Computer Networks 54 (2010) 1418–1425

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

A transmission control SCTP for real-time multimedia streaming

Kyung-Hoe Kim [a], Kwang-Min Jeong [b], Chul-Hee Kang [a], Seung-Joon Seok [c,*]

[a] Dept. of Electronics and Computer Engineering, Korea Univ., Anam-dong, Sungbuk-gu, Seoul 136-701, Republic of Korea
[b] Samsung Electronics, Co. Ltd., 416 Maetan-dong, Yeongtong-gu, Suwon-si, Gyeonggi-do 443-742, Republic of Korea
[c] Dept. of Computer Engineering, Kyungnam Univ., 449 Wolyong-dong, Masan-si, Kyungnam 631-701, Republic of Korea

ARTICLE INFO

Article history:
Received 31 October 2007
Received in revised form 3 September 2009
Accepted 30 November 2009
Available online 11 December 2009
Responsible Editor: Prof. Qian Zhang

Keywords:
SCTP
Multimedia streaming
Partial reliability

ABSTRACT

Multimedia streams over the Internet have a strict playback delay time, i.e., multimedia data arriving after the playback time cannot be played by the receiver and they are discarded. In this paper, we introduce a transmission control SCTP (TC-SCTP) which has a transmission control sub-layer (TCSL) in which the multimedia streaming server determines whether data can be played in the receiver before sending the data and decides whether to send data messages or not. In addition, TCSL employs differentiated retransmission policy depending on the type of multimedia. We evaluate the performances of SCTP, partial reliability (PR)-SCTP, and TC-SCTP via ns-2 simulator. The simulation results demonstrate that the TC-SCTP protocol can decrease the amount of transmissions and increase the video decodable ratio, compared with other protocols.

1. Introduction

Transmission control protocol (TCP) is a transport layer protocol that is widely used for various Internet applications, e.g., World Wide Web (WWW), email, and file transfer. However, the retransmission scheme in TCP is not appropriate for multimedia streaming applications because it can increase the end-to-end delivery latency. Therefore, user datagram protocol (UDP) is a commonly used transport layer protocol for multimedia streaming applications. UDP does not employ any flow control schemes in response to network congestion, and therefore it can burden other users on the network and, ultimately, lower its service quality [1].

To overcome this limitation, real-time transport protocol (RTP) and real-time control protocol (RTCP) can be adopted on the top of UDP in multimedia stream applications [2,3]. That is, the RTP/RTCP layer supplements the functions of UDP by correcting out of order data and controlling the volume of data transmitted by senders for congestion control. However, these actions rely on periodic reports between the sender's and receiver's RTCP, they cannot control by packets nor respond actively to network conditions.

A new transport layer protocol, stream control transmission protocol (SCTP) has been proposed by Internet Engineering Task Force (IETF SIGTRAN Working Group. Although it was first developed for telephone signaling, it is gradually expanded into a general-purpose transmission layer, and it has been standardized as RFC 2960 in 2000 [4]. Like TCP, SCTP provide reliable service and flow control mechanisms. In addition, similar to UDP, it can support unreliable transmission [5]. SCTP can provide multi-stream and multi-homing services for a single connection. In particular, it can differentiate the level of reliability provided to messages, which is called SCTP partial reliability (PR-SCTP) [6]. PR-SCTP has the function of setting the reliability level for specific messages. The preset reliability level is used to determine the timing when the retransmission of specific data message is stopped. The function can be effectively applied to traffic containing different types of data, such as I, P, B frames in MPEG streaming applications. However, PR-SCTP attempts transmission at least once even for messages that do not require any retransmission

* Corresponding author. Tel.: +82 11 255 7693; fax: +82 55 248 2554.
E-mail addresses: kyunghoe@widecomm.korea.ac.kr (K.-H. Kim), sjseok@kyungnam.ac.kr (S.-J. Seok).

1389-1286/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2009.11.014

due to the stringent delay constraint. In addition, if retransmission is given up, it has to send a Forward-TSN Chunk to the receiver.

Recently, multimedia streaming protocols are required to control its sending rate in response to the congestion condition of network [9–12,17–19]. It is because non-responsive streaming to network congestion, such as UDP, starves TCP flows under congestion condition. In this paper, we propose to use SCTP's congestion control for multimedia streaming.

Real-time data must reach the receiver within the tolerable playing time. Thus, a protocol like TCP, aiming for full reliability, cannot be used because a retransmission delay can exist. However, its alternatives, such as UDP and the combined UDP + RTP, do not attempt retransmission for lost packets. In this case, even if there is a chance for retransmission, based on the maximum playing time, they do not retransmit lost data. Considering this problem, the present study explains a multimedia data transmission protocol which is a modification of PR-SCTP.

