Integrating RTP into the World Wide Web

Randa El-Marakby, David Hutchison

Computing Department,
Lancaster University,
Lancaster LA1 4YR, UK
e-mail: <randa,dh>@comp.lancs.ac.uk

Abstract

In this paper, we outline an architecture and an implementation which enables the real-time delivery of continuous media within the WWW, monitors QoS of data transmission, and conveys minimal session control information to all participants in an audio or a video session. We describe our current implementation of the Real-time Transport Protocol (RTP) and its associated control protocol RTCP embedded into the WWW environment. We used the Web browser, Netscape, thus allowing Netscape users to access remote real-time audio and video media using RTP. An adaptive rate-based flow controller is also included to control and reduce network congestion and to adjust the bandwidth dynamically.

1. Introduction

Today, real-time applications transferring continuous media, i.e. digital audio and video, over the Internet, impose new Quality of Service (QoS) requirements in terms of bandwidth, delay, error rate, jitter, and throughput. The real-time applications must deliver data within a bounded time frame. They cannot operate using the retransmission method, used by TCP, in case packets arrive too late at their destination. Therefore, RTP has been designed and standardised by the IETF (Internet Engineering Task Force) for real-time continuous media transmission over the Internet.

The Internet is characterised by supporting heterogeneous computers. The WWW has been successful in supporting the Internet environment thus ensuring the interoperability. However, the protocol used by the WWW, HTTP (HyperText Transport Protocol), and the underlying transport protocol TCP are not suitable for continuous media delivery in real-time [FSVW96]. Currently there is ongoing research for the transfer of continuous media in real-time through the WWW and its associated browsers.

Because services on the Internet are best-effort, it provides no guarantees for resources nor does it guarantee performance measures, e.g. maximum delay, or minimum bandwidth [BT94]. To overcome this problem, QoS feedback information is sent from receivers back to senders. In our RTP/RTCP implementation, we utilise RTCP feedback reports to adapt the data transmission rate, from the source, thus reducing network congestion, as the bandwidth requirements of the source change dynamically [BT94].

This paper is divided as follows: In section 2, an overview of RTP with its associated control protocol RTCP is presented. In section 3, our RTP implementation, embedded into Netscape, is described followed by the conclusion.

2. Real-time Transport Protocol (RTP)

RTP provides end-to-end network transport functions which are suitable for applications transmitting real-time continuous data over multicast or unicast networks [SCFJ96]. These functions include: identification of payload data type, sequence numbering, timestamping, and monitoring QoS of data transmission.

RTP consists of two protocols: RTP and RTCP.

2.1 RTP Data Protocol

RTP is designed for the real-time data transmission. An RTP data packet consists of a header followed by the payload data, e.g. a video frame/several audio samples. Some important fields of the header are:

- *Payload type:* the format of RTP payload, e.g. H.261 for video streams
- *Timestamp:* the instant the data packet is generated. It is used in synchronisation and in jitter calculation.
- *Sequence number:* It is incremented by one for each data packet sent. It can be used by the receivers to calculate packet loss and out-of-sequence packets.

2.2 Real-time Control Protocol (RTCP)

RTCP is based on periodic transmission of control packets to all participants in the audio/video session. Its main control functionality is as follows [SCFJ96]:

- *Monitoring QoS:* RTCP provides traffic monitoring by transporting some QoS measurements as a feedback on the quality of the data distribution.

The feedback is performed by two types of RTCP reports: Sender Report (SR), and Receiver Report (RR). Both reports contain performance statistics on number of packets lost, highest sequence number received, jitter, and other delay measurements to calculate the round-trip time delay. The SR report provides more statistics summarising data transmission from the sender, e.g. timestamps, count of RTP data packets, and number of payload octets transmitted.

- *Identification:* RTCP conveys participant's identification information to all participants in a RTP session using RTCP SDES (Source Description) reports. SDES reports contain information as: CNAME[1] , user name, e-mail, telephone number, location, and application name. In addition, this SDES information is used in "loosely controlled" sessions to convey minimal session control information to participants that enter and leave without membership control or parameter negotiation.

3. Real-time Continuous media Implementation

In this section, we first outline a general overview of our implementation of RTP/RTCP. Then, we proceed to describe the implementation environment and the details of the continuous media real-time transmission in the WWW environment. Then, we explain our implementation of RTCP reports which are utilised by the adaptive rate controller.

