Reliable Multicast Tree Construction Algorithm *

Choonsung Rhee
NIA
Seoul 100-775, Korea
csrhee@nia.or.kr

Jungwook Song
Konkuk University
Seoul 143-701, Korea
swoogi@konkuk.ac.kr

Euijun Kim
Konkuk University
Seoul 143-701, Korea
euijun@cclab.konkuk.ac.kr

Sunyoung Han†
Konkuk University
Seoul 143-701, Korea
syhan@konkuk.ac.kr

ABSTRACT
In order to efficiently providing reliable multicast, this paper analyzed the existing TMTP and RMTP based on tree, and proposed extended architecture to solve problems grasped. The architecture can minimize implosion and exposure, and decreases delay time by tree reconstruction as partial tree reconstruction. We used early OMNeT++ simulator, and simulation did the environment which various, and was complicated to demonstrate effectiveness of the architecture and provided reliability-based the existing tree.

Categories and Subject Descriptors
C.2.2 [**Network Protocols**]: Applications

General Terms
Algorithms

Keywords
Multicast, Overlay Multicast, Tree Construction

1. INTRODUCTION
The reliable multicast ensures that packets transmitted from multicast source is delivered without errors to all recipients. As number of users increases and scale of communications grows larger, in the reliable multicast protocol occurs implosion and exposure problem. Methods to provide reliability can be divided in two ways at reliable multicast. First method is method of providing reliability through distributed recovery[1]. Distributed recovery method manages not only multicast source but also another recipient part. This method helps to reduce implosion. Another method to provide reliability is local recovery method. Local recovery method decreases bandwidth for retransmission and exposure [2] as restore regionally local loss in the neighborhoods. Tree-based a protocol can get naturally variance recovery method and local enemy recovery method [3, 4, 5, 6] using characteristic of the tree structure, and association. Tree-based protocol generates tree which is logically hierarchical in order to disperse responsibility of loss recovery. Intermediate nodes deals with feedback of child node and takes charge of retransmission regarding them, and provide reliability. It is important that the node which requires of retransmission is composed the logical tree at tree-based protocol so as to have different recipient in the stomach to parent than self on multicast routing trees. But relevant nodes can occur inversion problem that up and down layer is reverse by physical hop and logical hop. Multicast tree is taken into consideration in addition to the inversion and delay problem in order to efficiently provide reliable multicast at tree infrastructure, as mentioned. Multicast tree creation method suggested in this paper provides efficiently than exiting reliability multicast method as solving delay problem and tree reconstruction.

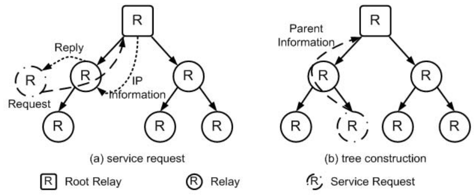
2. PROBLEM OF TREE-BASED STRUCTURE
The representatives of tree-based reliable multicast protocols are TMTP (Tree-based Multicast Transport Protocol) and RMTP (Reliable Multicast Transport Protocol). TMTP uses ESR (Expanded Ring Search) mode in order to compose a logical tree [6]. When new member joins to multicast group, multicast sends control packet which had a TTL price to relevant group. If there is not response from the existing member does again multicast using increased TTL price. Member sending response to related message become new node's parents. The tree configuration method using the TMTP's TTL cannot exclude inversion problem because indirection defect of TTL. There is problem that a logical tree turns on join / secession of member. Also retransmission that used Limited scope multicast uses TTL technology. There is the upper part even in addition to bases of a tree, and it is possible to occur for exposure problem. When there is multicast session, the RMTP DR (Designated Receivers) is decided to static, and assume [3, 4]. The sender and all DR multicasts periodically control packet having the same TTL

*This research was supported by the MKE(Ministry of Knowledge and Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the IITA(Institute for Information Technology Advancement)(IITA-2008-C1090-0804-0015).

