Structure and technical solutions for Web-conferencing system

Alexei Machovikov¹,²

¹Department of Informatics and Computer Technologies, National University of Mineral Resources 'Mining', 2, 21st Line, St-Petersburg, Russia, 199106
²Department of Software Development, Bradon Technologies Ltd., 5225 Orbitor Drive, Unit 22, Mississauga, Ontario, Canada, L4W 4Y8
amachovikov@gmail.com

Abstract: In this paper the structure of web-conferencing system named SAViiMeeting is described. This system contains server part which consists from web-portal, database, database-proxy, media-server, http-proxy, sip-gate and mobile-proxy and clients for Microsoft Windows, Microsoft Windows Phone, BlackBerry and Android. Also some technical solutions which have been developed for this system such as method of heterogeneous data synchronization, frame independent audio codecs, method of real-time data transfer through HTTP and UDP-based reliable data transfer protocol are considered. SAViiMeeting’s audio codecs have been noted in 2009 Unified Communications TMC Labs Innovation Award Winners and SAViiMeeting system in whole has been named as Unified Communications Product of the Year in 2008 and 2009.

Keywords: web-conferencing systems, frame independent audio codecs, real-time data transfer protocols.

Introduction

Web-conferencing systems such as Cisco WebEx and Adobe Connect Pro are increasingly used now for production management and distance learning. One such system has been developed by Bradon Technologies Ltd. (Canada) in 2006-2010. This system was called SAViiDesk and later renamed to SAViiMeeting (‘SAVii’ is abbreviation for Synchronized Audio Video Interactivity through Internet which is consonant with the word ‘savvy’). In this paper I would like to present the system structure and some technical solutions which have been designed by me.

System Structure

Like the overwhelming majority of Internet network services, SAViiMeeting system uses client/server technology.

The client end of the system is represented by applications for Microsoft Windows and Microsoft Windows Phone. There are also lightweight Java versions of clients for mobile devices of BlackBerry and Android types. Landline, mobile and IP telephones can be used as ‘only-audio’ clients.

The server end of the system is represented by a combination of servers, proxy and gateways operating under the control of Linux or Microsoft Windows. It includes:

- Web-portal of SAViiMeeting system which is responsible for user authorization, conference planning, distribution of invitations to conference participants, run client applications, etc.
- Database management system (DBMS) used for storage of information about SAViiMeeting clients, conferences held by them, etc. The fifth version of system can use MySQL or DB2.
- Database proxy is used for creation of an encrypted channel for communication between DBMS and other components of the server end. Such organization allows installation of the transmitting components at the customer’s site.
- Media-server is the main transmitting component of the system. Clients connect to the Media-server and send data through it. Interaction of clients and other components of the server end with the Media-server is realized with using of secure UDP-based protocol. The task of the Media-server also is archiving of conferences and reproducing them at the clients’ requests.
- HTTP-proxy that used to connect clients that failed to connect to the Media-server directly. The main reason of connection errors is caused by presence of a firewall which blocks the UDP protocol. The client operates with the HTTP-proxy by the HTTP protocol which is allowed by all firewalls.
- SIP-gate which is a gateway to IP-telephony provider. IP-telephony provider receives calls from ‘only-audio’ clients that are connected to it. If an IP-telephony client is based on SIP/RTP-protocols, it can connect to the SIP-gate directly.
- Mobile-proxy which serves light Java-versions of clients for mobile devices.

The fifth version of SAViiMeeting system can serve three types of customers: public (individual), small and medium business (SMB) and enterprise. The information flows in the SAViiMeeting system are shown in Fig. 1.
The work with SAViiMeeting system can be described in the following way. The conference host enters the Web-portal with using of the name and password received upon registration. Using tools implemented in the Web-portal, the host plans the conference by choosing the time for it and invites participants. New participants are added by e-mail addresses. Each participant receives an invitation e-mail with the link to run the client application and the telephone number with the pin code for connecting via the telephone. At the appointed time the host runs the client application for Microsoft Windows and waits for participants to connect. Participants also run its clients and enter the conference. Those who have a computer with a soundcard, a microphone and speakers/earphones run a full client. Later the host can transfer his rights to such participants. Others either run a client in the ‘deaf-and-mute’ mode and use the telephone or a mobile device. The client for Microsoft Windows Phone is fully functional. Java-clients for BlackBerry and Android can only receive images from the host’s desktop. After a sufficient number of participants have connected to the conference, the host starts the presentation. He turns on his microphone and a video-camera, places presentation materials on his desktop, marks them with a frame and enables screen capture. If required, he has a possibility of drawing with a marker inside the captured field. Participants viewing the host’s presentation and can speak with him and other participants if their microphones are not blocked. The participant can attract the host’s attention by ‘raising his hand’ (pushing a special button). If the host considers it necessary, he can transfer the right of leading the presentation to one of the participants. The conference can be archived for further viewing.

