ABSTRACT

Since 1960, many companies have invested more audio visual technology and the network to support it. This was due to recognition of importance of face to face meeting with key customers, suppliers and employees. Reduced travel cost and long distance calls as shortened decision making cycle and increased productivity of those systems justify their competition nowadays communication networks.

Rwanda is not behind this opportunities renovation where we find mostly absence or lack of based of communication infrastructures. The main problems of administration, management and control that face the Rwandese Government body and independent organisations are likely to be solved by implementation of good communication infrastructures based on New Generation Network (commonly said 3 Generation Network). This initiate requires less infrastructures and can be upgraded on current communication networks easily.

For that matters, KIST ICT has proposed to deal with this issues by organising student project related to this new technology to help the Government and other independent bodies to improve their Quality of work and easy administration. It is in that context that a project of Design and Implement of an audio-visual data network has been conducted with objective of getting convergence of data, video and voice on a single IP network.

The main goal of learning and participate in this research project constitute the main reason of formulation and composition of this report.

Executive Summary

The objectives of the project include understanding the potential of audio, video and audio-visual communication at bit rates up to 10 Mbit/s over IP networks as the quality of service provided by these networks . This is based mainly on evaluation of current network to allow audio visual quality. It is hoped to encourage the use of audio-visual communication on these networks. The project aims to evaluate methods for subjective and objective audio-visual quality assessment. This assessment specifies the compression as well as coding techniques to apply for multimedia transmission over IP network. Audio codec refers to PCM coding schemes or commonly called G.711 and G.723 specified on hardware card used for voice communication. Video coders supported in this work are Px64 coder such as H.261 and H.263 for conferencing and MPEG2 for video on demand and TV broadcast. RTP transmission is considered for audio and video stream transport over IP network. The Quality of service is maintained by RSVP and RTCP. The H.323 used for VoIP and Videoconference specifies RTP, RSVP and call signalling protocols to forward audio visual objects through the network with good quality. The development of this project aim to implement computer telephony communication network as well as visual face to face meeting and video entertainment

TABLE OF CONTENTS

ABSTRACT 1
Executive Summary 2
TABLE OF CONTENTS 3
List of Abbreviations 4
I. INTRODUCTION 5
II. AUDIOVISUAL NETWORKING BACKGROUND AND BASIC CONCEPTS 6
2.1 ANALOG VERSUS DIGITAL CONCEPTS 6
2.1.1 Analog signal 6
How are Voice and Video in analog communication? 6
2.1.2 DIGITAL SIGNAL 7
2.2.3 ANALOG ADVANTAGES 8
1.1.3 CODING 10
2.1.3.1 Coding voice and sound 10
2.1.3.2 Coding Video 10
2.1.4 COMPRESSION CONSIDERATION 12
2.1.5 STREAMING VIDEO 13
2.1.6 VOICE AND VIDEO CODING RECOMMENDATION TO BE CONSIDERED FOR OUR NETWORK 16
2.1.6.1 Audio Coding Recommendations 16
2.1.6.2 Real Time communication 17
2.1.6.3 Video coding recommendation 18
2.1.7 AUDIO-VISUAL APPLICATION OVER IP NETWORK 19
2.1.7.1 Computer Telephony 19
2.1.7.2 Videoconference interactive meeting 21
2.1.7.3 IP TV 22
III. NETWORK DESIGN AND REQUIREMENT PLAN 23
3.1.1 NETWORK REQUIREMENT 25
3.3 IP VIDEOCONFERENCE DESIGN 26
IV. NETWORK IMPLEMENTATION 30
4.2 VOIP NETWORK IMPLEMENTATION 32
4.2.1 Client PC setup 32
4.2.2Gateway setup 32
4.2.3 Call configuration 32
VoIP gateway configuration 35
V. SUMMARY AND CONCLUSION 37

List of Abbreviations

AVU	African Virtual University
CATV	Community Antenna TV
CD	Compact Disk
CIF	Common Intermediate Format
CODEC	Coding-Decoding
CPE	Customer Point Equipment
Db	Decibel
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name server
ICT	Information Communication Technology
IP	Internet Protocol
ISDN	Integrated Service Digital Network
ITU	International Communication Union
JPEG	Join Picture Experts Group
KIST	KIGALI INSTITUTE OF SCIENCE AND TECHNOLOGY
LAN	Local Area Network
MCU	Multiunit Control Unit
MPEG	Moving Picture Experts Group
NAT	Network Address Translation
NTSC	National Television Standards Committee
PAL	Phase Alternate Line
PSTN	Public Switch Telephone Network
QCIF	Quarter Common Intermediate Format
QoS	Quality of Service
RGB	Red Green Blue
RTCP	Real Time Control Protocol
RTP	Real Time Protocol
SECAM	Système Electronique Couleur Aven Memoire
STB	Set To Box
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
VoIP	Voice over IP
WINS	Windows Name Server

I. INTRODUCTION

The technology behind audio-visual computer network is based on a variety of ways to design and implement it. The main idea is to incorporate to current LAN/WAN data network with voice and video services. This issue can be operational using switching based technique, according to the performance and benefits of the technique. Such techniques are either circuits switched network or packet switched network.

