A BRIEF INTRODUCTION TO VOIP AND TVOIP

Nafeesa Bohra, Hermann De Meer
Chair of Computer Networks and Communications,
Faculty of Mathematics and Computer Science, ITZ/IH,
University of Passau. Innstr – 43, D-94032, Passau, Germany.
{bohra,demeer@fmi.uni-passau.de}

Abstract

The Internet is evolving into a universal communication network and the next generation of Internet together with data will carry voice and video traffic over an IP infrastructure. The way the Internet is evolving in homes and in offices it has become an essential part of life as water and electricity. A revolution is occurring as organizations of all sizes begin to implement IP-based voice and video communication systems. VoIP, VoD (Video on Demand) and TVoIP are currently the most emerging applications that can be integrated over a single IP infrastructure. This paper gives a brief introduction about VoIP, VoD and TVoIP. What are the limitations and QoS requirements in order to integrate these services over an IP infrastructure?

1. Introduction

Communication services such as VoIP, VoD, and TVoIP have gained much attention during the recent years and the next generation of Internet requires an integration of all these services over an IP infrastructure. Integration of these services is not an easy task as these services not only require a high speed Internet connection such as broadband DSL (Digital Subscriber Lines) but also require a mechanism that supports a satisfactory QoS, because these services are highly sensitive to end-to-end delay and delay variations.

The significant increase in the use of packetized audio/video over wide-area, packet-switched networks are indeed due to the fact that the IP-based services like: VoIP, VoD, TVoIP are available at low-cost and also high quality audio/video is supported by today's Internet. At present, indeed the Internet is being used to carry voice conversations (VoIP) as well as for multimedia streaming (VoD/TVoIP) over an IP infrastructure but the fact is Internet is a best-effort service, it does not provide guarantee towards end-to-end delay and delay variations, which is an important criteria to fulfill QoS requirements. As far as the integration of packet voice/video and telecommunication system is concerned, major technological issues are:

- Voice/Video must support telecommunication system signaling in order to achieve seamless support of voice/video.
- CODEC (coder/decoder) should be adapted in order to provide high quality voice/video over an unreliable transmission, as different CODECs are used for different applications over an Internet.
- Due to the best-effort characteristic of the current Internet delivery service, the quality of service (QoS) is not guaranteed. Hence, how to make Internet to support acceptable QoS guarantee, is an important research area.
- Besides QoS and high-speed Internet connection, making an efficient use of available bandwidth is also an important issue as far as the integration is concerned.

In this paper an effort has been made to provide an introduction about what is meant by VoIP, VoD, and TVoIP and what are the limitations and QoS requirements when these services are integrated over an IP infrastructure. The paper is organized as follows. Section 2 gives a brief introduction about VoIP and TVoIP while section 3 discusses the broadband technology. In section 4 and 5 limitations and network protocols require for the integration of VoIP and TVoIP are discussed respectively. Conclusion is given in section 6.

2. VoIP and TvoIP

VoIP (Voice over IP) is the transfer of voice conversations as data packets over an IP network. Unlike traditional circuit-switched calls on the PSTN (Public Switched Telephone Network), in VoIP calls, the telephone connection is packet-switched. In a packet-switched environment, multiple computer devices share a single data network. The communication takes place by sending packets of data to one another, each packet containing addressing information that specifies the source and destination computers. The packets within a single transmission can take different paths from end to end across a data network.

Figure 1 illustrates the data flow for VoIP. In order to transport audio over a packet-switched network, audio samples are compressed by an encoder, followed by insertion of sequence number and time stamp into the compressed packet data. The packet data travels through the Internet, and is received by the decoder enforced with a dynamic buffer. An underlying characteristic of the current Internet is that a packet often experiences delay, delay jitter and packet loss.

