2006

The conception of development of the scaled multifunctional software systems of voip telephony

The main idea, architecture, technological description and ways of development of the complex system of sound processing, converting and transmission, which is directed at the application as a multifunctional solution in the services of VoIP telephony are presented. The described conception implies development and usage of a set of program components while finding an efficient and flexible solution of the processing of audio data of various formats, voice, recording and playback of sound data in real time. The basis of the report is the experience of practical development of systems for the VoIP telephony and their improvement from the point of view of satisfying of changing demands of service providers and end users. The technologies, instrumental means and standards of the development used in the system are considered. The problem area, existing solutions, areas of practical approbation, applicability and system development are dealt with.

PhD A. Lipanov
Master Degree A. Kochetov

1. Introduction

In current conditions VoIP telephony as an alternative to analogue telephone lines assumes an all-embracing scale giving the Internet users all over the world a possibility of conducting free of charge telephone calls of computer-computer format or to call the analogue telephones at lowest rates.

Nevertheless the existing technological basis and standards used in VoIP telephony allow to broaden considerably the services spectrum and give the consumer and the provider-companies wider possibilities.

The positive experience of IP-services providers with the use of broadband connections in Internet such as Skype and others whose innovations attracted millions of users all over the world indicates that the further development of infrastructure is quite justified and promising.

At present most commercial systems on the basis of IP-records of transmission are directed towards rendering voice communication services to the subscribers. The systems of automatic processing of orders over telephone, automatized systems of technical support, systems toll-free of voice mail (free of charge calls for clients for the number 1-800), VoIP-conferences and others are carried out chiefly on the basis of expensive equipment. Nevertheless the described solution does not exclude usage of telephone equipment as a component of architecture. Moreover the developed system is oriented at the broadening of opportunities of infrastructure of IP-telephony systems, that is the technology of processing of in-calls and further control is in this case random and can include for instance the Cisco Systems equipment which has shown itself to advantage.

2. Short analysis of solutions existing

The absolute leader of the industry at present time is the Cisco Systems company, which gives a full set of telephone as well as switching equipment. Cisco Solutions have flexibility, scalability and reliability.

The company offers a complex approach to creating of the architectural systems with voice integration, video, including the net infrastructure, client places, server and user attachments.

The solution of the Cisco telephony consists of the following components: special-purpose digital IP-telephones, the operating server Cisco CallManager, voice locks for the attachment of IP-nets with the telephone net of the general use as well as user voice applications. The application area of the proposed VoIP-solutions embraces the user terminal equipment, items for the xDSL-lines, systems of allocated access giving the full processing of the traffic.

3. The conception of the proposed solution

The main purpose of the developed solution lies in the accomplishment of a step to the side from the usual IP-telephony, and to form an information communication system which allows to develop constantly new systems, directed at different kinds of services.

A significant feature of the proposed conception is that the main emphasis is made on the development of powerful and reliable software services using existing technologies, net architecture and so on. Such method is much more flexible and universal in comparison with the use of the equipment.

As the main task at this very point the Infostroy company has set the development of a ready-to-use scaled system which gives wide opportunities as to processing of the in-calls, connection of the analogue and IP-lines, converting different formats of audio-information and playback in analogue and digital telephone lines.

In general case the VoIP-net infrastructure can be built as directed only at switching of the Internet users or oriented at combining IP-telephony lines and analogue lines. It is obvious that the second case may demand the involvement of additional equipment or services of the IP-lock provider and analogue nets.

The simplest conception of the VoIP-infrastructure directed exceptionally at the IP-communications is presented in diagram form at figure 1.