The desktop of host’s computer with SAViiMeeting application interface is shown in Fig. 2.

Some technical solutions

During development of SAViiMeeting systems some serious technical problems have been solved. I would like to describe four such problems and their solutions.

Problem 1. Synchronization of heterogeneous data.

In the web-conferencing system there is a problem of synchronous playback of heterogeneous data such as speech of the host, its video image, images from his desktop, etc. To solve this problem a lot of different solutions were proposed. First, many try to pass all the data in a single stream and create a tight synchronization between heterogeneous data. However, this method of transmission has serious problems with interruption of the data flow in case of bad communication channel. These interruptions the stronger, the greater the amount of data transmitted per unit of time. So, the bulk data which do not require very stronger continuity (for example, images from the desktop) brings a negative effect to the transfer of data that require it, but take a small amount in the stream (for example, speech). Another more correct way is to pass different data in different flows. In this case synchronization between data streams is organized, as a rule, with the using of time stamps. This method is realized, for example, in RTP (Real-time Transport Protocol). But the transport layer there is no information.
about the physical meaning of transmitted data, so data which require continuity (speech) can be stopped before the arrival of the data which do not require continuity (image from the desktop).

So, I came to the idea of ‘natural’ data synchronization. Each data stream I try to transfer and play in real time. The data which are late are discarded. Request for the retransmission of the lost and corrupted data is not done. Accordingly, all the data are played simultaneously. Many years of successful operation of the SAViiMeeting system is confirmed the validity of this idea.

**Problem 2.** Audio compression for transmission over a channel with losses.

In 1999-2002, we with Kirill Stolyarov are developed some low speed frame independent audio codecs for 2400, 4800, 9600 and 19200 bps (Machovikov et al., 2001) which can be used in communication channels with losses without additional methods of data protection, such as FEC (Forward Error Correction). These codecs are used in SAViiMeeting system.

Consider briefly the algorithms of our codecs.

**2400 bps codec.** The recorded speech signal (8 kHz, 16 bit) divided by 180 ms windows is passed through third-order high-pass and low-pass Butterworth filters with cutoff frequencies 100 Hz and 3900 Hz. Filtered window is divided into 8 sub-windows for 22.5 ms. Each sub-window is divided into five frequency bands {0-500, 500-1000, 1000-2000, 2000-3000, 3000-4000} with using of sixth-order Butterworth filters. For the first four ranges an indication tone/noise is calculated. If the first two bands have the tone indication the pitch period is calculated. Also for each band the gain coefficient is calculated. Additionally, for the entire sub-window the jitter coefficient and linear spectral pairs are calculated. These parameters are packed into a bit buffer and transmitted to the remote side. At the remote side on the basis of pitch period, jitter coefficient, gain coefficient and tone/noise indicator the excitation signal is created. After interpolation, this signal is fed to the synthesis filter based on linear prediction coefficients calculated from the linear spectral pairs. The output of this filter is the synthesized speech signal.

**4800 bps codec.** Processed as well as 2400 bps codec signal is divided into six 30 ms sub-windows. For each sub-window the linear prediction coefficients are calculated and converted into linear spectral pairs. Then, the original signal is passed through the inverse filter, based on the linear prediction coefficients. Obtained signal is divide into two parts. For the first part (180 samples) parameters of adaptive codebook initialization are calculated. By the second part (1260 samples) parameters to generate adaptive and algebraic codebooks are calculated. These parameters are packed into a bit buffer and transmitted to the remote side. On the remote side on the base of parameters for adaptive and algebraic codebooks the excitation signal is synthesized. This signal is passed through a filter based on the linear prediction coefficients. The output of this filter is obtained synthesized speech signal.

**9600 bps codec.** This codec is generally similar to the 4800 bps codec, but it does not use an algebraic codebook. Instead, amplitudes and positions of pulses which substitution into the excitation signal allow minimize the deviation of the synthesized signal from the source are calculated. The algorithm of sequence of pulses calculation is our ‘know-how’. Using this sequence of pulses can significantly increase the quality of synthesized speech.
19200 bps codec. The algorithm of this codec is completely identical to the 9600 bps codec algorithm. More high speed of this codec allows increase the number of optimal pulses and, as result, of synthesized speech quality.

