.1% from among all traffic over the monitored link.

5. PIM-SM Performance Analyzer

5.1. Procedures to Analyze Multicast Datagrams and PIM-SM Messages

5.1.1. Measuring Multicast Datagtams. Similarly to the DVMRP performance analyzer, the PIM-SM performance analyzer counts datagrams stored in the PDU log on a source-group basis. The datagrams are counted using the unit of bytes per second. It displays traffic-vs-time graph and router's internal information through a graphical user interface.

5.1.2. Estimating Router's Internal Information. The analyzer estimates two tables as a router's internal information. One is the PIM-SM Router Table that maintains RP information. The other is the PIM-SM Forwarding Table that maintains the relationship among the multicast group, the source address, the RP address and the address of upstream router. Figure 11 shows an example of PIM-SM Router Table. It is estimated by analyzing Bootstrap messages.

The PIM-SM Forwarding Table is estimated by analyzing Register messages and Join/Prune messages stored in the PDU log. Figure 12 shows an example. Figure 13 shows the procedure for estimating PIM-SM Forwarding Table. If the analyzer finds a Register message in the log, it extracts the source IP address and the multicast group address from the encapsulated multicast datagram. It then looks up the Router Table and obtains the RP's IP address corresponding to the multicast group. Based on this information, it creates an entry in the Forwarding Table if there are no corresponding entries in the table. At this moment, the upstream router in this entry is null. If a Register-stop message corresponding to the Register message is detected, the analyzer starts Register-Suppression-Timer for this entry.

If the analyzer finds a Join message in the PDU log, it

Group Address	Group Mask	RP Address	Priority	Holdtime (sec)
224.0.1.39	255.255.255.255	192.168.10.20	0	0
224.0.1.40	255.255.255.255	192.168.10.20	0	0
224.0.0.0	240.0.0.0	192.168.10.20	0	120

Figure 11. Estimated router table

Group Address	Source Address	Upstream Router	RP Address	Join Timer (sec)	Register Suppression Timer (sec)
224.1.1.1	INADDR_ANY	192.168.40.20	192.168.10.20	0	-
224.1.1.1	192.168.60.200	192.168.40.60	192.168.10.20	0	90
224.3.3.3	INADDR_ANY	192.168.40.20	192.168.10.20	150	-
224.3.3.3	192.168.50.200	192.168.40.20	192.168.10.20	150	-
224.3.3.3	192.168.60.200	192.168.40.60	192.168.10.20	150	0

Figure 12. Estimated forwarding table

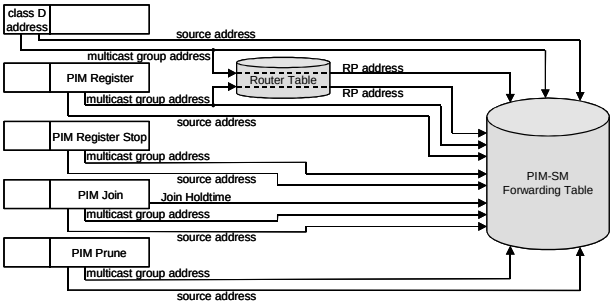


Figure 13. Procedure of estimating forwarding table

extracts the source IP address, the multicast group address, the address of upstream router and Join Holdtime. It searches for the corresponding entry and creates a new entry if there are no corresponding entries. Then, the analyzer updates Join-Timer to Join Holdtime in the Join message. It also sets the address of upstream router in the entry. If the wildcard bit is set in the message, the source IP address in the entry is set to INADDR_ANY. If this bit is not set, the analyzer decides that this Join message is used to join SPT and the source IP address in the entry is set to that in the Join message.

If the analyzer finds a Prune message in the log, it extracts the source IP address and the multicast group address, and resets Join-Timer value in the corresponding entry in the PIM-SM Forwarding Table. If the wildcard bit in the message is set, the analyzer set the source IP address in the entry to INADDR_ANY.

5.1.3. Mapping between Traffic and Estimated Information. Since PIM-SM supports two transmission schemes, with RPT and with SPT, the analyzer estimates that the captured multicast datagrams are transmitted by RPT or SPT. This mapping is performed by analyzing Join and Prune messages. Table 2 indicates the rules to relate Join and Prune messages to RPT or SPT.

Table 2. Rules to related Join/Prune mesages to RPT or SPT

Message	WC-bit	RPT-bit	Description
Join	0	0	Source address in this message is sender's IP address. This message is Join for SPT.
Join	0	1	Invalid message.
Join	1	0	Invalid message.
Join	1	1	Source address in this message is RP's IP address. This message is Join for RPT.
Prune	0	0	Source address in this message is sender's IP address. This message is Prune for SPT.
Prune	0	1	Source address in this message is sender's IP address. This message is Prune, and switches from to SPT.
Prune	1	0	Invalid message.
Prune	1	1	Source address in this message is RP's IP address. This message is Join for RPT.