The following are some definitions of terms used, as shown in fig. 1:

- *Server/Sender:* The continuous media server that runs continuously and sends audio/video file to be played back at the receiver
- *Client/Receiver:* The continuous media client that receives the audio/video file, from the continuous media server, to be played out.
- *Netscape Client:* Client running Netscape on same machine where the continuous media client application runs.
- *Netscape Sender/Server:* Sender running Netscape on same machine where continuous media server application runs.
- *WWW Server:* The Web Server that stores and runs the CGI programs (to be explained later). In our implementation, the Web Server is on a different machine than the sender and the client.

3.1 Overview

RTP is integrated within a multimedia C application that sends real-time digital audio and video streams over the Internet. It works over UDP/IP. The application is embedded within Netscape thus allowing Netscape users to access remote audio and video and to play them back in real-time. In order to ensure interoperability all over the heterogeneous Internet environment, the HTTP protocol of the WWW is used as mentioned in the paper of Fry et al [FSVW96]. However, the HTTP protocol is only used at start up time.

RTP audio/video data packets are sent in real-time, initially at a constant bit rate, from sender to receiver to be played out at the receiver. During the data transmission, RTCP control reports are sent periodically. The RTCP RR (Receiver Report) is sent only from receiver back to sender. It contains measurement statistics that are used as feedback on the quality of data distribution. SDES (Source Description) reports are sent from sender to receiver and vice versa. It carries identification information followed by participant's image.

An adaptive rate control scheme is also implemented for the dynamic adjustment of the bandwidth requirements of the continuous media application [BDS96].

3.2 Implementation Environment

Our experimentation was first carried out in a LAN environment, with two Sun Sparcstations running Sun OS 4.1.4 and connected through a 10 MBit/s Ethernet. Afterwards, our environment was extended and we were transmitting from an HP700 series machine at Edinburgh to our Sun Sparcstation at Lancaster via SuperJanet/SMDS .

[1] CNAME (Canonical Name) is a persistent unique identifier used to identify each participant.

3.3 General Architecture

Fig. 1 shows the general architecture with Netscape running at the client and the sender. In addition, there are two independent continuous media components of the application, one at the client and the other at the sender. Moreover, the WWW server, storing and running CGI programs, is running at a third machine[2] .

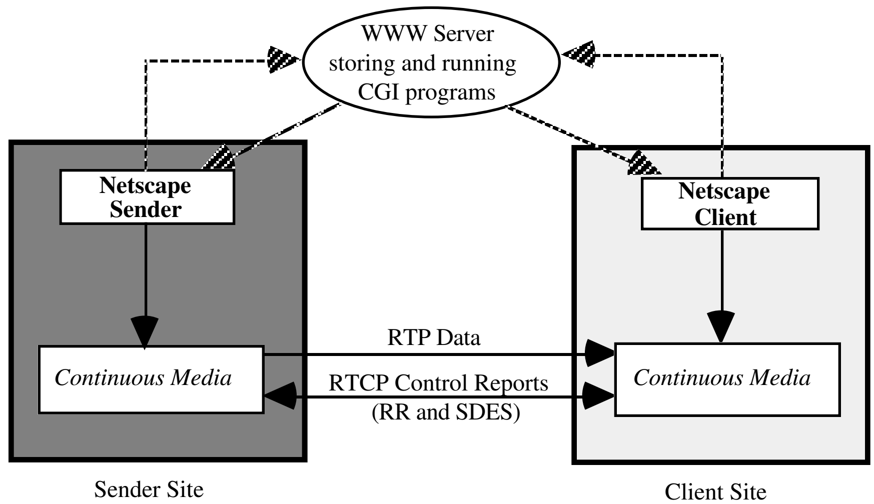

Fig. 1 Architecture of real-time continuous media transmission embedded into Netscape

CGI, Common Gateway Interface, is an interface between a Web Server and external programs. It allows those programs to process HTML forms of WWW browsers, e.g. Netscape. Then, it sends a response back to the browser answering its query.

3.4 Transmission of Real-time Continuous Media in the WWW environment

The following summarises the transmission of real-time continuous media embedded within Netscape:

- Fig. 2 shows the Netscape page of the server/sender. The sender chooses the type of RTP session (audio, video, or both) by clicking at the appropriate choice. The server runs continuously.