Nowadays, with high advances in computer technology with also limitation of circuits switched network in terms of high demand and high prices of audio-visual services, real time based computer network are more considered in design and implementing a two ways multimedia network.

In this project, the H.323 technology is considered as basis of our network, because of its standardisation, good management and easy to incorporate to the current LAN/WAN of on a Campus.

Objective

The main objective of this project is to design, implement an audio-visual computer network on current LAN data network offering basically Computer based telephony, audio-visual conference meeting as well digital audio and video entertainment.

To achieve this objective, the project is subdivided in different stages assembled in the followings sections:

- Section 1 gives a brief overview of audio-visual technology concepts and background.
- A sub section describes the main applications of audio-visual objects over IP network
- Section 2 describes the network design and requirement plan.
- Section 3 focuses on Voice over IP implementation on corporate LAN and VoIP gateway configuration to connect to other communication networks.
- Section 4 gives brief summary and conclusion.

The project shows how it is reliable to deploy voice and video communication services using the LAN data network previously working only on data communication with guaranteed Quality of service. The document can be also used by someone else who wants to develop other audio and visual services .

II. AUDIOVISUAL NETWORKING BACKGROUND AND BASIC CONCEPTS

The general audio-visual communication network can be design or implemented in two ways:

analog communication network or

digital communication network

2.1 Analog versus Digital concepts

As said above, communications can be classified in two categories, analog and digital. In the analog form of electronic communications, information is represented as a continuous electromagnetic wave form. Digital communications represents information in binary form (1s and 0s) through a series of discrete blips or pulses.

2.1.1 Analog signal

Analog is best explained by examining the transmission of a natural form of information, such as sound or human speech, over an electrified copper wire. In its native form, human speech is an oscillatory disturbance in the air which varies in terms of its volume or power (amplitude) and its pitch or tone (frequency). As sound compression waves fall onto a transmitter, analogous (approximate) variations in electrical waveforms are created over an electrical circuit. Those waveforms maintain their various shapes across the wire until they fall on the receiver or speaker, which converts them back into their original form of variations

in air pressure.

A similar, but more complicated, conversion process is employed to transmit video over networks. In its native form, video is a series of still images captured and transmitted in rapid succession in order to create the illusion of fluidity of motion; image information is reflected light waves. Analogous variations in electrical or radio waves are created in order to transmit the analog video image information signal over a network from a transmitter (TV station or CATV- Community Antenna TV source) to a receiver (TV set), where an approximation (analog) of the original information is presented.

Information which is analog in its native form (voice and image) can vary continuously in terms of intensity (volume or brightness) and frequency (tone or color). Those variations in the native information stream are translated in an analog electrical network into variations in the amplitude and frequency of the carrier signal. In other words, the carrier signal is modulated (varied) in order to create an analog of the original information stream.

How are Voice and Video in analog communication?

Voice

A voice grade channel is approximately 4,000 Hz, or 4 kHz. Approximately 3.3 kHz (200 Hz to 3,500 Hz) is used for the voice signal itself. The remaining bandwidth is used for purposes of network signalling and control, and in order to maintain separation between information channels. While human speech transmission and reception encompasses a much wider range of frequencies, 3.3 kHz is considered to be quite satisfactory and is cost-effective. Band-limiting filters are used in carrier networks to constrain the amount of bandwidth provided for a voice application

Video

A CATV video channel is approximately 6,000,000 Hz, or 6 MHz. Approximately 4.5 MHz is used for information transmission, while the balance is used for guard bands to separate the various adjacent channels riding the common, analog coaxial cable system.

2.1.2 Digital signal

While the natural world is analog in nature, computers (which are decidedly unnatural beings) are digital in nature. Computers process, store, and communicate information in binary form. That is to say that a unique combination of 1s and 0s has a specific meaning in a computer language. A bit (binary digit) is an individual 1 or 0. Multiple bits travel across a network in a

digital bit stream.

