REAL TIME MULTIPLE CODECS SWITCHING ARCHITECTURE FOR VIDEO CONFERENCING

By

USMAN SARWAR

**Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science**

MAY 2008

REAL TIME MULTIPLE CODECS SWITCHING ARCHITECTURE FOR VIDEO CONFERENCING

By

USMAN SARWAR

Thesis submitted in fulfilment of the requirements for the degree of Master of Science

MAY 2008

ACKNOWLEDGEMENTS

بســــــــــــم الله الرحمن الرحيــــــــــــم
"اقرأ باسم ربك الذي خلق"

I would like to take this opportunity to convey my sincere thanks and deepest gratitude to my supervisor: Assoc. Prof. Dr. Sureswaran Ramadass, and to my co-supervisor Dr.Rahmat Budiarto, for all the help and valuable guidance provided to me during the preparation of this thesis. I consider myself privileged to have had the opportunity to work under their guidance.

Moreover, I would like to convey my appreciation to the MCS ver. 6 core team, all Network Research Group members, School of Computer Sciences, Institute of Postgraduate Studies, the university library and the Mlabs staff for their help and support.

Finally and most important of all, my sincere gratitude also goes to my parents; Dr.M. Sarwar, Riffat Bano, Song Jing Sheng and Dong Feng, my wife Shen, brother (Imran), sisters (Ainee and Ayesha), uncle (Dr.Tahir), aunt (Fatima) and friends (Ali, Arshad and many more). I thank them for their support, understanding, and encouragement during every step of my study and my writing of this thesis.

.

TABLE OF CONTENTS

CHAPTER TWO : LITERATURE REVIEW

CHAPTER THREE : RMCS ARCHITECTURE

CHAPTER FOUR : IMPLEMENTATION

CHAPTER FIVE : TEST RESULTS AND DISCUSSION

CHAPTER SIX : CONCLUSION

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

UDP	User Datagram Protocol
MCS	Multimedia Conferencing System
SIP	Session Initiation Protocol
MTU	Maximum Transmission Unit
TCP/IP	Transmission Control Protocol/Internet Protocol
LAN	Local Area Network
RMCS	Real Time Multiple Codec Switching

LIST OF PUBLICATIONS & SEMINARS

1.1 Usman Sarwar, “Presentation on RMCS architecture” in Prince of Songkla University, Thailand. 3 May 2007.

SENIBINA PENUKARAN PELBAGAI CODECS PADA MASA SEBENAR UNTUK PERSIDANGAN VIDEO

ABSTRAK

Trend terkini yang berhubung kait dengan khidmat video internet telah merubah dengan mendadak cara kita berfikir dan berkerja. Perkhidmatan-perkhidmatan ini telah memperluaskan ruang lingkup teknologi serta memberi peluang kepada perkembangan idea-idea penyiaran tanpa sempadan dan memendekkan jarak komunikasi secara geografi melalui telesidang video. Perkhidmatan-perkhidmatan ini harus dicapai dengan kaedah yang betul dan efisyen bagi mengelakkan ia menyekat penggunaan sumber terhad yang di kongsi bersama terutama yang berhubung kait dengan lebar jalur.

Penyelidikan yang dilaporkan dalam tesis ini mencadangkan satu arkitektur bagi menyelesaikan isu-isu telesidang video yang berkaitan dangan penggunaan lebar jalur dan sumber sistem secara optima. Codec video yang berlainan mempunyai kesan yang berbeza terhadap lebar jalur serta kualiti gambar. Arkitektur ini membenarkan penggunaan pelbagai codec untuk sistem video telesidang secara serentak. Pendekatan ini membolehkan pengguna telesidang video menggunakan codec video yang berbeza bagi melaraskan lebar jalur tujuan pelarasan berdasarkan sumber sistem dan kepelbagaian jenis rangkaian. Ia juga membolehkan penyelarasan secara dinamik dan pertukaran codec video bagi membolehkan penggunaan optimum lebar jalur rangkaian semasa sesi telesidang berlangsung. Arkitektur cadangan juga menambah codec video yang piawai dan bukan piawai kepada sistem telesidang video asas melalui pendekatan mudah pasang.