Fig. 2 Netscape Page of Sender/Server

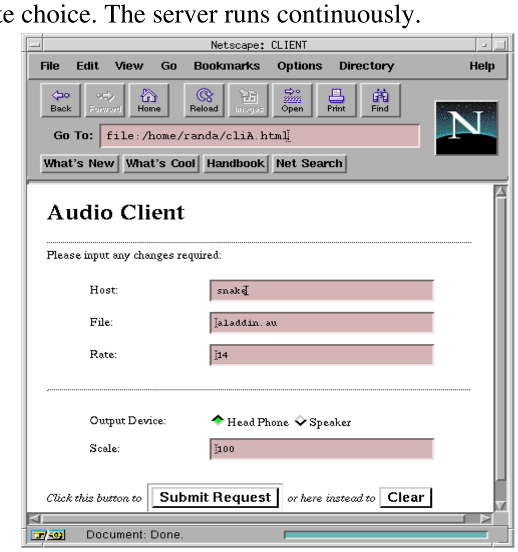

Fig. 3 HTML Form of Audio Client

- On the other side, the client inputs some data in the HTML form, i.e. name of host to connect to, required audio file to be played, and transmission rate. Fig. 3 shows the client HTML form for an audio session. The client can also choose the output audio device (speaker/headphone) and can change the volume. Then, the client clicks at the "Submit Request" button.
- If the client wants a video file to be played out, the video HTML form will appear. The form contains also Host, File, and Rate.

---

2 The Web Server can run on any machine including the Netscape Server or even the Netscape Client machine. Our Web Server is chosen to run on a seperate third machine to avoid performance problems [BH96].

- The Netscape client's request is submitted to a CGI program, running at the WWW Server, where it is executed in real-time by the Web daemon. The request contains mainly parameters in the HTML form.
- The output of the CGI program is sent back to the client
- The continuous media client then starts
- After the sender chooses the type of RTP session in the first step, a request is sent to a CGI program, at the WWW Server, and the output is delivered back to the sender.
- The continuous media server then starts

The main idea of the continuous media transmission is that the server runs continuously waiting for a connection from client(s) for transmitting audio/video. The client process requests the transmission of a certain audio/video file from the server process to be played out at this client as shown in the following:

- The continuous media client creates a RTCP channel to deliver some initial control data to the continuous media sender in a RTCP application specific packet that contains the following:
  - Required rate of data transmission
  - Name of requested audio/video file to be played out
  - Data port number created at the client so that the sender can send the data packets to.
- After receiving the initial control data, the sender forks a process to get the required file. The data packets are sent, at the specified rate, to the data port at the client to be played out. Every data packet contains a RTP header that is stamped with the current time and contains a sequence number.
- After arrival of the data packets at the client, audio packets are added into a circular buffer, before being played, in order to smooth out jitter. Video packets are passed to the modified MPEG player program [YGHS96], implemented at the University of California in Berkeley, that decodes and displays them on the screen.
- During data transmission, RR reports, carrying QoS measurements, are transmitted from client back to sender which are used by the adaptive rate control scheme to adjust the rate to the current state of the network. Moreover, SDES reports are sent on the RTCP channel in both directions.

During the real-time data delivery, the user at the client side is not aware that he is not using the Netscape client directly for the accessing of the audio/video file, as mentioned in [BH96]. The continuous media components of our implementation are transparent to the users at the sender and at the client sites.

As mentioned earlier, the HTTP protocol is only used at the starting time, when the user, having filled the HTML form, submits his request for transmission to the WWW Server. The real-time continuous media transmission is actually done using RTP not HTTP, as mentioned by Bolot and Hoschka [BH96].

3.5 RR Report

In our implementation, after arrival of some data packets at the receiver, some traffic QoS measurements are computed. The measurements include:

- *Total Number of Packets Lost* since the beginning of transmission.
- *Fraction of Packets Lost* in the interval since sending the previous RR and the current one.
- *Transit Time Delay* which is the difference between a packet’s RTP timestamp, at the sender, and the receiver’s clock at arrival time [SCFJ96]. It is used in the rate control scheme and in jitter calculation.
- *Interarrival Jitter* which is the mean deviation (smoothed absolute value) of the difference in packet spacing at the receiver compared to the sender for a pair of packets [SCFJ96].

