IEICE Electronics Express, Vol.7, No.21, 1635–1639

Customized Forwarding in application layer multicast for real-time services

Winam Kwon[a)], **Hojin Ghim, and H. Yoon**

Korea Advanced Institute of Science and Technology,

335 Gwahak-ro (373-1 Guseong-dong), Yuseong-gu, Daejeon 305-701, Republic of Korea

a) *winam.kwon@samsung.com*

Abstract: We propose a customized forwarding (CF) scheme for tree-based application layer multicasts to support real-time interactive multimedia services. We show that CF reduces residual loss rate and play pause frequency significantly.

Keywords: tree-based application layer multicast, error recovery

Classification: Electron devices, circuits, and systems

References

[1] Y.-H. Chu, et al., "Enabling conferencing applications on the internet using an overlay multicast architecture," *Proc. ACM SIGCOMM Computer Communication Review*, San Diego, CA, USA, pp. 55–67, Oct. 2001.

[2] Z. Fei, et al., "A proactive tree recovery mechanism for resilient overlay multicast," *IEEE/ACM Trans. Netw.*, vol. 15, no. 1, pp. 173–186, 2007.

[3] G. Tan, et al., "Improving the fault resilience of overlay multicast for media streaming," *IEEE Trans. Parallel Distrib. Syst.*, vol. 18, no. 6, pp. 721–734, 2007.

[4] NS2, [Online] http://www.isi.edu/nsnam/ns/

[5] D. Pendarakis, et al., "ALMI: An application level multicast infrastructure," *Proc. 3rd USNIX Symposium on internet Technologies and Systems*, San Francisco, CA, USA, pp. 49–60, March 2001.

[6] W.-P. Yiu, et al., "Lateral error recovery for media streaming in application-level multicast," *IEEE Trans. Multimedia*, vol. 8, no. 2, pp. 219–232, 2006.

[7] Ongoing IP network conditions, [Online] http://ipnetwork.bgtmo.ip.att.net/pws/index.html

1 Introduction

Tree-based application layer multicasts (ALM) can provide real-time interactive services such as network conference, network broadcasting, and network education over wide-area networks [1]. In tree-based ALMs, when a parent leaves or gets fault, its descendants should rejoin and experience significant

DOI: 10.1587/elex.7.1635
Received September 13, 2010
Accepted October 02, 2010
Published November 10, 2010

amount of data losses. This also occurs during periodic ALM tree adjustment [2, 3].

We categorize losses which can occur during rejoin or tree adjustment into three types. First, detection loss is the loss of data packets occurred while a member detects that its parent leaves or gets fault. Second, rejoin loss is the loss of data packets occurred before a member receives the first data from its new parent since it initiates rejoin process. Third, phase loss is the loss of data packets detected when a member receives the first data from the new parent. It occurs because members are receiving and forwarding different data packets as per their tree delays from the source. The amount of the difference between a rejoining member and its new parent turns into phase loss at the end of the rejoin process. It gets more bulky as the tree height gets longer, because the diversity of the tree propagation delays to members gets greater. It may give serious impact on the service quality especially for real-time interactive services over wide area networks because both playback and end-to-end deadline are so short. We propose a customized forwarding (CF) to remove phase loss, and show that CF enhances the quality of service significantly by simulation with NS-2 [4].

2 Customized forwarding

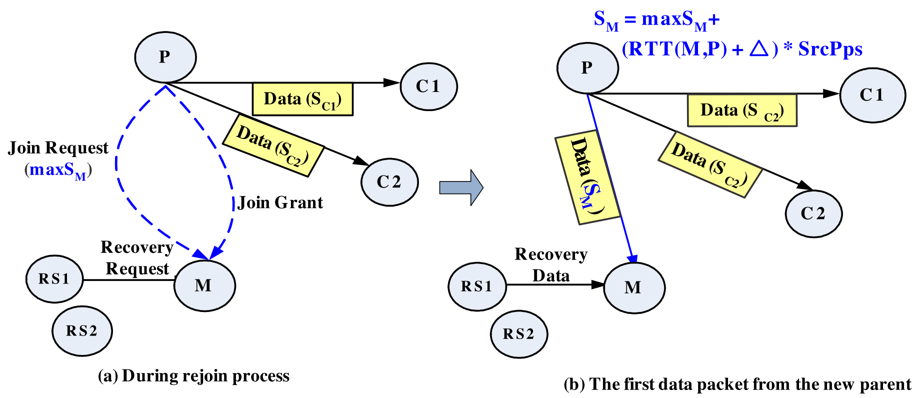


Fig. 1. Customized forwarding

In traditional forwarding (TF), a parent replicates and forwards the data of the same sequence number to children. In constrast, in customized forwarding (CF), a parent can forward to different children the data of different sequence numbers as per their requests.