Arkitektur cadangan ini telah diintegrasi ke dalam sistem persidangan multimedia (MCS). Ia telah diuji dan di bandingkan dengan H.323 dan sistem berasaskan SIP.

REAL TIME MULTIPLE CODECS SWITCHING ARCHITECTURE FOR VIDEO CONFERENCING

ABSTRACT

Current trends related to internet video services have drastically transformed the way we think and work. These services have expanded the new scope of technology and have given opportunity for broadcasting ideas without any boundaries as well as shortened the vast geographical distances for communications through video conferencing. However these services must be accessed in a proper and efficient way in order to avoid clogging finite shared resources especially those related with bandwidth.

The work reported in this thesis proposes an architecture for resolving issues with video conferencing related with optimal bandwidth utilization and system resources. Different video codecs affects on bandwidth as well as picture quality. This architecture allows the use of simultaneous multiple codec support for a video conferencing system. This approach allows the video conferencing users to use different video codecs simultaneously to configure bandwidth according to heterogeneous network environments and system resources. It also allows for dynamic configuration and switching of video codecs during the conference to consume optimal available network bandwidth in video conferencing sessions. The proposed architecture also adds standardized and non-standardized video codecs to baseline video conferencing systems in a pluggable approach.

The proposed architecture was integrated into the Multimedia Conferencing System (MCS). It was tested and compared with the H.323 and SIP based systems.

CHAPTER ONE
INTRODUCTION

1.1 Introduction

Current trends in communication have massively involved the usage of video based services on Internet. Services like video conferencing, online TV, video blogging, video on demand are getting widespread which invokes the limitation of bandwidth available. Moreover there are various types of computers which are accessing these services. If these services are not properly accessed, it can cause disturbance to the Internet users. To cope with this situation, different types of video compression technologies have been developed.

1.2 Problem statement

This thesis focuses on the following three problems related with the video conferencing systems:

- Lack of extensibility in video codec architecture for video conferencing.

  Video conferencing system lacks extensible and pluggable video codecs architecture for adding video codecs.

- Single video codec support during the video conferencing session.

  Video conferencing system heavily depends on network bandwidth. Change in network bandwidth can disrupt the usage of the conferencing system. Current systems are restricted to their baseline architecture to use a single preferred video codec during the conferencing session; this may get affected in the case of change in bandwidth. Hence video conferencing systems are not able to configure dynamically according to bandwidth.

- Different video codec users cannot communicate with each other

  Currently there are various types of devices which are capable of doing video conferencing but with different video coding capabilities. Due to limitation of using single video codec for all the participants in a session can restrict diversity of video conferencing usage.

1.3 Objectives of this thesis

This thesis proposes a real time multiple codec switching (RMCS) architecture for video conferencing systems. The proposed architecture has following objectives:

- Extensible software based video architecture which allows multiple video codecs to have variable configuration for network bandwidth and picture quality. The architecture allows adding of new codecs with little modification in a pluggable manner.
- Different video codecs users in a video conferencing session can communicate with each other without any handshaking procedure. RMCS architecture proposes a reliable multiple codec change architecture in real time. The codec change is made without handshaking procedure which means that in video conferencing two or more participants do not require any handshaking for availability of codecs.
- To allow real time multiple codec change without handshaking procedure. This aspect will allow video conferencing systems to configure dynamically according to available bandwidth.

1.4 Background

Recent advancement of technology has drastically changed the usage of multimedia and communication technology. Video as a media of communication has affected the communication and entertainment industry. Initially video was transmitted and stored in an analog format for example television broadcasting and radio; magnetic tapes as storage media. The dawn of digital transmission has revolutionaries the media to new horizons. Telephone network was used for interactive communications but with the advent of the Internet, interactive communication has reached to new horizons. The Internet has become a vital infrastructure for modern communications including video conferencing. Video and multimedia conferencing has revolutionized the way people communicate, do business and enjoy entertainment.