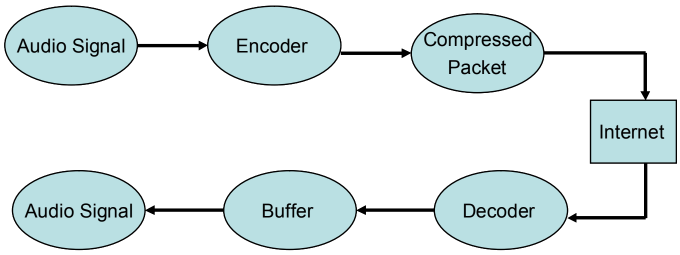

Figure 1. VoIP Data Flow

As far as VoIP call is concerned the audio portion of the call itself needs to be converted from analog to digital, cut into packets, sent across the network still in packet format, reassembled, and converted from digital back to analog. The conversion process from analog to digital and back into an analog form to be audible to human ear is accomplished by CODECs at either ends.

One of the advantages of VoIP [1] is that it can avoid the tolls charged by ordinary telephone service by utilizing fixed charge IP network services such as broadband. Currently, VoIP is based upon the Client-Server architecture. Existing VoIP client-server architecture is based upon IETF's (Internet Engineering Task Force) SIP (Session Initiation Protocol) [2, 3] or ITU-T recommendation H.323 [4] which typically employ a registration server for every domain. The user agents (or IP phones) of the users in the domain register their IP addresses with the server so that the other users can reach them. Scalability and reliability [5, 6] of such server-based systems are achieved using traditional redundancy and failover methods such as: DNS, IP address takeover, MAC address takeover or application layer switches. The majority of the system cost is in maintenance and configuration which is taken care of by a dedicated system administrator in the domain. It also means that quickly setting up the system in a small environment e.g. for a conference is not easy.

VoIP performs the following basic functions:

- Resource Location: Is responsible [7] for identifying and locating other users so that conversations can be established.
- Session establishment and management: Creates, controls, and terminates text or multimedia sessions between users and provide firewalls.
- Presence: Is the ability of users to monitor other users' connection status and to be notified when users join or leave.
- Application-level multicast: Group communication or multicast is the key feature of the distributed interactive applications in order to deliver data from the sender to the receiver.

Multicast is the mechanism that allows the distribution of information from one-to-many and can be used by applications like VoIP/TVoIP and VoD. Multicasting is based upon the concept of distribution trees where each source is referred as the root of the tree while the receivers are treated as the leaves of the tree. Routers replicate the packets at each branch of the tree and that point is called as the bifurcation point. It means that only one copy of the packet would travel over any particular link in the network and this makes multicasting extremely efficient in order to distribute the same information to multiple receivers. This is the greatest advantage of multicasting as it not only reduces the bandwidth usage but also reduces the source-processing load, which is not possible with unicast mechanism. Multicast can be [8] either best-effort or reliable. Best-effort means that the packet is delivered to all the group members but without any QoS guarantees i.e. a member may or may not receive the packet while reliable means that the packet is being delivered to all the group members with QoS guarantees. Details about multicast and multicast protocols are given in [9].

An IP – based platform also provides an advantage to integrate TV with other IP- based services like VoIP. TVoIP describes a system where a digital television service is delivered to subscribing consumers using the Internet protocol over a broadband connection. This service is often provided in conjunction with video on demand and may also include Internet services such as web access and VoIP where it may be called triple play and is supplied by a broadband operator using the same infrastructure. Triple play is an expression that is used by service operators which describes a consumer package including voice, data and video. A simpler definition would be television content that, instead of being delivered through the traditional format (i.e. via satellite, cable, and terrestrial networks), is being received by the viewer through the technologies used for the World Wide Web.

TVoIP can be divided into two broad categories. Live TV (multicasting) as well as stored VoD. Live TV broadcast in which the client may choose one out of the many available channels, whereas VoD is a stored audio/video application which a client selects out of many stored audio/video files and the file begins to play out few seconds after it is being received from the server. In both cases the playback requires either a personal computer or set-top-box connected to a TV. Video content can be delivered through MPEG2TS (MPEG over IP transfer system) via multicast. The underlying protocols used for TVoIP are IGMP (Internet Group Management Protocol) (RFC 988) for channel change signaling i.e. for live TV and RTSP (Real Time Streaming Protocol) [10] for video on demand. IGMP [8, 11] is particularly required for establishing IP multicast group. IGMP is the transport-layer multicast protocol based on class D addressing scheme. It is used to establish membership in multicast groups. In order to make an efficient use of network resources, the network sends multicast packets to only those networks and sub networks that have users that belong to the multicast group. IGMP allows terminals or hosts to show interest in receiving a multicast transmission. The working of IGMP is quite simple. There are three types of messages generated by IGMP. The join, query, and leave.