Computational cost of our codecs in MIPS (Million Instructions Per Second) are presented in Table 1. The table shows that the computational costs are low, especially in the decoder. This means that our codecs should be used in the web-conferencing systems where the parallel decoding of multiple flows is required.

In 2009 our codecs were shown as Unified Communications TMC Labs Innovation Award Winners (Dinan, M., 2009).

<table>
<thead>
<tr>
<th>Codec</th>
<th>Encoder, MIPS</th>
<th>Decoder, MIPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2400 bps</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>4800 bps</td>
<td>6</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>9600 bps</td>
<td>8</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>19200 bps</td>
<td>10</td>
<td>&lt; 1</td>
</tr>
</tbody>
</table>

Recently, due to the improvement of communication channels quality, some wideband speech codecs for 16 kHz speech signal have been developed. Sampling increasing results in a significant increase in intelligibility and naturalness of speech. The classic representative of this group is the codec SILK, which is used in the Skype system. In 2010 we have developed our own wideband codec for 50 kbps which provides excellent quality of the reconstructed speech with extremely low computational cost (Encoder - 0.5 MIPS, Decoder - 0.2 MIPS). This codec we will use in the next version of SAViiMeeting system.


All real-time data transfer protocols (standard as RTP and special as media transfer protocol for SAViiMeeting system) are working on the base of UDP protocol. But unfortunately, UDP protocol is always blocked in corporate networks. In these networks only TCP protocol and based on it HTTP/HTTPS are allowed and the data is transferred only through caching HTTP proxy. In this regard, I am faced with a problem: how to organize the real-time data transfer with using of a protocol that is not designed for this? I have proposed the following solution.

To transfer each type of data its own TCP-connection is established. Data transmission is organized with using of POST and GET methods of HTTP protocol. In case of bad communication channel using the TCP protocol rapidly increases the data delivery delay and leads to loss of real time. To eliminate this effect, new TCP-connection is created with a specific time interval (in the SAViiMeeting system – 10 sec.). This new connection is used to continue the data transfer. Data that have not been transferred through old connection is removed. This method can be used in the presence of corporate HTTP-proxy. In this case, no-caching option must be indicated in the request and response headers. HTTPS protocol also can be used.

Exploitation of the SAViiMeeting system is showed an efficiency of the proposed method. The actual link quality, of course, seriously degrades with the deterioration of channel quality, but the external channels in the corporate sector have good quality.

Problem 4. Data transfer through poor communication channels.

In contrast to the corporate sector, in private sector there are practically no problems with UDP protocol blocking, but a problem of a channel quality is often present. Of course, if the channel quality is completely very bad, the real-time communication is impossible. However, users with such channels should be able to download and view archives.

In contrast to the ‘live’ conference, playback does not require the transmission of data in real time and it can be implemented using the TCP protocol. However, the data transfer speed with using of this protocol seriously degrades with deterioration of the communication channel quality. It is caused by the fundamental flaw of the TCP protocol, when in the delivery confirmation packet sent only the expected number of the next packet. So, if for example, the packet number five was lost and six, seven and eight were not lost, all of them will be re-transmitted, because the transmitter does not have information about their successful delivery.

To eliminate this problem and reduce useless data transfer through poor communication channel, I have proposed a proprietary protocol based on the UDP which provides a guaranteed data delivery. The essence of it is that for each received packet a confirmation with information on eight previous packets is sent. If the packet is delivered, it sets a flag 1, if not - 0. Transmitter analyzes this information and resends only undelivered packets.

Testing of this protocol is showed that, in comparison with the TCP protocol, a real data delivery speed provided by them is decreased significantly slowly with deterioration of channel quality.

Conclusion

In this paper I described the structure of SAViiMeeting system and give a brief description how to use this system. The system contains server part which consists from the following services:

- web-portal;
- database;

http://aict.itf.ltu.lv
The client part is presented by clients for Microsoft Windows, Microsoft Windows Phone, BlackBerry and Android. Also some technical solutions which have been developed for this system, for example:

1. method of heterogeneous data synchronization;
2. frame independent audio codecs;
3. method of real-time data transfer through HTTP;
4. UDP-based reliable data transfer protocol.

SAViiMeeting’s audio codecs have been noted in 2009 Unified Communications TMC Labs Innovation Award Winners (Dinan, M., 2009) and SAViiMeeting system in whole has been named as Unified Communications Product of the Year in 2008 (Viscusi, 2009) and 2009 (Harrison, 2010). The system is available for to thirty days of trial use and exploitation. The link is http://www.saviimeeting.com.

References