Fig. 1 describes the steps a parent calculates the sequence number from which a rejoining child expects to receive. During rejoin process, M (rejoining member) lets P (new parent) know the maximum sequence number of data it has received (called $maxS_M$) by enclosing it into rejoin request message, as shown in Fig. 1 (a). As shown in Fig. 1 (b), when P starts forwarding to M, it decides from which sequence number to forward to M (called S_M). S_M is chosen as the sequence number that M will expect to receive at the time this first data arrives at M as follows: $maxS_M + (RTT(M,P) + \triangle) * SrcPps$,

DOI: 10.1587/elex.7.1635
Received September 13, 2010
Accepted October 02, 2010
Published November 10, 2010

IEICE Electronics Express, Vol.7, No.21, 1635–1639

where $RTT(M,P)$ is the round-trip time from M to P, $\triangle$ is the elapsed time since receiving rejoin request message from M, and $SrcPps$ is the source's encoding rate in packet per second (pps). RTT is known by the underlying overlay networks in advance and $SrcPps$ is known when M joins the service for the first time. Finally, M receives the first data of S_M from its new parent and continues receiving the subsequent data.

P keeps forwarding M as many data packets as its bandwidth allows until the biggest sequence number that P has received is forwarded. Consequently, in CF, M doesn't experience phase loss, whereas in TF, it suffers phase loss from S_M to $maxS_P$ and should recover them from its recovery servers.

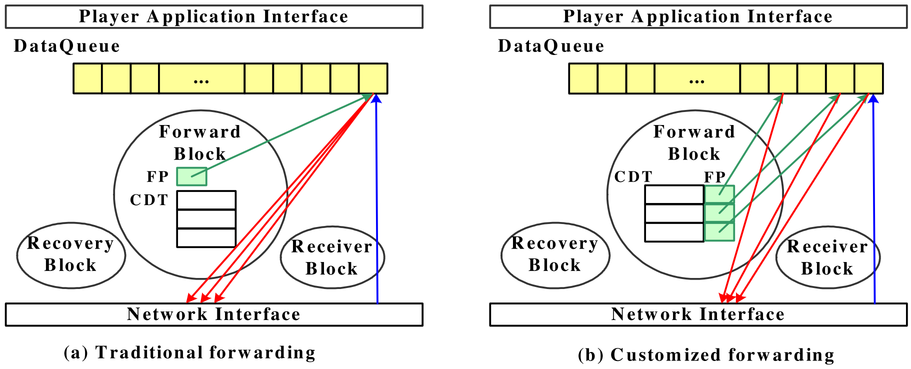


Fig. 2. The block diagrams of CF and TF

We adopt an fundamental ALM infrastructure, ALMI [5], which most ALM schemes with error recovery imply. Fig. 2 (a) shows its simplified block diagram of ALMI, with focusing on data forwarding and error recovery. The main resource components are a data queue (called DataQueue), a pointer to an element in DataQueue (called FP), and a child descriptor table (called CDT). DataQueue maintains not only the received data but also their control information such as sequence numbers and time stamps. FP directs to the entry of DataQueue to forward to children next time. CDT maintains per-child information for data forwarding such as socket descriptor. The main functional components are ReceiverBlock, ForwardBlock, and RecoveryBlock. ReceiverBlock receives incoming data packets and stores them into DataQueue. RecoveryBlock requests lost data packets to recovery servers and transmits the requested data by other members. ForwardBlock forwards data packet directed by FP to all children.

TF maintains a common pointer, FP, for all children, whereas CF maintains a separate one for each child in the child descriptor table as shown in Fig. 2 (b). In CF, ForwardBlock forwards different data packets to different children with separate FPs in the CDT. FP is initialized when forwarding the first data to a newly joined child as shown in Figure 1. Therefore, the additional space overhead of a member is the memory size for m pointers, where m is the number of children. And, the additional time overhead of a member to multicast one data packet is m memory read operations to get FP. However, the amount is negligible because the number of children of a

DOI: 10.1587/elex.7.1635
Received September 13, 2010
Accepted October 02, 2010
Published November 10, 2010

IEICE Electronics Express, Vol.7, No.21, 1635–1639

member is not large.