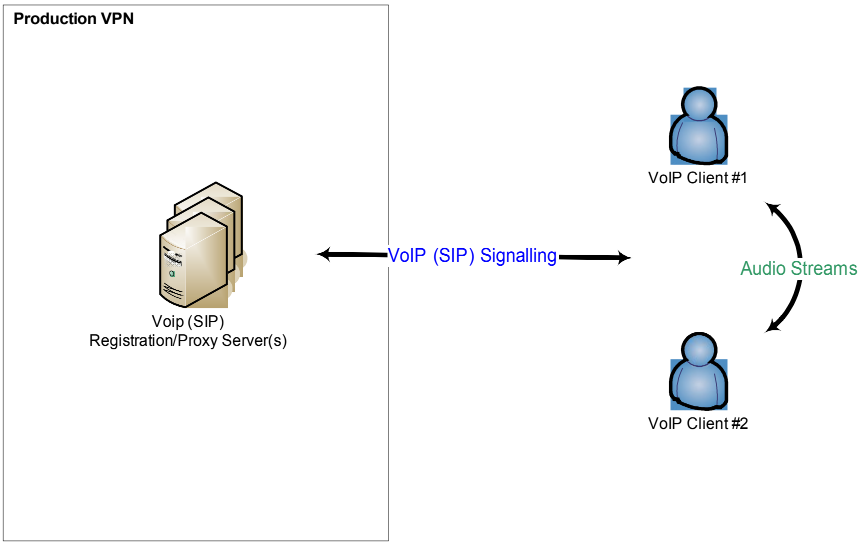

Figure 1. The scheme of a simplest VoIP-net

The expansion of the system with the purpose of the ensuring additional services of creating conferences, retransmission of conferences and calls in real time, recording of calls and conversion of audio data media-formats into VoIP voice stream require involvement of additional system components (figure 2). Physically it is a server or a cluster of servers supplemented with an attachment or a device which assign traffic load.

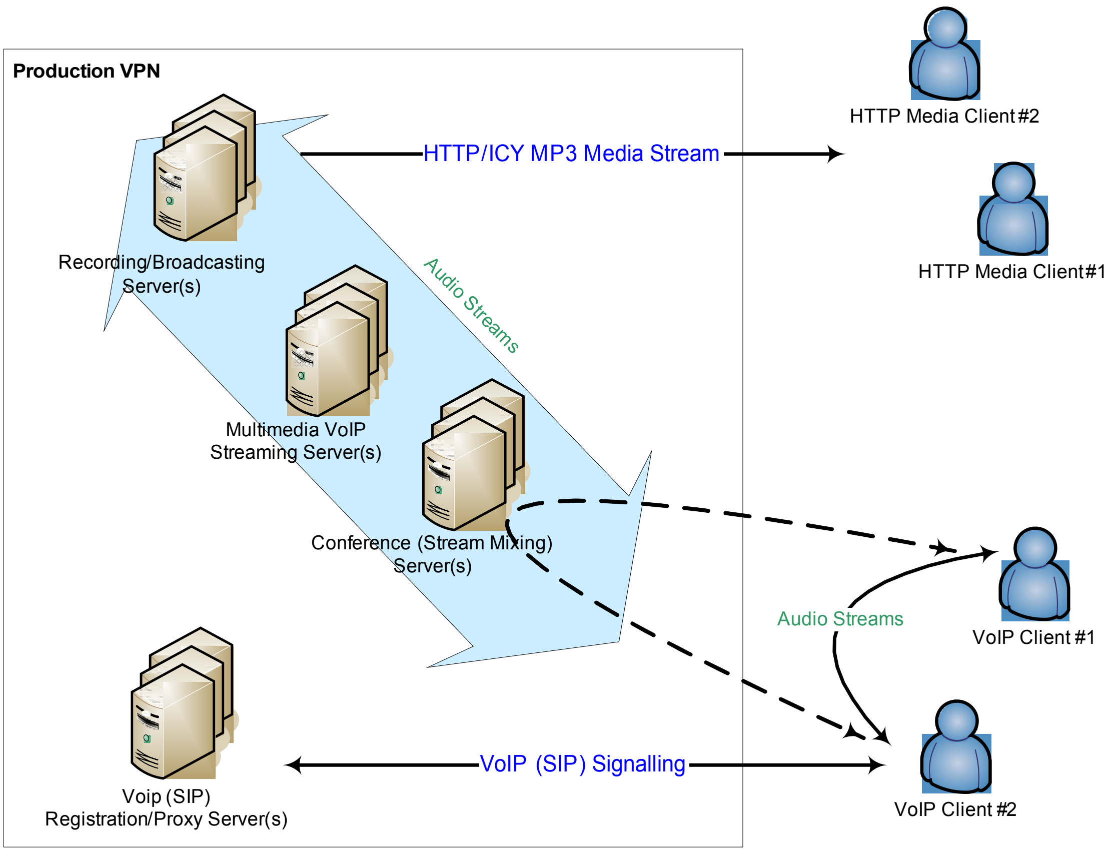

Figure 2. The scheme of the multifunctional system of the VoIP-telephony (switching of the Internet users)