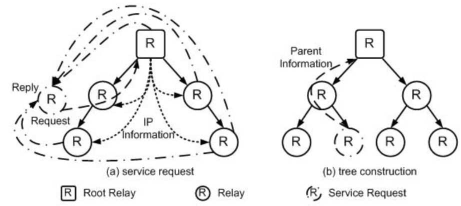
†Corresponding Author

price. New member will compose a logical tree as DR selects parent to have TTL of a price largest among this. Tree configuration method of RMTP does not occur inversion that the control packet spread periodically uses sub-tree multicasting, and makes inversion of a routing tree. But this work is possible if sub-tree multicasting is adapted to all multicast router. DR of current RMTP uses tool called Tracer, and can be selected dynamically. Tracer can chase hops to sender using IGMP MTRACE function. It can select suitable parent as exchange path advertisement and response. However, the periodic hop prompt which used ERS large overhead.

TMTP and RMTP which are multicast protocols to provide reliability, have inversion, various problems such as overhead for implosion, exposure, multicast tree bookkeeping. On the other hand the technique proposed in these papers proposes the architecture which there has expandability for reliable multicast-based a tree. As composed logical tree similar routing tree solved inversion, and presented construction method efficient tree.



(a) New relay joining



(b) When there is no item

Figure 1: Multicast tree construction method

3. RELIABLE MULTICAST STRUCTURE

In reliable multicast proposed in these papers, a multicast tree is composed around relay, and each relay uses NTP (Network Time Protocol), and fit synchronization in time. We will show about multicast tree construction method of architecture proposed in these chapters and partial multicast reconstruction method.

3.1 Method for multicast tree construction

Relay which manages multicast service in proposed architecture knows address for multicast source. If relay is operated at initial, it sets up a TTL price to 254. Multicast service requests a message with IP address of self does unicast to root relay which connected multicast source.

During request message transmission to root relay which is connected multicast source from related relay, it saves message information for route. Root relay gets from service request message and searches existing relays to relay near location requested service in order to compose a logical tree as multicast routing tree. Root relay compares information for relay composing entire multicast tree and route of transmitting message from relay which want new join. Compare of route accomplishes from relay which wants new join and selects upper relays existing in same location or closer location. Root relay have to manage mapping table for TTL values from IP address of each relay to root relay. Information for each relay managed by root relay is used in order to decide parent node of relay and solve inversion of a multicast tree.

Information of IP and TTL is transmitted to root relay periodically. Comparing with hierarchical structure of multicast tree, when the value of under layer is bigger than the value top layer, relay operates more than twice continually regards invension problem. And then location of related relays in tree changes each other result in relays near multicast source locate top layer always. If the candidate relays which will be becoming the parent nodes are decided request service, the root relay transmits the IP information about relay that requests service to the candidate relays. The relay that received the IP information is the relay that requests service sends a message with TTL value 63 to the relay that requests service as reply.

The relay that received response message chooses a relay with maximum TTL value as a parent node and sends this information to the root relay. Figure 1(a) shows the procedure of new relay joining the multicast tree. When the candidate relay cannot be selected because there is no corresponding item are same between multicast tree information preserved in the root relay and request message that joining the service, the root relay demands all relays managed by the root relay to send the reply message to the corresponding relay. For this case, all relays including root relay will fix the TTL value to 254 and send respond message to the relay that request service. The relay that request service will choose a relay with the maximum TTL value as parent node between the respond messages it has received and select it as parent node and send this information to root relay. Figure 1(b) shows procedure of new relay joining the multicast tree when there is no item corresponding to multicast tree information preserved in the route relay.

3.2 Reconstructing Partial Multicast Tree

All relays can obtain information from the node that surrounding the parent node and high-ranking nodes according to the information that is obtained multicast tree information collected in the root relay when the multicast tree is constructed. According to this information, a relay can make information about candidate relays that are possible to connect, in order to restructure partial tree when the trouble occurs in its parent relay. The information about the candidate relays will be made according to information collected by root relay as reference, and the information involves the multicast data and will be transmitted from the root relay periodically to other relays each. The information about the candidate relays will be transmitted when the root relay senses that the tree structure is changed and after some times it. Table 1 explains the point of information about candidate relay that is sent by the root relay.

Table 1: Sending candidate relay information

Events	Description
New relay join	Detecting changes of multicast tree
Relay leave	Detecting changes of multicast tree
Relay failure	Detecting changes of multicast tree
Solve inversion	Detecting changes of multicast tree
Periodic update	When receipt check information at root relay is full

Figure 2 shows a restructuring partial multicast tree method. When the trouble occurs in a specific relay, the child nodes of the trouble relay transmit information about candidate relays and connection request message is transmitted to the corresponding relays. The candidate relays which received connection request send the response message with the fixed TTL value to the relay that requests connection. The relay that receives the response message will connect to the candidate relay with the maximum TTL value. The relay that succeeds in restructuring partial tree will send this information to the root relay. When the root relay that perceives the change in the multicast tree updates whole multicast tree information and list of each relay's candidate relays again and send to the lower-level relays.