- Join: A host uses a report message to join new group.
- Query: It is used to discover which hosts are members of the given group.
- Leave: This message is sent when the host wishes to leave a given group.

Further details regarding IGMP are given in [9].

No matter whether it is a live TV or VoD the key issues are QoS and how to make an efficient use of available bandwidth. Simultaneous delivery of channels is an important concern in order to compete TVoIP with the traditional system of cable or satellite TV. At present, most television [12] services are carried over satellite, cable and terrestrial networks and users select one out of many available channels that are available over a particular frequency but in order to provide TV content over IP it means that there must be separate IP address for each channel. Assigning each channel a separate IP address is again an important issue.

Despite of the issues discussed above it is a fact that with the advancement of technology the TV of future will become more interactive and will involve electronic delivery over packet-switched networks over a high speed network like broadband DSL.

Broadband access to the Internet is getting more and more common. This increases the use of and the demand for new high-speed services such as VoIP/TVoIP and VoD. Since the Internet consists of many interconnected networks, data sent from one point to another may need to pass many of these networks. In each network the data passes through a number of network devices such as routers, which act as traffic crossings, on its way to the destination. Just like in motor traffic the flow through the crossings needs to be controlled. Similarly TV of next generation can be interactive just like Internet is today by the use of broadcasting over IP networks.

Broadband IP means a distribution system where the last-mile access into the home is via a high-capacity bi-directional IP network. In other words broadband IP is when bandwidth intensive services such as TV and VoD are carried into the room in IP datagrams.

In the past it was hard to say that TV content can be provided through Internet and this concept was mainly due to the slow dial-up download speed of the Internet. It is possible now because of the high-access Internet speed that can be provided by a broadband. No doubt, TVoIP is at its infancy and is growing with a brisk pace but with the broadband available to almost 100 Million house hold world wide therefore it is possible to have the triple play (VoIP, VoD and TVoIP) all in the room in IP datagrams.

For IP voice and video as described in [13] communications systems to work properly, the bandwidth should be as large as possible while the end-to-end delay, jitter and packet loss must be minimized. Lower end-to-end delay leads to a more satisfactory, natural-feeling experience while large delay values leads to unnatural conversations with long pauses or sentences. Large jitter values may cause network data packets to arrive in the wrong sequence causing jerky video or stuttering, popping audio.

It can be said that in order to provide hundreds of TV channels, audio and video on demand, VoIP, Internet access, much higher bandwidth will be required and it is only possible with high-speed Internet access also referred as broadband Internet access or simply broadband. Next section will describe the broadband.

3. Broadband Technology

High-speed broadband is nothing but the services that provide higher bandwidth than standard telephone services. Broadband allows users to access the Internet and other IP-based services such as: (voice, video and data) at significantly higher speeds than traditional modems. The rapid growth as discussed in [14], demands for high speed Internet/web access and voice/video for residential and small business customers have created a demand for broadband access. Broadband access is currently available through xDSL (Digital Subscriber Lines), cable, and BWA (broadband wireless access).

DSL is written as xDSL as it indicates that it is a growing family of standards and technology and each DSL technology provides different data communication speed capabilities. One of the most promising implementations is the ADSL (Asymmetric Digital Subscriber Line) [15]. It offers up to 8 Mbps over about 2km of copper, with typical downstream bitrates from 1.5 to 6 Mbps and upstream bitrates are from 64 to 640 kbps. The xDSL have all been designed [16] so that it:

- Must work over the existing twisted-pair copper telephone-lines.
- Must not affect customers existing telephone and fax machines.
- Must be faster than 56kbps.
- Must be always on, with just a monthly charge but no per-minute charge.