1.4.1 Multimedia real time interactive services

Interactive audio and video services usage refers to interactive communication with different users. Internet telephony and internet video conferencing are good example. Previously telephone and teleconferencing were used for interactive communication.

In telecommunication, teleconference is the live exchange of information among persons and machines, apart from one another but linked by a telecommunications system. The telecommunications system may support the teleconference by providing audio, video, and data services by one or more means, such as telephone, telegraph, teletype, radio, and television [13]. Although teleconferencing has been used for decades, the deployment was difficult and expensive.

A videoconference (also known as a video teleconference) is a set of interactive telecommunication technologies which allow two or more locations to interact via two-way video and audio transmissions simultaneously. It has also been called visual collaboration and is a type of groupware [14]. The major advantage for internet based video conferencing is that it uses the existing internet infrastructure with easy deployment and it is inexpensive as compare to traditional teleconferencing.

1.4.2 Digitizing video

Digitizing or digitization is the process of turning an analog signal into a digital representation of that signal. Analog signals are continuously variable, both in the number of possible values of the signal at a given time, as well as in the number of points in the signal in a given period of time. However, digital signals are discrete in both of those respects, and so a digitization can only ever be an approximation of the signal it represents. The digital representation does not necessarily lose information in this transformation since the analog signal usually contains both information and noise [15]. Audio and video signal cannot be sent through internet without digitizing it.

Video is a three dimensional array of color pixels. Two dimensions serves as spatial i.e. vertical and horizontal directions of the moving picture and one dimension represents the time domain. Set of all pixels that corresponds to a single point in time is known as frame. A video consists of a sequence of frames. We get the impression of full motion video when individual frames are displayed fast enough which deceives the human eye that cannot differentiate rapid flashing frames as individual one as motion picture.

Each frame consists of small grids which is the fundamental unit of picture known as pixels. For black and white displays like TV or monitors, each 8-bits pixel represents one of 256 different gray levels. In color displays, each pixel is 24 bits, with 8bit for each primary color i.e. red, green and blue.

Bits per seconds for specific resolution can be calculated for instance if the video at 640x480 resolutions with 25 frames per second is required to transmit over the network following bit rate will be required for transmission:

2 x 25[frames] x 640 [width] x 480 [height] x 24 [color] = 368640000 = 368.64Mbps

As we have seen, to send raw video, requires a very high bandwidth technology for transmission. The solution for optimize transmission at low data rate is to compress the video with some video codec.

1.5 Codec

A codec is a device or program capable of encoding or decoding a digital data stream. The word is the combination of compressor-decompressor or coder-decoder. Codecs are often used in streaming media and video conferencing applications. There are two types of codecs.

1.5.1 Hardware codec

Hardware codec is a special device which is used for compressing or decompressing digital audio or video data. Hardware codecs are embedded into devices that are connected to or installed in a computer such as capture card.

The main advantage of using hardware codec is that it handles the compression and decompression computations at high speed without depending on the computer's CPU. There are also some disadvantages; the cost of developing and manufacturing hardware codecs are generally expensive and often have compatibility issues as some manufacturer support their own format.

1.5.2 Software codec

Software codec is software by which a digital audio or video data is compressed or decompressed using the software algorithm using computer.

Growing computing power of the personal computers' have allowed to implement the software version of the codec with good performance. Software codecs can offer the same quality of hardware codecs. There are numerous advantages of using this approach such as hardware independency, low cost development, deployment and flexibility of updated version of codec.

1.6 Video Compression

Video compression refers to reducing the quantity of data used to represent video content without excessively reducing the quality of the picture. It also reduces the number of bits required to store and/or transmit digital media. Compressed video can be transmitted more economically using a smaller carrier [16].

Digital video transmission requires lots of bandwidth as well as powerful hardware. The better the picture, the more data is required to transmit or store. Video compression algorithms reduce the quantity of data i.e. number of bits required to store or transmit data without excessively reducing the picture quality. Video compression algorithms can be categorized into two groups:

1.6.1 Lossless compression

Lossless data compression is a class of data compression algorithms that allows the exact original data to be reconstructed from the compressed data hence the result is a bit-for-bit perfect match with the original data.