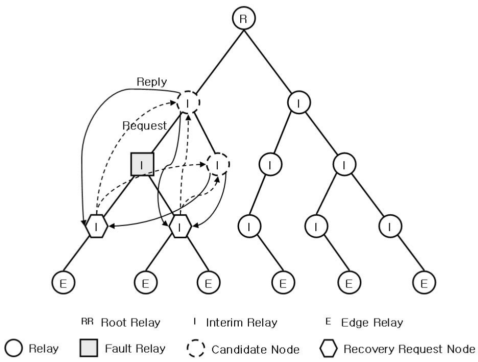


Figure 2: Partial multicast tree reconstruction

Even though such a Partial restoration multicast tree offers shorter delay than the overall recovery cannot offer to maintain in best situation. When the trouble occurs in two relays at the same time or more than two relays the trouble occurs in the same region and the subordinate position relay in the region to solve this problem, the whole multicast tree is restructured. This purpose is in order to minimize the inefficient multicast tree decreasing performance.

3.3 Method for recovery

The proposal structure in this paper, succeeding reception or loss of the multicast data and information on the in by a high-level relay. Using sequence number and the flag bit in the packet to confirm the reception and to reduce the overhead of reception confirmation. Figure 3 shows newly proposed structure for reception confirmation information.

The sequence number means the number in the beginning packet's reception confirmation of the multicast packet. The bit flag for the reception confirmation is expressed by the bit

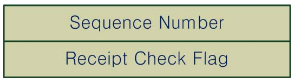


Figure 3: Receipt check information structure

of packet, it has been confirmed or not. The flag bit is set in the first stage, and everyone set to 0 and when the reception of the package of the correspondence is confirmed, it is set to 1. Such reception confirmation information sends to a when the packet there is a request to retransmission in a terminal and there is no corresponding packet in the cache of the relay and requests retransmission to high-level node. Only when all dosages of reception confirmation information are filled, reception confirmation information is transmitted to a high-level relay and a normal confirmation message and the request package to retransmission exist in the cache contrary to him. The size to reception confirmation information for efficient operation was tested to the simulation in the performance evaluation part.

4. PERFORMANCE EVALUATION

We use Network simulator OMNeT++ for performance evaluation against the proposal structure.In the simulation, the receivers can participate to the session dynamically with proposed technique and TMTP to evaluate extendibility. It has the arbitrary value between 20ms to 5ms in the simulation for the measurement latency in consideration of the characteristics of the network between each relay. The loss probability of the link is from 0.01 to 0.2. Therefore give it to each link and the loss occurrence is assumed to be arbitrary. The value of the timer to send NACK used for the loss report of the multicast packet with TMTP is decided with RTT between the service terminal and the relay. And, an arbitrary price was selected by the price of the control timer for suppression of NACK between values that until the half of RTT from 0.

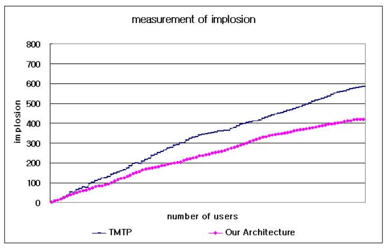


Figure 4: Average implosion

Figure 4 is a graph which is the feedback information implosion was measured with proposed structure and TMTP. As it is shown, feedback implosion proposes by the number of users are growing and the size of the multicast session is increase by TMTP. It was able to be confirmed that the proposed structural overhead appeared from the implosion overhead of TMTP however small by the seen thesis, and the size of the multicast session was increasing larger and it was confirmed because of such an increasing user.

Figure 5, the retransmission packet to restore the packet lost by proposed structure and TMTP passed or a result of measuring the collection as for the link of average though is not

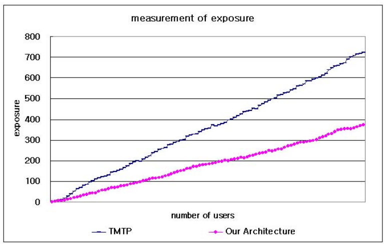


Figure 5: Average exposure

known. The result of the exposure measurement also shows a sartorial increase by growing of the price of growing of the session of the multicast as looked like the implosion measurement graph and the loss probability. It can be confirmed that the overhead of the proposed technique appeared from the exposure overhead of TMTP, The multicast session increase sees a greater difference from the measurement result of implosion.