To the scheme the following components are included:

- directly clients who gain access to the IP-telephony services by means of software IP-telephones.

- the server or a cluster of servers with the load tightrope walker conducting the client logging in the system and responsible for the initialization of calls. As a basic record of communication SIP (Session Initiation Protocol) is suggested. The absolute majority of software IP-telephones and VoIP proxy-and logservers support this record in this or another form.

- the server or a cluster of servers conducting the conversion of mp3, wav, asf formats into audio format which is used in the digital line of the IP-telephony and the further retranslation of the coming voice stream to the clients of the IP-telephony. For the most VoIP-attachments and equipment it is crucially important that within the limits of the set session (call) between two terminal points no additional sources of voice streams appear. Each server in the cluster uses concealment of the initial IP-address of the packet (the modification of the IP-packet headline in such a way that the receiver of the packet could identify the IP-address of the sender in another way than if the packet would not have been changed), that is why the IP-client does not feel the interference of the third side (beside two VoIP-clients who are in the state of an active session/call).

- the server or a cluster of servers for retranslation of the audio-stream of the active session (call) between two or more clients for a potentially unlimited number of the HTTP/ICY clients.

In case of the supporting of the digital and analogue lines switching the scheme of the multifunctional VoIP-net looks a little bit different (figure 3). Besides the mentioned above components there is also a switchboard or a bridge which conducts the converting of different types of signals.

If the telephone line is used mainly for transmission of the voice information and the IP-nets for transmission of data, the technology of the IP-telephony combines these nets by means of a device named a lock or a bridge. The lock is a kind of device into which from one side the telephone lines are switched and from the other the IP-net (for example, Internet). The device conducts converting from the analogue format into the digital and vice versa providing in that way switching of different types of the lines.