Lossless data compression is mostly used for digital video storage. Examples of these video codecs are Huffyuv, SheerVideo, MSU Lossless Video Codec, H.264/MPEG-4 AVC

1.6.2 Lossy compression

Lossy data compression method is one where compressing data and then decompressing it retrieves data that may well be different from the original, but is "close enough" to be useful in some way. Lossy data compression is most commonly used to compress multimedia data (audio, video, still images) especially in applications, such as streaming media, Internet telephony and video conferencing hence real time compression employs lossy compression. Examples of these video codecs are H.261, H.263

1.7 Digital video paradigm

Digital video systems are replacing all existing analog video systems and making possible the creation of many new telecommunication services such as direct broadcast satellite, digital television, high definition television, video teleconferencing, telemedicine, e-commerce, Internet video that are becoming an essential part of world economy. Objective metrics for measuring the video performance of these systems are required by the industry for specification of system performance requirements, comparison of competing service offerings, network maintenance, and optimization of the use of limited network resources such as transmission bandwidth.

1.8 Contribution

This thesis proposes as architecture which promises optimal utilization of bandwidth. This architecture allows different types of devices to use different video codecs to do video communication seamlessly. The video conferencing users with low processing devices can now use less complex video codec for communication and vice versa. Video traffic is bandwidth intensive and can clog networks. RMCS architecture, as a software architecture is easy to parameterize and modify and thus provides a simple way to respond dynamically to various network architectures. These contributions benefit the field of video conferencing for using the network resources optimally with good picture quality.

1.9 Organization of the thesis

The thesis is organized into six related chapters.

Chapter 1 is an introduction and background on the work to be presented within this thesis.

Chapter 2 gives an overview of video conferencing systems. This naturally leads to a study on existing multiple codec architecture including their strengths and weaknesses. This provides the benchmarks for the proposed algorithm. A summary of the principle problem is given, followed by the proposed solution. This provides an introduction to the main architecture which is presented in chapter 3.

Chapter 3 presents the architecture design and discusses the details of the proposed solution. This includes the design details of the new architecture for changing codecs in real time during a conference. The chapter also describes the communication suite between the video capture and the video playback.

In **Chapter 4**, we discuss the implementation details. The hardware platforms and operating systems used are presented for both video capture and video playback. Programming language and libraries chosen are explained. An insight into the program organization is also given, followed by a description of the tests to be carried out and the analysis of the result obtained.

Chapter 5 describes the tests conducted and analysis of results. It also describes the implementation of the proposed RMCS architecture into multimedia conferencing system v6 (MCS v6), a commercially available video conferencing system.

Chapter 6 presents a summary of the discussions. An attempt is made to outline possible future work that can be done to enhance and continue the research presented here.

CHAPTER TWO
LITERATURE REVIEW

2.1 Introduction

In the past years, a number of compression standards have emerged and currently numerous standards are in development. The new standards are developed to integrate the latest techniques which are optimized for factors like picture quality and bandwidth and can take advantage of enhanced processing power.

This section will discuss the existing multiple video codec architectures for video conferencing systems. It briefly describes each system, displays its architecture (if available), and discusses the capabilities of each system with regards to extensibility of video codecs as well as simultaneous multiple codec support. It also provides a summary of all architectures with reference to this thesis problem statement and proposed solution.

2.2 Multimedia Communications

'Multimedia' means being able to communicate in multiple ways. It is a media that uses multiple forms of information content in the form of image, text, graphics, animation, audio and video to inform or entertain users. It is the combination of two words i.e. multi and media. 'Media' refers to a form of human interaction that is amenable to computer capture and processing, such as video, audio, text, graphics, images, whereas 'multi' signifies that several of those 'media' are present in the same application. Some people define a multimedia system as a computer (or computers) based system which allows end-users to share, communicate and process various information media in an integrated manner [17]. Others say that an application would be considered as 'multimedia' if it involves at least one time-continuous medium, such as video and audio, and at least one discrete one, such as text, image or graphics. Multimedia synchronization denotes a temporal, spatial or even logical relationship between objects, data entities or media streams [18]. Multimedia communication refers to communicate between different users using multimedia contents like text, audio and video etc.