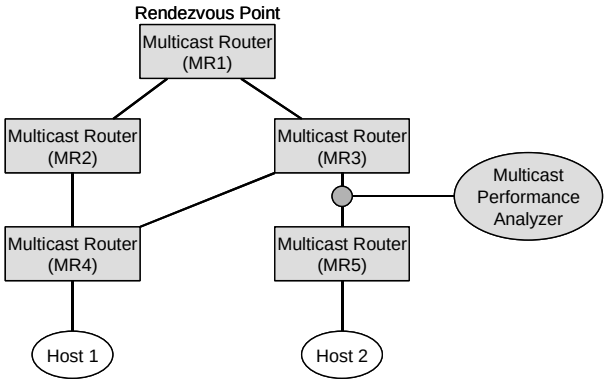


Figure 14. Experimental network structure

5.2. Results of Experiments

We have also implemented the PIM-SM performance analyzer and have applied it to an experimental network shown in Fig. 14. All routers run RIP2 as a unicast routing protocol. MR1 works as RP and a Bootstrap router. Host1 and Host2 behave as a sender and a receiver, respectively. Host1 joins to some multicast groups that are sent by Host2. Also Host2 joins to some multicast groups that are sent by Host1.

Figures 15 and 16 show some results of these experiments. When the analyzer monitors a link, multicast traffic shown in Fig. 15 was observed. This is an aggregated traffic over the monitored link. In this figure, the black line shows native multicast traffic and the gray line shows Register messages which encapsulate multicast datagrams. The analyzer then separated this traffic on a source-group basis. The separated traffic mapped with RPT/SPT information is shown in Figs. 16 (a) through (c). The top graph shows the multicast traffic for every second, the second one shows the Join-Timer for SPT, the third one shows the Join-Timer for RPT, and

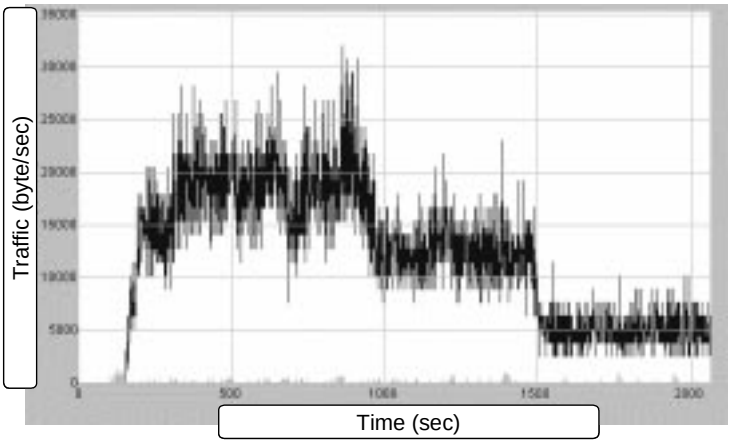
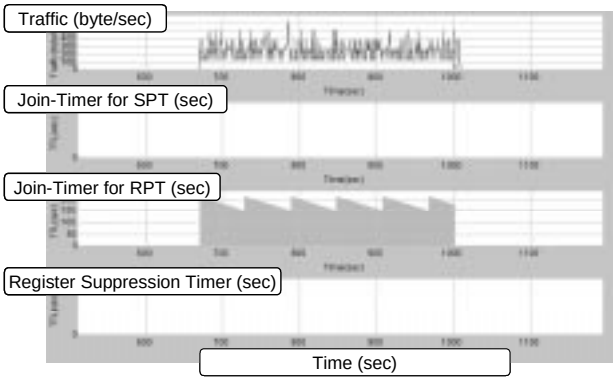
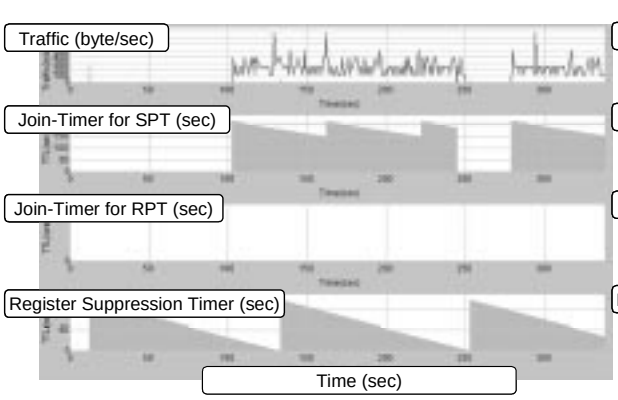