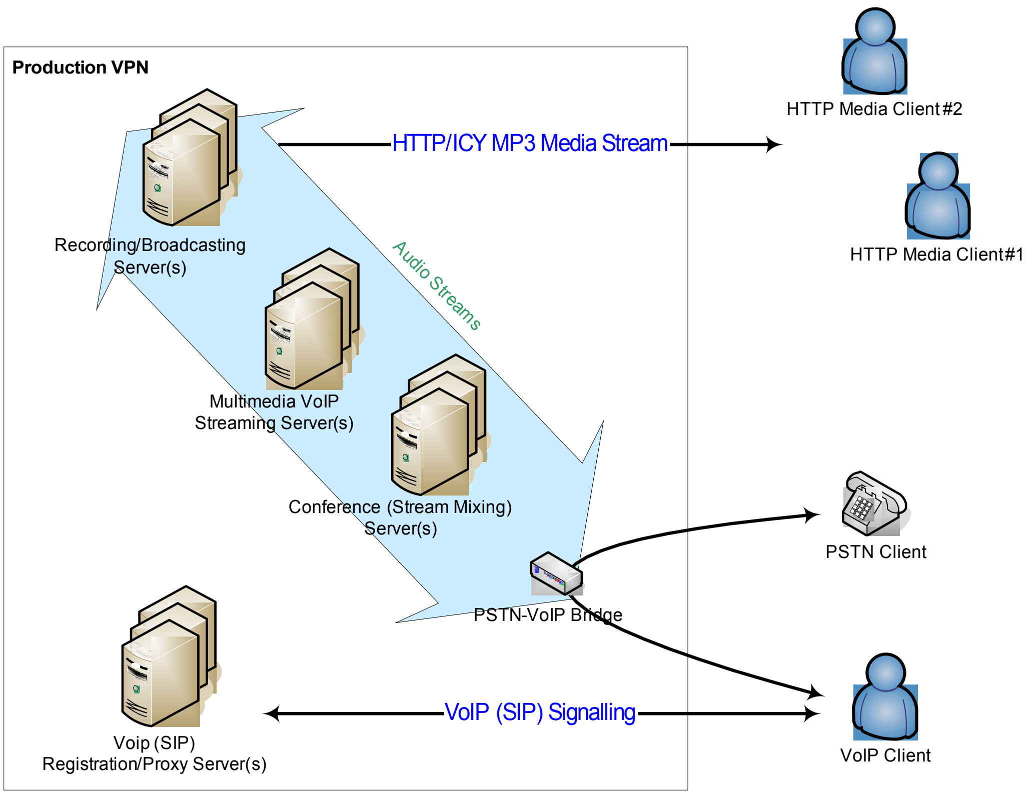

Figure 3. The scheme of the multifunctional system of the VoIP telephony (the switching of the Internet users and the users of the analogue lines)

Modern systems used by VoIP-operators must functionally and technologically be up to the barest requirements of the business demands. For implementation of the flexible intellectual routing of the calls into the analogue lines from the VoIP nets one can include to the architecture commercial SIP locks situated in different spots of the globe. This allows to choose the route of the call with minimal value and highest possible quality of communication what raises the profitability of the net.

4. Description of the solution

In brief outline the initialization of the session (call) and transmission of voice in the VoIP-net occurs in the following way. In case of the presence of VoIP and analogue line switching the in-call and the signal information from the telephone net are transmitted to the borderline net device named the telephone lock and processed by the special card of the device of the voice maintenance. The lock using the managing records retransmits the signal information to the VoIP-server which is at the receiving side of the IP-net.

After establishment of the connection the voice at the net in-device get digitized (if it was not digital), encoded in accordance with the standard algorithms ITU, such as G.711 or G.729, compressed, encapsulated into the packets and sent according to its intended purpose to the distant device with the use of the record stack TCP/IP or UDP. The IP-packets which arrive at the receiving lock are transformed again into a telephone signal.

4.1. Voip Registration/Proxy Server

VoIP-proxyserver provides transmission of the signal information to the VoIP-client. In addition to that the server is responsible for the registration of the VoIP-clients in the system and takes part in the process of session (call) initialization between two or more clients. Usually standard VoIP proxyservers do not give enough flexibility for the penetration into the pioneering business solution. For example, the alternative process of creating a conference may demand creating virtual clients on the VoIP-proxyserver for joining of several VoIP and analogue calls. In such cases a new server is developed or the existing one is improved in order to satisfy new requirements. As a basic record in the described architecture SIP (Session Initiation Protocol) is used which is a universal popular and perspective record of video and audio sessions control.

Extra services, such as server of media-data conversion into a digital VoIP signal or server of retransmission of calls and conferences, do not take part in the process of initialization and call control. That’s why for distant control of such services random protocols of upper level on basis XML are used.

4.2. Conference Server

In the network of VoIP telephony the creation of conference is a combination/mixing of several sessions (calls) on the level of session initialization signals as well as on the level of audio data. The processing of initialization signals of the session is the task of VoIP proxy server, performed on basis of protocol SIP. But in general case VoIP client may not support the conference mode. That’s why an alternative solution is used as an alternative. The possibility of conference request is then included (brought into) the system. For example, it can be implemented as web-page, where the existing users can request a new conference. As an output data in this case we have a number, which is temporarily registered on VoIP proxy server, and the calls on this server are automatically being connected to the conference.

The next is to connect several sound streams of VoIP clients into one. To solve the following problem we use a mixing service of VoIP protocols. The logic of its functioning is easy. The service orientates on protocol RTP (Real-time Transport Protocol) and performs application of signals from different channels. In this case for each conference we select a separate UDP port, which receives signal of all connected VoIP clients. After signals application, the total signal is transmitted to each channel of output audio stream.