Digital communications dates to telegraphy, in that the varying length of making and breaking an electrical circuit resulted in a series of dots and dashes which, in a particular combination, communicated a character or series of characters. Early mechanical computers used a similar concept for input and output; contemporary computer systems communicate in binary

mode through variations in electrical voltage.

Digital signalling, in an electrical network, involves a signal which varies in voltage to represent one of two discrete and well-defined states: such as either a positive (+) voltage and a null, or zero (0), voltage (unipolar); or a positive (+) or a negative (-) voltage (bipolar). The receiver monitors the signal, at a specific frequency and for a specific duration (bit time) to determine the state of the signal. Various data transmission protocols employ different physical states of the signal, such as voltage level or voltage transition. Because of the discrete nature of each bit transmitted, the bit form is often referred to as a square wave

In the digital world, bandwidth is measured in bits per second (bps). The amount of bandwidth required depends on the amount of raw data to be sent, the desired speed of transmission of that set of data, and issues of transmission cost.

Additionally, data is routinely compressed by various means in order to enhance the efficiency of transmission and to reduce transmission costs. Additionally, analog voice commonly is converted to a digital bit stream, requiring a maximum of 64 Kbps for full fidelity or quality.

Note that the voice bandwidth in analog form is 4 kHz. By Nyquist's principal , the frequency bandwidth must be twice the maximum required to produce good quality voice. Thus results in producing 8 kHz for voice bandwidth.

By comparing both techniques, we state advantages and limitations of each techniques as follows:

2.2.3 Analog advantages

Analog Data.

Analog transmission offers advantages in the transmission of analog information. The process of transmission of such information is relatively straightforward in an analog format, whereas conversion to a digital bit stream requires conversion equipment.

Bandwidth

Additionally, it is more bandwidth-conservative and is widely available. A raw information stream consumes less bandwidth in analog form than in digital form. This is particularly evident in CATV transmission, where 50 or more analog channels routinely are provided over a single coaxial cable system. Without the application of compression techniques on the same cable system, only a few digital channels could be supported.

Presence

Finally, analog transmission systems are already in place, world-wide

Digital Advantages

Digital Data

Digital transmission offers advantage in the transmission of digital information. It is better to transmit digital information in a digital format. Digital transmission certainly has the advantage where binary computer data is being transmitted.

Compression

Digital data can be compressed relatively easily, thereby increasing the efficiency of transmission. As a result, substantial volumes of voice, data, video and image information can be transmitted using relatively little raw bandwidth.

Once we have digital data we may convert it to standard format that could be quickly transmitted. For that, may techniques are available to compress audio such as:

Broadcast audio distribution with sample rate of 32 kHz (to provide a 15 kHz bandwidth) and with 14 bit-linear coding.

Professional recording with sample rate of 48 kHz with 16 bits or 20 bits coding

Compact disk recording with a sample rate of 44.1 kHz with 14-16 bits coding .

Security

Digital systems offer better security. While analog systems offer some measure of security through the scrambling, or intertwining of several frequencies, scrambling is fairly simple to defeat. Digital information, on the other hand, can be encrypted to create the appearance of a single, pseudo-random bit stream. Thereby, the true meaning of individual bits, sets of bits, or the total bit stream cannot be determined without having the key to unlock the encryption algorithm employed.

Quality

Digital transmission offers improved error performance (quality) as compared to analog. This is due to the devices that boost the signal at periodic intervals in the transmission system in order to overcome the effects of attenuation. Additionally, digital networks deal more effectively with noise, which always is present in transmission networks.

Attenuation

Electromagnetic signals tend to weaken, or attenuate, over a distance; this is particularly true of electrical signals carried over twisted pair copper wire, due to the level of resistance in the wire. It is also particularly true of microwave radio and other terrestrial radio systems, due to matter in the air. Attenuation is sensitive to carrier frequency, with higher frequency signals attenuating more than lower frequency signals.

Noise

Signals also tend to pick up noise as they transverse the network. Again, this is particularly true of twisted pair, copper wire systems. Such wires tend to act as antennae and, therefore, absorb noise from outside sources of ElectroMagnetic Interference (EMI). Thus, the quality of the signal degenerates as it is distorted by the noise.

NB

The SNR(signal to noise ration)=dB $= 10 \log \frac{\text{Power out}}{\text{Power in}}$

dBm $= 10 \log \frac{\text{(Power, measured in mW)}}{1 \text{ mW}}$

The reference power used in describing noise is –90dBm

To describe the noise in circuits, we use dBrn= dBm + dB

To get 0 dBrn= -90dBm+x dB, we need x= 90. The SNR is 90 dB for 0 dBrn

To get good quality sound, most of Sound cards use 90 dB SNR.