DSL can be implemented on existing infrastructure therefore it is possible to utilize DSL for VoIP, TVoIP and VoD for residential as well as businesses. DSL also offers high-speed Internet connectivity for interactive gaming, on-demand streaming audio/video entertainment, VoIP/TVoIP and also downloading large files in seconds instead of minutes. Although, DSL technology can be a right choice for voice, video and data, there are certain limitation which the packet network faces. These limitations are discussed in the next section.

4. Limitations

One of the major limiting factors is how to make an efficient use of available bandwidth as the Internet bandwidth is shared among all users; some sort of unreliability must be encountered during packet transmission. Hence, one may experience that the voice/video quality is acceptable at one time, but is terrible at another time. In conjunction with a hop-by-hop resource reservation protocol such as RSVP [17], end-to-end capacity can be set aside for real-time traffic.

The current dominant protocol for transmitting multimedia in packet-switched networks is RTP/RTCP [18]. RTP, the real-time protocol, is a generic mechanism for carrying data with real-time properties, e.g., audio and video. The headers of RTP provide the sequence number and time stamp information necessary to re-assemble a real-time stream from packets, but RTP still can not provide any QoS guarantee.

Network QoS can be evaluated by considering the four key parameters. They are:

- Bandwidth: It is measured in kbps or Mbps, is measured as the average number of bits per second that can travel successfully through the Internet.
- End-to-end delay or latency: It is the average time taken for a network packet to traverse the network from one end point to another.
- Jitter: It is the variation in the end-to-end delay of sequential packets.
- Packet loss: It is measured as the percent of transmitted packets that never reach the intended destination. Transmitted packets may be lost due to several reasons but the most important cause is the congestion in the network routers. When too many packets are simultaneously sent to a router, it will simply discard some packets, assuming that the application that has sent the packets will transmit it again.

Among all the factors given above the most important is the end-to-end delay or latency as far as voice is concerned as in order to ensure good voice quality, latency of about 200ms is acceptable. Although, since the future IP networks are going to provide information, entertainment and communication applications over an all IP infrastructure, which makes the user more demanding and the latency of 200ms would become too much. Hence, there is a need to reduce the end-to-end delay. The main causes as is discussed in [19], of delay are:

- Processing at switch/router.
- Transmission time or time to put packets online.
- Propagation delay, or the actual time it takes to pass between two switches.
- Variable delay or jitter introduced when packets get out of order and must be buffered and reordered before play.
- Speech encoding, compression and decompression.

Among other factors discussed above the major limiting factor in implementing the TVoIP is the bandwidth as video requires almost thirty times more bandwidth as compared to that of audio. Although by the use of DSL technology higher bandwidth can be achieved but still the solution is not effective and efficient.

After discussing the limitations another key factor in implementing the VoIP/TVoIP is the suite of protocols as several protocols would be required to cater these services. Next section discusses the protocols that will be required for VoIP and TVoIP.

5. Network Protocols to Support VoIP/TVoIP

ITU-T recommendation H.323 and IETF SIP are the currently used protocols for signaling and control for VoIP. SIP [20, 22] and H.323 [20, 21, 22] presents different approaches towards VoIP. H.323 addresses the traditional circuit-switched approach to signaling based on the ISDN Q.931 protocol and earlier H-series recommendations, whereas, SIP is a light weight Internet approach based on HTTP. SIP is an application level protocol and can run over TCP or UDP while H.323 runs over TCP as shown in figure

2. SIP describes how to set up Internet telephone calls, video conferences and other multimedia connections. Unlike H.323, which is a complete suite of protocols, SIP is a single module, but it has been designed to interwork well with existing Internet applications.

Figure 2 shows different protocols. TCP/IP is a connection oriented protocol and is not suitable for real-time applications.