We should take into account the fact, that each of VoIP clients can code its input audio signal differently. That’s why the service has support modules of voice codec G711 ulaw/alaw, G726, G729.

It is obvious that such a system must have the highest performance and stability. To provide the maximal number of conferences and clients processed by one mixing service, the conception Input/Output Completion Port is used, which helps to increase not only the performance level of network application but also to use the advantages of multiprocessor platforms in full.

4.3. Multimedia Streaming Server

Multimedia server solves the most resource-intensive tasks among all the components of the system. The main function of multimedia server is to convert media data in mp3, wav and asf formats into VoIP network format (RTP as well). The source of media data can be a file on a disk or on internet server as well as live MMS or ICY stream of media data.

In addition to conversion of media data into audio stream VoIP, server support back transformation. So that every RTP stream can be converted be the server into a stream of mp3 frames and sent as HTTP/ICY stream directly to a media client (Windows Media Player, Macromedia Flash, Real Player) or to the retransmission server.

The process of conversion of media data into RTP stream itself supposes the performance of the following sequential operations:

- to read the stream from a file. Internet resource or from media server;
- to decode the stream;

- to decrease the frequency of sampling;
- to code with voice codec;
- RTP transmission

The first two operations have standard solutions and are not difficult. Though the decoding of a stream provides in most cases the highest traffic on the service.

The decrease of sampling frequency is a reduction of a current frequency of media stream (usually 44100 Hz for mp3) to the sound frequency in VoIP network (usually 4000 or 8000 Hz).

Coding with voice codec is a crushing of voice signal with algorithms G711, G726, GSM and others. Most of the voice codec are developed with respect to the necessity of definite and simple implementation on microprocessors of the equipment. That is the reason why they are easy in program implementation and provide high efficiency.

The transmission in accordance with the specification RTP requires adhering of timeouts and sending of data with respect to the host limited buffer. It means that it is impossible to transmit media stream that is 5 minutes long in one second. From one point of view, this fact provides time periods, when each stream "stands", and from the other point of view, it requires to look for specific solutions at the realization attempt of the concept of data transmission on basis of Input/Output Completion Port (IOCP).

4.4. Recoding/Broadcast Server

Server of retransmission and recording works in connection with Multimedia Streaming Server. Its main task is to record output mp3 stream from Multimedia Streaming Server into mp3 file or to retransmit it to mane media clients.

5. Experience and analysis of results

The result of the work is a flexible system of processing of IP telephony signals and audio information, which will provide an effective and quick solution for the modern service provider that uses or plans to use signal and sound transmission through IP protocols and networks.

The set of allocated system components provides the following range of functional services:

- processing of initialization signals and signals of VoIP session (calls);
- recording of audio signals into mp3 format in real time;
- retransmission of calls and conferences in real time with support of media clients Windows Media Player, Macromedia Flash, Real Player;
- representation of mp3, wav, asf formats in real time with transmission on IP telephony channels;
- mixing servers of input VoIP sound streams to make conference;
- drivers of network transmission, which allow (without efficiency reduction) to reduce network traffic for those applications that do not support silence filter in output channel.

During the work on project solutions we have used the concepts and technologies, which have shown themselves to good advantage and which can be considered a starting point for new researches in this field:

- realization of high efficient services on basis of Completion Port concept of Microsoft Company, which allows supporting the biggest number of clients' connections and which is optimized for work on multiple processor platforms.
- full application of the opportunity to enlarge the system by implementing protocol stack of server control as system components, which allow configuring and complementing server clusters, dividing the load according to the round-robin principle, randomly of by minimal traffic.

Conclusion

The proposed solution architecture allows using system components in order to provide services for VoIP line users as well as for users of similar lines having bridge or gateway. Such services are: voice mail with recording and representation of messages in mp3 format, creating, recording and retransmission of conferences.

At present this platform is being used to create VoIP lines on basis of SIP and RTP protocols. In future we plan to enlarge the described solution and to direct it of a wider market spectrum. The fact, that mobile networks of new generation will be based on SIP, will simplify convergence of VoIP networks and therefore will provide new demands in this area and will stimulate new wave of development multifunctional systems of VoIP telephony.