In this paper, we introduce a sub-layer, called as transmission control sub-layer (TCSL), between PR-SCTP and the application layer. The TCSL sub-layer maintains as many buffers as the number of data types, and stores messages from the transmission server depending on the type. After that, TCSL of TC-SCTP evaluates the necessity of transmission before transmission. If the message does not need to be sent, it is simply removed. Compared with PR-SCTP with the fixed reliability level, TC-SCTP can dynamically consider network condition in each packet transmission.

The rest of this paper is organized as follows. Section 2 reviews existing video data transmission protocols. Section 3 describes a proposed scheme for streaming multimedia data, called TC-SCTP. Section 4 evaluates the performance of TC-SCTP and Section 5 concludes this paper.

2. Existing protocols for streaming multimedia data

As mentioned above, PR-SCTP can differentiate its retransmission service for each message. Differentiated retransmission can be achieved by applying different retransmission limits depending on the data type, e.g., delay-sensitive data and delay-tolerable data. The differentiated retransmission policy in PR-SCTP can be applied to MPEG-type video transmissions. In order to differentiate the reliability of each frame in MPEG video transmissions, [7,8] restrict the number of transmissions according to the type of frame to be transmitted. For example, only one retransmission is allowed for I frames and P frames with stringent end-to-end delay bound because additional retransmissions are not necessary due to excess playing time. On the other hand, no retransmission is provided for B frames because losses of B frames do not have any significant effects on the quality of video streaming. In this mechanism, each frame data is retransmitted only once or not retransmitted at all. However, because all data have to be transmitted at least once, if the network is congested, data may be queued for a long period of time, which may in turn lower the overall performance of the video transmission. Moreover, when the retransmission is given up, the information on the corresponding packet should be delivered to the receiver using forward TSN, which takes a long one-way trip time, and may cause faster retransmission than required, resulting in a reduction of the size of cwnd (congestion window). Furthermore, if the volume of network traffic is small and delay characteristics are favorable, the quality of video transmission is lower than the case where the number of retransmissions is not restricted.

Like [7,8], Media-SCTP [5] also restricts the number of retransmissions according to the frame reliability. Media-SCTP assumes that MPEG frames belonging to the same GOP have the same playout deadline. Because each GOP is composed of the same number of frames, the playout deadline appears at a regular interval. The receiver's frame dropping filter compares the current time with the playout deadline for each retransmitted frame. The difference between the two points of time is the time allowed for retransmission. Because it takes around 0.5 round trip time (RTT) in average for the retransmitted frame to be delivered from the sender to the receiver, the frame dropping filter compares the time allowed for retransmission with 0.5 RTT. If the allowed time is longer than 0.5 RTT, the corresponding data is retransmitted. If not, the receiver sends the Forward-TSN chunk to the sender so that the corresponding data chunk should not be retransmitted. On receiving the Forward-TSN, the sender takes an action to prevent the time-out of the corresponding data and the consequent reduction of the congestion window. That is, Media-SCTP's receiver side decides data retransmission only when the playback time of the data remains more than 0.5 RTT regardless of the type of the data frame. In the proposed TC-SCTP, however, the differentiated retransmission restricts the maximum number of retransmission in the consideration of remaining time of data frame until its playback and of data frame type.

Ahmed and co-workers [13,14] propose a sub-layer above RTP which is referred to as MP-RTP. Like Media-SCTP, MP-RTP's receiver side has responsible for retransmission request and also sender side considers frame's lifetime before sending low priority frame. However, high priority frame is retransmitted through all available path simultaneously. In this paper, proposed TC-SCTP control sending multimedia data according to data types. TC-SCTP does not request any receiver's message for retransmission and based on SCTP's multi-streaming properties, not multi-homing's.

Wu et al. [15] introduce six special issues about Internet streaming video. Application-layer QoS control issue of these issues is classified into congestion control and error control. In particular, this paper introduces a Delay-Constrained Retransmission as an application layer error control method. This Delay-Constrained Retransmission is similar mechanism to Media-SCTP [5,13,14]. In this scheme, receiver's application requests packet retransmission and sender's application retransmits the lost packet when arrival in time for display is expected.

The RaDiO in [16] determines which packets to select, when to transmit and how to transmit to minimize decoded video distortion at the receiver. This algorithm is a general scheme for data transmission and is mainly based