2.3 Video Conferencing

"Teleconferencing systems can generally be classified into three main categories. The first category is audio conferencing where the system provides a shared aural space between participants using the telecommunication network. The second category is the visually augmented audio conferencing system. In this type of system, a shared visual space, in addition to the aural space, is provided. The visual space provides the means for displaying, modifying, and interacting with images. These images can be text, graphics, or photographic still images and can, in some cases, involve some form of animation. The third category is video conferencing. This type of system provides means for communicating live (moving) pictures of conference participants thus expanding the shared visual space. It also subsumes the first and second categories." [19]. Video-conferencing is an efficient means for distributed collaboration especially for people separated by substantial distance. With the increasingly pervasiveness of broadband wide area networks such as the Internet2, it is easy to expect an increasing demand for video-conferencing over these wide area networks in the future.

2.4 Existing video framework for video conferencing

This section will now discuss about the existing systems architectures, and compares them on the basis of video codec extensibility, simultaneous multiple codec support and real time multiple codec change without handshaking procedure.

2.4.1 ITU-T H.323 Video Conferencing Suite

H.323 is a recommended standard protocol from International Telecommunications Union (ITU-T) that provides specification for computers, equipment and services for multimedia communication like voice, video and data conferencing over the packet based network for video conferencing. It is commonly used in Voice over IP (VoIP) and internet protocol (IP) based video conferencing. Users can connect with other users over the internet and use various products which supports H.323. H.323 video conferencing suite utilizes several other ITU-T protocols including H.245, H.460, H.239.

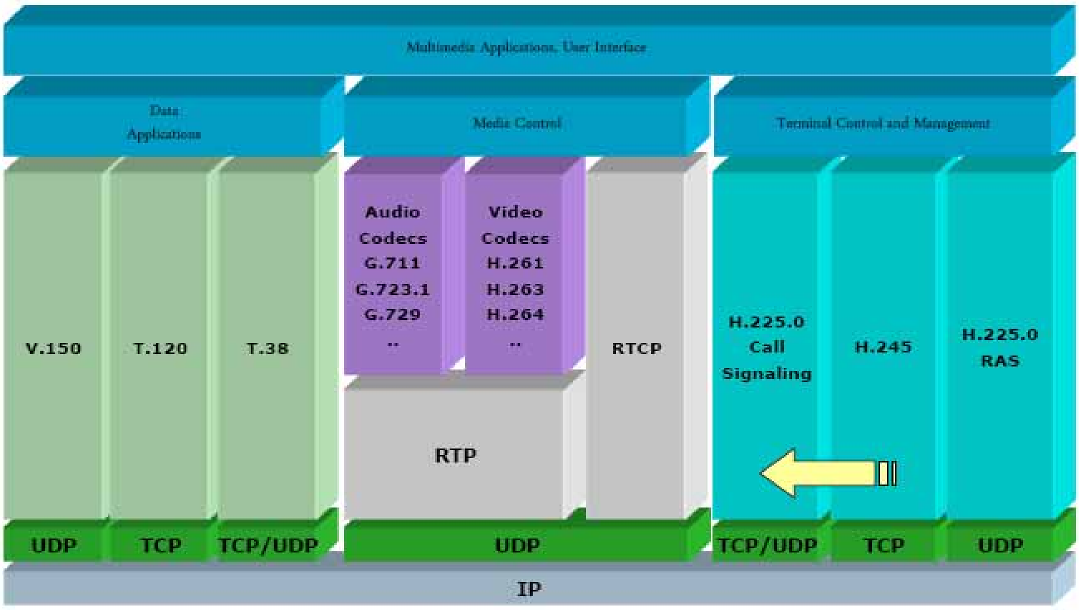

Figure 2.1: H.323 Stack

Architecture

The standard supports wide variety of different protocols under the umbrella like Real-Time Protocol (RTP), Real-Time Control Protocol (RTCP) and User Datagram Protocol (UDP). Currently, H.323 supports three video compression standards which are H.261, H.263 and H.264. There are both hardware based and software based implementation of the standard. Hardware based end products are from companies like Polycom and software based products are Microsoft Net Meeting and Ekiga.