Cost

The cost of the computer components required in digital conversion and transmission has dropped considerably, while the reliability of those components has increased over the years.

Upgradability

Since digital networks are comprised of computer (digital) components, they are relatively easy to upgrade. Such upgrades can increase bandwidth, improve error performance, and enhance functionality. Some upgrades can be effected remotely over a network, eliminating the need to dispatch expensive technicians for that purpose.

After comparing them, digital data network offer more possibility to perform also video and voice services on the Campus IP network. Generally speaking digital networks can be managed much more easily and effectively due to the fact that such networks consist of computerised components. Such components can sense their own level of performance, isolate and diagnose failures,initiate alarms, respond to queries, and respond to commands to correct any failure. Further, the cost of so enabling these components continues to drop.

But the conversion of digital to analog (and reverse conversion) is a must due to the predominance of analog presence in multimedia communication. The project mainly focuses on digital to analog conversion needed to add video and voice services on a corporate IP LAN network. This process is called CODING.

1.1.3 Coding

Coding is simply analog to digital signal conversion. The material used in analog to digital conversion is commonly called CODEC (coding-decoding). The reverse process is called modem(digital to analog).

Encoding is the process of converting an analog information stream (e.g., voice or video) into a digital data stream. The voice or video signal is sampled at frequent intervals with each sample of amplitude then being expressed in terms of a binary(computer) value, which is usually a 4-bit or 8-bit byte. The reverse process of decoding takes place on the receiving end,

resulting in recomposition of the information in its original form, or at least a reasonable approximation thereof.

2.1.3.1 Coding voice and sound

In telephony, the sampling rate is 8000 samples per second (8 khz). Each sample is quantized in the of range 0 to 4095 and compressed in value between 0 and 255 (having 256 values), filling into one byte= 8 bit. The throughput will be 8000 samples* 8 bits= 64 000 bits= 64 Kbit. This technique is called Pulse Code Modulation as G.711 standards.

For music, today sound cards allows to convert with 16 bit of 22050 Hz (sampling rate of CD Player). By Nyquist, we need 44100 samples per second. The required memory buffer or network bandwidth will be 2 bytes(16 bit)*44100= 88200 byte/s = 705600 kbit/s. For stereo, we use 4 byte*88200= 176 Kbytes/s. Those configurations are the basis of the VoIP implementation in this project with only the difference of using G.723.1 coding on the level of sound cards.

2.1.3.2 Coding Video

In video coding , more basics have to be considered. The basic concepts are those of frame rate, scanning, resolution, aspect ratio, luminance, chrominance, and synchronization.

Frame rate

refers to the rate at which frames of still images are transmitted. Video is a series of still images that are transmitted in succession in order to create the perception of fluidity of motion. If transmitted in rapid succession, the perception is one of complete fluidity—24 frames per second (fps) is considered to be motion picture quality, while 30 fps is considered broadcast quality for this project. If the frames are transmitted at a slow rate, the result is a herky-jerky video of poor quality.Particularly below 15 fps, quality suffers quite noticeably.

Scanning

refers to the process of refreshing the screen in horizontal lines. Odd lines are refreshed in one scan, and even lines in the next. Each set of odd and even lines refreshed constitutes a frame refreshed, with the scanning rate being a function of the power source of the receiver. For example, the American

NTSC standard provides for 30 fps, involving 60 scans, which relates directly to the 60 Hz of the U.S. power source.

Resolution

refers to the definition, or sharpness, of the image. Resolution is determined by the number and density of the pixels (picture elements), which essentially are dots of picture. The greater the number and density of the pixels, the better the resolution. If the same number of pixels are spread over a greater area, the result is a grainier picture, as one can readily see by sitting close to a big-screen TV.

Aspect ratio

refers to the relationship between the width and the height of the image. The 4:3 (4 wide to 3 high) aspect ratio specified by the American NTSC standard is rooted in the early days of television, when round picture tubes made effective use of this approach.

Ex: 640x480 is the reference for our desktop videoconference consideration.

Luminance

refers to intensity, or brightness, which can vary within an image. A video transmission varies the luminance by varying the power level, or amplitude, of the signal.

Chrominance

refers to color, with different standards allowing for varying levels of color depth. Clearly, the video image is more pleasing and life-like when the variation of color is as broad as possible.