The receiver dynamically set to the technique and TMTP proposed for the extendibility evaluation by a user increase that he or she participated in the session and it evaluated it. Figure 6 is a graph where the average delay time of the multicast data by a user increase. It discriminated 20ms to 5ms, and the service request time and the connecting location were arbitrarily set and measured by the simulation for the enlargement time measurement at the enlargement time between each relay application.

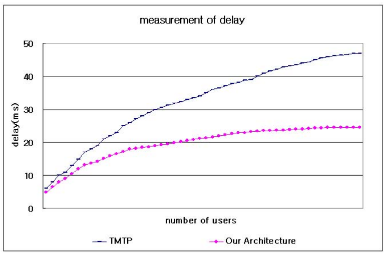


Figure 6: Average delay

Service it was enhanced to whole network, the number of users was increased, and the structure proposed with the thesis thought to be more and more appearance by the simulation result showed the shorter overall latency than the TMTP structure.

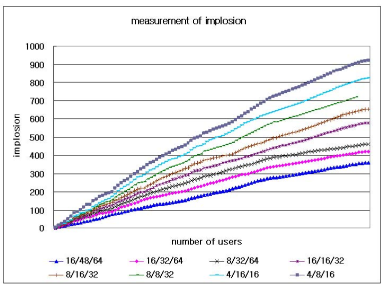


Figure 7: Implosion by receipt check structure size

Reception confirmation information is a result of the Root relay's setting the value to 32KB to the Edge relay and the Interim relay by 8KB for the third graph in high rank. The reason why the graph of the correspondence is not connected with the end is that the trouble occurs in the middle where the multicast service is provided. The trouble that exceeded the dosage of the buffer by the Interim relay in the middle because the Edge relay had been set to 8KB equally to the reception confirmation information size of the Interim relay occurred.

It is a way that reception confirmation information is large in a hierarchical structure of the relaycomposes the multicast tree and the trouble doesn't occur without fail. However, setting the reception confirmation information size of a high-ranking relay more greatly than the size of reception confirmation information on the subordinate position relay can decrease danger to the failure by the excess of the dosage of the buffer. Sizing reception confirmation information in the result of the measurement by the simulation showed the price between 64KB every 48KB in the route relay and when the price of 16KB or less was used in the price between 48KB every 16KB and Edge Relay, better performance was shown in Interim Relay.

5. CONCLUSIONS

The proposed structure proposed a dynamic tree composition method by changes of relays which composed the multicast tree and tree structures to get over the weak point of silence. The basic technique it proposed a regional multicast based method. Implosion and exposure can be generated by a large-scale session for the multicast with reliability through such a method can be minimized, and the enlargement time by the tree restructuring can be decreased by providing continued service as it is a method of restructuring a partial tree and a big change in the overall tree structure doesn't exist. Moreover, the performance of a whole network is advanced by supplementing the weak point of the TTL technique of an existing tree base and trust multicasts, solving the forward problem, and composing a logical tree similar to Multicast. It was simulated in various, complex environments were assumed to prove proposed structural using OMNet++ Simulator. Through the numerical values of result. We can get how dose proposed structure progress with it's advantages. And also can prove that it is much better than existing tree structure based multicast structure which is support reliability.

6. REFERENCES

[1] S. Floyd et. al. A reliable multicast framework for light-weight sessions and application level framing. *ACM SIGCOM 95*, Aug 1995.

[2] S. K. Kasera, et. al. Scalable reliable multicast using multiple multicast groups. *ACM SIGMETRICS*, 1997.

[3] S. Paul et. al. Reliable multicast transport protocol (rmtp). *IEEE Journal on Communications*, April 1997.

[4] J. C. Lin and S. Paul. Rmtp: A reliable multicast transport protocol. *IEEE INFOCOM 96*, March 1996.

[5] R. Buskens, et. al. Reliable multicast of continuous data streams. *Bell Labs Tech. Journal*, 1997.

[6] R. Yavatkar, et. al. A reliable dissemination protocol for interactive collaborative applications. *Proc. ACM Multimedia*, 1995.