H.245

H.245 is a control signaling protocol in the H.323 multimedia communication architecture. It is responsible for the exchange of end-to-end H.245 messages between H.323 terminals. The protocol specifies control messages including the call setup and capability exchange. The control messages are carried over the H.245 control channels. The messages carried include messages to exchange capabilities of terminals and to open and close the logical channels. After getting connected through call signaling process, the control protocol is used for resolving the call media type as well as establishes the media flow, in pre call establishment.

Capability Exchange Procedure

The capability exchange allows two endpoints to exchange information about what media capabilities they possess, such as G.711, G.723, H.261, and H.263. Along with the type of media, specific details about the maximum number of audio frames or samples per packet is exchanged, information about support for silence suppression (VAD), etc. are also exchanged. Using this capability information, endpoints can select preferred codes that are suitable to both parties. When endpoints advertise capabilities, they also advertise which capabilities may be performed simultaneously. It may not be possible, due to bandwidth limits, to open a high bit-rate video codec at the same time as a high bit-rate audio codec

1 – G.723.1
2 – G.711
3 – H.261
4 – H.264
5 – T.38

Figure 2.2: Capability table

The capabilities of codecs are numbered in a table called capability table. All the attributes like frames per packet etc are included in the capability table. Figure 2.2, illustrates the example of the table.

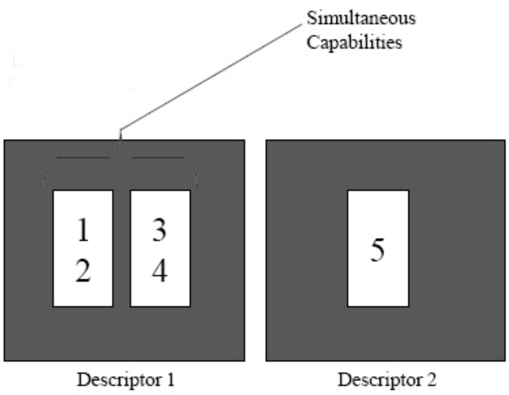

Figure 2.3: Capability Descriptor

In the case of multiple capabilities, H.245 has capability descriptors. As illustrated in Figure 2.3 which shows that there are simultaneous capabilities in the description 1. These capabilities are taken from the capability table in Figure 2.2. At one time, one of the capabilities will be used.

Once capabilities are exchanged, the endpoints negotiate master and slave roles. The master in a point to point conference really only has the power to indicate when channels are in conflict. The slave device must yield to the requests of the master device and reconfigure channels appropriately

Extensibility

H.323 has two ways for extensibility. There is a non-standard data fields found in H.323 and H.245 which can be extended by vendors. The non-standard data fields are octet strings which can be filled with any type of data. Mostly vendors use ASN.1 data structures that are encoded within this non-standard data field.

Second way for extensibility is known as generic extensibility framework or GEF. It was designed in a way that it will not affect the base H.323 documents but still provides the capabilities as if an extension is added to the based document. GEF based extension documents are defined in series of documents number H.460.x H.323 is extended with non-standard features in such a way as to avoid conflicts between vendors.

Simultaneous multiple codec support

The conference is initiated by the master who indicates the media types or video codec to be used in the conference. During the handshaking, if any of the participants does not have the indicated video codec, all the participants drop back to default H.261 video codec.

Real time multiple codec change without handshaking procedure

In H.323, master of the conferencing session, enforces the codec to be used in the conference. There is no procedure for master or participants to change codec during the conference.

2.4.2 Session Initiation Protocol (SIP)

The session initiation protocol also known as SIP provides signaling and control functionality for variety of multimedia services. SIP was published as an Internet Engineering Task Force (IETF) proposed standard [31] in March 1999. It is being used in wide variety of applications like internet telephony, instant messaging and multimedia conferencing. SIP is part of multimedia architecture by IETF which includes the Real-Time Transport Protocol (RTP), Real-Time Streaming Protocol (RTSP), Media Gateway Control Protocol (MGCP) and Megaco (also known as H.248), Session Descriptive Protocol (SDP) and the Session Announcement Protocol (SAP). SIP initiates, modifies and terminates network sessions.