3 Evaluation

3.1 Setup

We have evaluated CF with NS-2 [4]. Networks are generated by GT-ITM's transit-stub model like in [3, 6] and the propagation delay of transit nodes follows uniform distribution (0.05, 0.4) in seconds with referencing the real IP backbone measurement in [7] and that of stub nodes follows uniform distribution (0.001, 0.03) in seconds. A member can reside in a stub node. We have evaluated on various topologies with 50, 100, 200, and 300 network nodes and 20, 50, 100, 150, and 200 members. Source's encoding rate is 64, 128, 256, 384, and 512 Kbps in constant bit rate. Evaluation setup models a general video/audio conference that has large audience with infrequent interaction. All members have the same amount of outgoing bandwidth. When a member joins or rejoins, it is attached to ALM tree so that tree delay can be minimized [1]. Tree adjustment is not included. We assume that a member can succeed in rejoining in the round-trip time to the new parent with proactive mechanisms [2]. The recovery scheme for detection losses and rejoin losses is CER [3], LER [6], or none. Members pause playing when the buffer has no more data to play and resume playing after playback deadline. Other details are listed in Table I.

Table I. Simulation setup and metrics definition

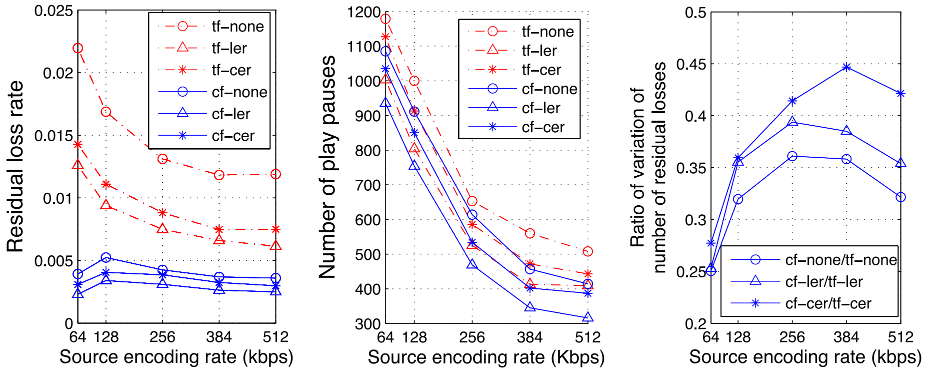
Simulation Setup	
Data packet size	1 Kbyte
End to end deadline	1 second since the source transmits data
Playback deadline	0.3 second before playing the data
Maximum tree degree	3 children per member
Recovery bandwidth	30% of source's encoding rate per member
Member's fault rate	1 fault per second in average
Metrics definition	
No. of play pauses	number of times that data playing is paused
No. of residual losses	number of times that reading data to play fails
Residual loss rate	$\frac{\text{(number of residual loss)}}{\text{(number of received data)} + \text{(number of residual loss)}}$

3.2 Results

Fig. 3 (a) and (b) show that CF reduces residual loss rate and frequency of play pause significantly. This is mainly because in CF, phase loss is effectively prevented by new parents, whereas in TF, phase loss occurs and should be recovered from recovery servers. Recovery from recovery servers may fail because members may have stale information about recovery servers and the number of recovery retrials is limited by short playback deadline. It's also noted that as source's encoding rate becomes slower, residual loss and play pause frequency becomes higher. The reason is that playback deadline

DOI: 10.1587/elex.7.1635
Received September 13, 2010
Accepted October 02, 2010
Published November 10, 2010

IEICE Electronics Express, Vol.7, No.21, 1635–1639



(a) Residual loss rate (b) Number of play pauses (c) Ratio of variation of number of residual losses

Fig. 3. Simulation results of 100 network nodes and 50 members

expires even with a small number of losses when the interval between data packets is long.

Fig. 3 (c) shows that the variation of the number of residual losses in CF decreases under 45% of that in TF. Consequently, with CF, members can be provided with more stable quality of service.

4 Conclusion

We have categorized data losses during rejoining and proposed a new forwarding scheme to remove phase loss. The evaluation results showed that the proposed scheme reduced residual loss rate and the frequency of play pause significantly.

DOI: 10.1587/elex.7.1635
Received September 13, 2010
Accepted October 02, 2010
Published November 10, 2010