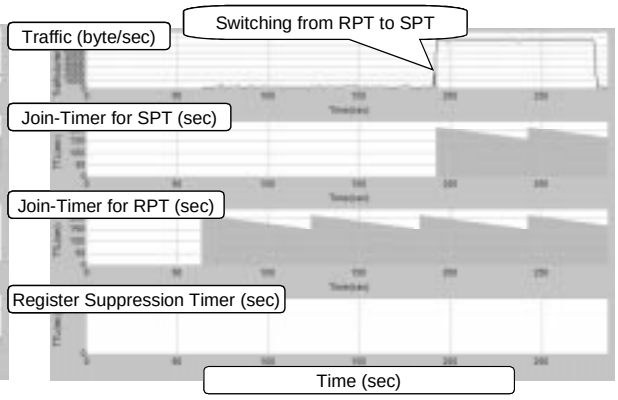
Figure 15. Aggregated traffic over monitored link



(a) Traffic with associated RPT Join message



(b) Traffic with associated SPT Join message and Register with Register-stop message



(c) Switching delivery tree from RPT to SPT

Figure 16. Results of per-source-group analysis based on router's internal information

the bottom one shows Register-Suppression-Timer. If a Join message is detected, the Join-Timer value in the graph is increased to the value in the message and is decreased to zero as the time elapses. If a Prune message is detected for the pair of source and multicast group, the timer value is decreased to zero at this moment. In the Register-Suppression-Timer graph, if a Register-stop message is detected, the Register-Suppression-Timeout

value is increased to the predefined value. In this graph, the predefined value is assumed to be 120 seconds.

As a result, we obtained three types of multicast traffic. The first type is that for which Join messages only for RPT are detected (shown in Fig. 16 (a)). In this case, the analyzer estimates that the traffic transmitted by way of RPT. The second type is that for which Join messages only for SPT are detected, and Register and Register-stop messages are detected (shown in Fig. 16 (b)). In this case, the analyzer estimates that the traffic transmitted by way of SPT. It also estimates that the traffic transmitted by Register messages is no more required. The third type is that for which downstream routers sent Join messages for RPT initially and, after that, it sent Join messages for SPT (shown in Fig. 16 (c)). The analyzer estimates that initially this traffic was sent by way of RPT and then, since multicast traffic increased, a router switched delivery tree from RPT to SPT.

6. Conclusion

In this paper, we have described the multicast performance analyzer, which monitors a network for multicast IP datagrams and multicast routing protocol messages. It measures the bytes of datagrams transmitted in every second for an individual pair of source address and multicast group address. It also analyzes a multicast routing protocol in order to investigate the behaviors of multicast IP datagrams in detail. We have implemented multicast performance analyzer for DVMRP (Distance Vector Multicast Routing Protocol) and PIM-SM (Protocol Independent Multicast - Sparse Mode). The DVMRP performance analyzer analyzes the DVMRP Probe, Report, Prune and Graft messages and estimates the DVMRP routing table and forwarding table of detected routers. As a result, it can estimate the reasons for starting and stopping multicast datagram transmissions, such as sender’s starting transmission or the prune timeout. The PIM-SM performance analyzer analyzes Register, Register-Stop, Join, Prune and Bootstrap messages and estimates PIM-SM Router table and Forwarding table. As a result, it can estimate the multicast traffic changes between RPT (Rendezvous Point Tree) and SPT (Shortest Path Tree). We believe that this detailed analysis is helpful to investigate the multicast traffic behaviors and to design and operate the Internet from the multicast traffic point of view.

7. Reference

[1] T.Maufer and C.Semerica, "Introduction to IP Multicast Routing," draft-ietf-mboned-intro-multicast-03.txt, Jul. 1997.

[2] Vinary Kumar, "MBone," New Riders Publishing, 1995.

[3] "Manta Ray," http://www.caida.org/Tools/Manta/

[4] "UCLA/Xerox PARC Multicast Route Monitor," http://www.ganef.cs.ucla.edu/~masseyd/Route/

[5] "MultiMON - an IPmulticast Monitor," http://www.merci.crc.ca/mbone/MultiMON/

[6] D. Waitzman, C. Partridge and S. Deering, "Distance Vector Multicast Routing Protocol," RFC1075, Nov. 1998.

[7] T. Pusateri, "Distance Vector Multicast Routing Protocol," draft-ietf-idmr-dvmrp-v3-08.txt, Aug. 1999.

[8] D. Estrin, et al., "Protocol Independent Multicast - Sparse Mode (PIM-SM) : Protocol Specification," RFC2362, Jun. 1998.

[9] S. Deering, et al., "Protocol Independent Multicast - Sparse Mode (PIM-SM) : Motivation and Architecture," draft-ietf-idmr-pim-arch-04.ps, Oct. 1996.

[10] N. Shichijo, "The Current Status and Perspectives of MBone," http://www.race.u-tokyo.ac.jp/~shichi/apan. htm