Architecture

"SIP is part of the overall IETF multimedia architecture that has emerged over the past few years. This architecture includes the Real-Time Transport Protocol (RTP) for transporting audio, video and other time-sensitive data, the Real-Time Streaming Protocol (RTSP) for setting up and controlling on-demand media streams, the Media Gateway Control Protocol (MGCP) and Megaco (also known as H.248) for controlling media gateways, the Session Description Protocol (SDP) for describing multimedia sessions, the Session Announcement Protocol(SAP) for announcing multicast sessions, Telephony Routing over IP (TRIP) for locating the best gateway between the Internet and the PSTN, as well as a suite of resource management and multicast address allocation protocols."

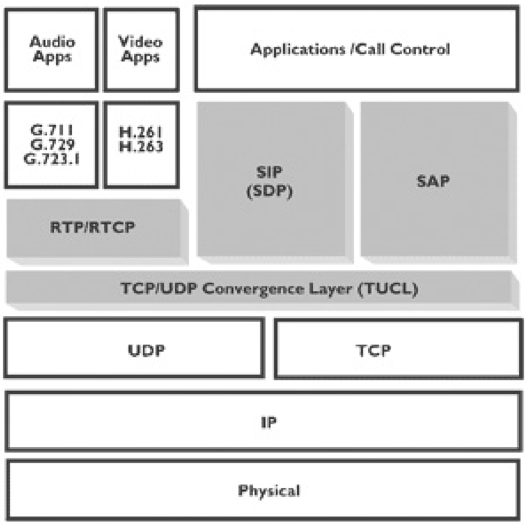

Figure 2.4 SIP Architecture

The SIP INVITE request to join a conference or phone call contains a listing of the media types and associated encodings that the calling party is willing to use. The called party simply responds with a subset of media types and encodings that it is willing to use. Thus, in most cases, negotiation incurs no further delay. Thus, SIP adds an OPTIONS request by which the organizer of a conference call can inquire about a terminal's capabilities without actually initiating a session. [20]

Extensibility

SIP is extended with adding new headers or message bodies that may be used by different vendors to serve different purposes, thus risking interoperability problems. SIP is extended by the standards community to add new features to SIP in such a way as to not impact existing features. However, new revisions of SIP are potentially not backward compatible (e.g., RFC 3261 was not entirely compatible with RFC 2543 [21] [22])

Simultaneous multiple codec support

The one-shot negotiation does work between two participants but fails to work when a multiparty conference is to be set up, as the session description agreed upon between the conference initiator and the first callee may not be applicable to all participants. Hence the codec enforced by the conference initiator is used for all the participants.

Real time multiple codec change without handshaking procedure

SIP also does not support real time multiple codec change without handshaking procedure.

2.4.3 vic - Video Conferencing Tool

vic, is a real-time, multimedia application for video conferencing over the Internet developed by Network Research Group at the Lawrence Berkeley National Laboratory in collaboration with the University of California, Berkeley.

Architecture

“Vic was designed with a flexible and extensible architecture to support heterogeneous environments and configurations. For example, in high bandwidth settings, multi-megabit full-motion JPEG streams can be sourced using hardware assisted compression, while in low bandwidth environments like the Internet, aggressive low bit-rate coding can be carried out in software" [23]. Vic supports three video coding standards that are H.261, H263 and H.263+.

Vic is based on the Draft Internet Standard Real-time Transport Protocol (RTP) developed by the IETF Audio/Video Transport working group. RTP is an application-level protocol implemented entirely within vic. Although vic can be run point-to-point using standard unicast IP addresses, it is primarily intended as a multiparty conferencing application. Vic provides only the video portion of a multimedia conference; audio (vat), whiteboard (wb), and session control tools are implemented as separate applications.