A RR (RTCP feedback Receiver Report) carries the QoS parameters back to the sender via the RTCP channel. RR reports are transmitted periodically during the RTP session to monitor QoS.

3.6 SDES Report

In our implementation, RTCP Source Description (SDES) reports are transmitted periodically from sender to receiver and vice versa. A SDES report contains identification items about its sender, e.g. CNAME, sender name, e-mail, telephone number, and application name used. The SDES report can include also the geographical location and a notes item describing the current state of the sender [SCFJ96]. In addition, a new item which is the image of the sender is also sent as part of the identity information.

As shown in fig. 3, the SDES information with the sender's image sent, are displayed on the Netscape page of the other participant who received the SDES report. The page is refreshed automatically everytime a new SDES report and the image arrive, and the new information is redisplayed. The refreshing of the Netscape page is done from within our continuous media application. Periodic refreshment of the page enables the receivers of the SDES report to see a slowly animated image of the sender.

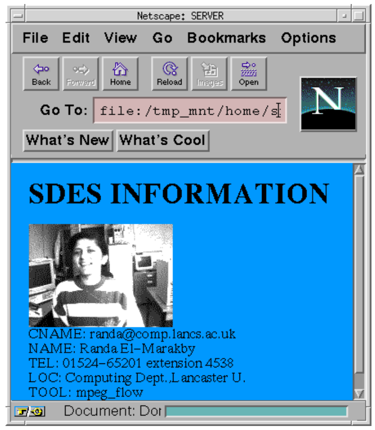

Fig. 3 Netscape page, at the server, showing identification and image of client

Items in the SDES report should be optional except the CNAME item. The transmission rate depends on how much bandwidth is used for control information transported by RTCP.

3.7 Adaptive Rate-based End-to-End Flow Control Scheme

Because the Internet is a best-effort service, it doesn't provide guarantees for resources nor does it guarantee performance measures, e.g. maximum delay, or minimum bandwidth [BT94]. In the implementation, RR feedback reports are sent periodically from receivers back to senders where they are used by the sender in order to adapt the data transmission rate to the current state of the network. The adaptive rate controller uses the packet loss parameter as an indicator of network congestion. The rate decreases if packet loss, detected during the interval between receiving the current RR and the previous one, exceeds a certain threshold. Afterwards, if packet loss became tolerable, i.e. within acceptable boundaries, then the rate is increased again, gradually, by a small fraction. The rate can vary between a pre-defined minimum value and the value specified by the user at start up time.

Beside dynamic control of the bandwidth, the adaptation scheme does not require special support from the network, e.g. admission control, resource allocation, etc. [BT94].

4. Conclusion

In this paper, we presented a proof of concept about the possibility of the integration of real-time transmission of audio/video into the WWW. We described our RTP/RTCP implementation, embedded in Netscape, that performs real-time transmission of continuous media over the Internet. Netscape ensures interoperability and is user friendly. RTCP provides QoS monitoring and identification information. An adaptive rate controller, that makes use of RTCP feedback reports, is implemented in order to adjust the rate of data transmission to the current state of the network thus reducing congestion.

References

[SCFJ96] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "RTP: A Transport Protocol for Real-time Applications", RFC 1889, January 1996.

[BH96] Jean Bolot, Philipp Hoschka, "Sound and Video on the Web", 5th WWW conference, Paris, May 1996.

[YGHS96] N. Yeadon, F. Garcia, D. Hutchison, D. Shepherd, “Filters: QoS Support Mechanisms for Multipeer Communications”, To appear in the IEEE Journal on Selected Areas in Computing (JSAC) forthcoming issue on Distributed Multimedia Systems and Technology, 3rd Quarter 1996.

[BT94] Jean-Chrysostome Bolot, Thierry Turletti, “A rate control mechanism for a packet video in the Internet”, Proc. IEEE INFOCOM '94, June 1994, Toronto, pp. 1216-1223.

[FSVW96] Michael Fry, Aruna Seneviratne, Andreas Vogel, Varuni Witana, "Delivering QoS Controlled Continuous Media on the World Wide Web", 4th International IFIP Workshop on Quality of Service - IWQoS96, Paris, March 1996, pp. 115-126.

[BDS96] I. Busse, B. Deffner, H. Schulzrinne, “Dynamic QoS Control of Multimedia Applications based on RTP”, Computer Communications, vol. 19, no. 1, January 1996.