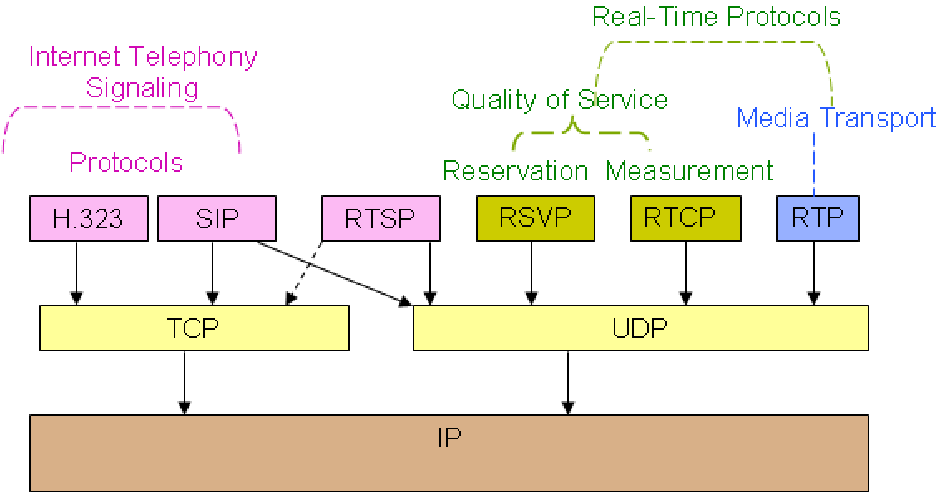

Figure 2. Protocol Suite

On the other hand UDP (User data Gram Protocol) is a connection less transport layer protocol which runs on top of IP and is used for real-time applications (voice, video) to avoid acknowledgements for packet loss as acknowledgement trigger undesirable retransmission of packets and increase network traffic (end-to-end delay) and thus effect the QoS. RTP [18] is an application layer protocol. RTP is an Internet protocol used for the transmission of real-time data (audio/video). Whereas RTCP (Real Time Control Protocol) [18] is an Internet protocol to monitor the quality of service and to convey information about the participants in an on-going session. RTCP works in conjunction with RTP. RTCP can be used to control stored media servers, or devices capable of playing and recording media from the server (i.e. pause/resume, repositioning of playback, fast forward and rewind).

Among H.323 [20, 22] and SIP, SIP would be the best choice due to its simplicity. SIP is basically an emerging signaling protocol for:

- Initiating,
- Managing, and
- Terminating data, voice, and video sessions.

Here the question arises why there is a need to use SIP?
Following are the reasons:

- To determine the capabilities of end points before start of media communication.
- Session management: Negotiates change in capabilities or requirements of end points during communication
- Session termination: Tears down communication, efficient network bandwidth management.

6. Conclusion

This paper gives a brief introduction about VoIP and TVoIP and different issues concerning the implementation of VoIP/TVoIP are discussed. From the discussion it is concluded that in order to implement VoIP/TVoIP it is necessary to select an efficient and affective communication media like DSL. Correct choice of protocols is required and it is also important to maintain QoS requirements in order to fulfill the demands of the user. Since, information, entertainment, advance telephony, interactive gaming, audio/video on demand, TV will all be delivered at the doorstep in IP frames which makes the user more demanding and would eventually makes a task more difficult.

References:

1. D. Tarrant, "VoIP: An Insight into a Progressing Technology", 4th Annual Multimedia System Conference, Electronics and Computer Science, University of Southampton, 2004.
2. H. Schulzrinne and J. Rosenberg, "Internet telephony: Architecture and Protocols: An IETF Perspective,"*Computer Networks and ISDN Systems*, Vol. 31, No. 3, pp. 237–255, Feb. 1999.
3. J. Rosenberg, H. Schulzrinne, G. Camarillo, A. R. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session Initiation Protocol," RFC 3261, Internet Engineering Task Force, June 2002.
4. J.Toga and J.Ott, "ITU-T Standardization Activities for Interactive Multimedia Communications on Packet Based Networks: H.323 and Related Recommendations," *Computer Networks and ISDN Systems*, Vol. 31, No. 3, pp. 205–223, Feb. 1999.
5. K. Singh and H. Schulzrinne, "Failover and load sharing in SIP telephony," Tech. Rep. CUCS-011-04, Columbia University, Computer Science Department, New York, NY, USA, Mar. 2004.
6. H. Bryhni, E. Klovning, and O. Kure, "A Comparison of Load Balancing Techniques for Scalable Web Servers," *IEEE Network*, Vol. 14, No. 4, July 2000.
7. D. A. Bryan, B. B. Lowekamp, C. Jennings, SOSIMPLE; A Serverless, Standards-Based, P2P SIP Communication System, AAA-IDEA 2005.
8. Z. Sun, Satellite Networking, Principles and Protocols, John Wiley & Sons Ltd., 2005.
9. N. Bohra and H. De Meer. "IP Multicast Protocols and Algorithms for TVoIP", In Transactions Automated Control Systems, Vol 1, pp 199–209, Tbilisi, Georgia, 2006.
10. H. Schulzrinne, A. Rao, R. Lanphier, "Real Time Streaning Protocol", RFC 2326, 1998.
11. S. Deering, "Host Extensions for IP Multicasting", RFC 1112, 1989.
12. R. Ollmar, H. Stephansen, White Paper, "The Complete Guide to TVoIP", 2002, www.tendbergtv.com.
13. E.B. Kelly, "Quality of Service in Internet Protocol (IP) Networks", International Communications Industries Association, Infocomm 2002.
14. H. Bolcskei, Arogyaswami, J. Paulraj, K.V.S. Hari, and Rohit U.Nabar, Stanford University W. W. Lu, Siemens-Infineon, "Fixed Broadband Wireless Access: State of the Art, Challenges, and Future Directions", IEEE Communications Magazine, Jan 2001.
15. I. Kompatsiaris, M. G. Strintzis, and I. Orfanos, "Using Asymmetric Digital Subscriber Line (ADSL) for Fast Internet and Multimedia Services", http://www.iti.gr/~ikom/papers/ICCE00.doc
16. A. S. Tanenbaum, " Computer Networks", 4th Edition.
17. R. Braden, L. Zhang, S. Berson, Jamin, "Resource Reservation Protocol (RSVP)", RFC 2205, Sep 1997
18. H. Schulzrinne, S. Casner, R. Fredrick, V. Jacobson, „"RTP: A Transport Protocol for Real Time Applications", RFC 1889, 1996.
19. U, Varshney, A., Snow, M. McGivern, and C. Howard, "Voice over IP", Communications if the ACM, Vol. 45, No. 1, Jan 2002.
20. M. Handley, H. Schulzrinne, E. Schooler, J. Rosenberg, "SIP: Session Initiation Protocol", RFC 2543, 1999.
21. H. Liu, and P. Mouchtaris, "Voive Signalling: H.323 and Beyond, IEEE Communication Magazine, Vol. 38, No. 10, pp 142-148, 2000.
22. H. Schulzrinne, J. Rosenberg, "A Comparision of SIP and H.323 for Internet Telephony", 1998, http://citeseer.ist.psu.edu/schulzrinne98comparison.html

КРАТКОЕ ВВЕДЕНИЕ В VOIP И TVOIP

Бора Нафесса, Де Меер Герман
(Германия, Университет Пассау)

Резюме

Интернет развивается в универсальную сеть коммуникации, и следующее поколение Интернета вместе с данными будет нести голос и видео-изображения по IP инфраструктуре. Интернет все шире примениется в домах и офисах, это стало существенной частью жизни, как вода и электричество. Революция происходит, поскольку организации начинают осуществлять голос на основе IP и видео системы коммуникации. VoIP, VoD (Видео по требованию) и TVoIP - в настоящее время большинство появляющихся апликаций, которые могут быть объединены по единной IP инфраструктуре. Дается краткое введение в VoIP, VoD и TVoIP. Каковы ограничения и требования QoS, чтобы объединить эти услуги по IP инфраструктуре ?