Figure 2.5 VIC

<u>Extensibility</u>

Vic was designed with a scope of flexibility and extensibility in its architecture to support heterogeneous environments and configurations but it still lacks the flexibility of adding of newer video codecs [23].

Simultaneous multiple codec support

Although vic supports three video coding standards but there is no information available about different participants of the conference that can simultaneously communicate with different video codecs [23].

Real time multiple codec change without handshaking procedure

The vic supports three video coding standards but it does not provide any specification which allows it to change codecs in real time [23].

2.4.4 INRIA Videoconferencing System

IVS (INRIA Videoconferencing System) is one of the first software videoconferencing tools for the Internet available in July 1992. IVS is a software system to transmit audio and video data over the Internet using standard workstation.

Architecture

IVS supports software based audio and video codecs. It uses standard Internet technology to transmit video/audio-data. This is achieved by implementing PCM and ADPCM audio codecs and H.261 codec. [24]. IVS can configure the H.261 video codec for three parameters that are refresh rate, the quantizer, and the movement detection threshold.

Extensibility

INRIA Videoconferencing System only supports H.261 video codec and there is no method for adding more video codecs [24].

Simultaneous multiple codec support

Since IVS supports only one video codec hence there is no way of holding conference with multiple video codecs [24].

Real time multiple codec change without handshaking procedure

IVS only supports one video codec [24].

2.4.5 Vaudeville

Vaudeville is an ATM based video conferencing system. It also supports multiple simultaneous multiparty conferences using a scalable multicast mechanism.

Architecture

Vaudeville features NTSC video, voice activated audio transmission, audio bridging of two streams, and voice activated video switching. Audio and video encoding is achieved by using hardware device. “Vaudeville was built using the Project Zeus ATM testbed and the MultiMediaExplorer (MMX). The MMX is an ATM desktop multimedia hardware interface that is host independent, and requires no computing resources from the host workstation for processing of multimedia streams. It is controlled over a RS-232. It allows users to simultaneously transmit and receive audio and video. Video is full-motion, JPEG compressed, NTSC video (60 fields/30 frames per second at 640x480 resolution). The MMX has two audio modes, a default mode and a conferencing mode. The default is CD-quality stereo audio. In conferencing mode, the audio is voice-quality monaural audio.” [25]

Extensibility

Vaudeville uses special hardware for multimedia streams hence there is no prospect of adding new video codecs [25].

Simultaneous multiple codec support

Vaudeville uses only one video codec hence there is no way of holding conference with multiple clients with different video codecs [25].

Real time multiple codec change without handshaking procedure

Vaudeville only supports one video codec [25].

2.4.6 ALX video conference

ALX video conferencing provides high quality audio and video over the WAN. It solves the problem associated with real time audio and video transmission over the wide area IP network [27]

Architecture

It states high quality audio and video over the WAN by utilizing especially designed hardware called adaptation layer translator (ALX). It uses Vaudeville System as a baseline for high quality video conferencing on local area IP network. An end Station in Vaudeville System generates ATM Adaptation Layer 0 (AAL0) video and audio cells. ALX translates and encapsulates these AAL0 cells into IP packets for transmission over wide area IP based networks. On the contrary, multimedia streams of audio and video when received from wide area IP based networks translated back to AAL0 cells so that end systems can understand. [28] [26]. It uses a codec MMX, and ATM enabled motion-JPEG codec.

Extensibility

ALX video conferencing is based on specially designed hardware device for encoding and decoding due to which it is not possible to extend the conferencing application with more codecs unless hardware is updated. Furthermore, there is no information available for supported video coding standards [27].

Simultaneous multiple codec support

This research provides no information for supported video coding standard. Moreover encoding and decoding process also depends on a special hardware. There is no information related with simultaneous multiple codecs with different participants available [27].

Real time multiple codec change without handshaking procedure

ALX research does not provide any details for video codec supported and the procedure for real time multiple codec change [26].

2.4.7 Seodang

Seodang is a desktop video conferencing application uses Windows NT platform. It was developed for collaborative learning. Seodang uses low cost components hence appropriate for individual or small group usage.