Synchronisation

includes vertical and horizontal synchronization, both of which are critical. Vertical synch is required to keep the picture from scrolling or flipping. Horizontal synch keeps the picture from being twisted.

Major analog standards are NTSC and PAL, which are compared to digital HDTV in this table

Television standards compared: NTSC, PAL, and HDTV.

	NTSC	**PAL**	**HDTV**
Analog/Digital	Analog	Analog	Digital
Scanlines	525	625	1,920 or 1,280
Synchronization	40	49	N/A
Resolution (pixels per line)	640	640	1,080 or 720
Frame Rate (fps)	30	25	60, 40, or 24
Aspect Ratio	4:3	4:3	16:9

NTSC ((National Television Standards Committee) , established in the United States as the first standard in 1953 ,is characterized as analog in nature with 525 scanlines. There are 640 pixels per line, 485 of which are dedicated to the active picture. The frame rate is 30 fps and the aspect ratio is 4:3 .

PAL (Phase Alternate Line), established in western Germany, The Netherlands, and the United Kingdom, PAL is characterized as analog, with 625 scanlines. There are 640 pixels per line, with 576 dedicated to the active picture. The frame rate is 25 fps, and the aspect ratio is 4:3.

The third standards is SECAM (Systeme Electronique Couleur Aven Memoire) developed in France.

By coding the video, the Digital TV network can be also implemented on computer networks. Digital video is nothing more than data, offering advantages in terms of processing, storage, and manipulation. These advantages include editing, alteration (morphing), reproduction, compression, transmission, storage, and store and forward capability. Digital video requires that the information be digitized through the use of a video codec. Because a broadcast quality video signal is extremely bandwidth-and storage-intensive, the resolution of such digital video requires 640 horizontal pixels and 480 vertical; chrominance and luminance require 24 bits/pixel; and the frame rate is 30 fps. The video signal alone requires 7,372,800 bits per frame. At a frame rate of 30 fps, the bandwidth requirement is 221,184,000 bps. Clearly, it is not within the realm of reason to transmit a broadcast quality video signal requiring 221+ Mbps. It is clear that the storage requirements are incredible and compression is critical if the networks are to support digital TV.

Despite all the advantages of digital technology as applied to television production, storage and transmission, all of the TV sets are analog. A conversion from analog to digital broadcast TV is revolutionary and completely unthinkable, unless there exists a compelling reason for the changeout of untold millions of TV sets. That compelling reason could (but probably won't) take the form of HDTV (High Definition TV). While the standards have yet to be fully determined, HDTV will be digital, with an aspect ratio of 16:9. The number of scanlines will be either 1,920 or 1,280; resolution will be either 1,080 or 720 pixels per line; and the frame rate will be either 60, 40 or 24 fps.

2.1.4 Compression consideration

The bandwidth required for video transmission is significant, affected by frame rate, resolution, color depth, aspect ratio, and audio. Broadcast-quality TV requires about 6 MHz in analog form, of which the signal occupies about 4.5 MHz. As little can be done to compress an analog signal, analog TV is limited to airwave broadcast or CATV transmission over analog coaxial cable—in other words, analog is doomed!

Digital video, on the other hand, can be compressed fairly easily. Because uncompressed, broadcast quality, digital video requires between 90 Mbps and 270 Mbps, compression is critical. Without compression, a 1 Gbps fiber optic network could accommodate no more than 11 digitized NTSC channels. To digitize and compress the video information stream, the analog video and data signal first must be digitized through the use of a codec. Clearly, the amount of bandwidth required to transmit digital video, and the amount of memory required to store it, can be reduced by reducing the frame rate, resolution, or color depth. However, the result is less than pleasing. In order to maintain the quality of the video presentation, therefore, the data must be compressed using an appropriate and powerful data compression algorithm.

Lossless compression allows the video signal to be reproduced faithfully with no data loss and compression rates are in the range of 10:1. However, even the relatively new ITU-R 723 standard reduces the data rate to only 45 Mbps. *Lossy* compression results in a degraded picture, but allows compression at rates up to 200:1 [14-8]. Actually, compressed video currently can be transmitted with quite acceptable quality at T1 speeds of 1.544 Mbps or less. MPEG uses lossy compression in the form of DCT. There are a number of steps involved in video compression, including filtering, color-space conversion, scaling, transforms, quantization and compaction, and interframe compression.

Filtering reduces the total frequency of the analog signal through a process of averaging the values of neighboring pixels or lines. For instance, a black pixel and a white pixel become a gray pixel. *Taps* are the number of lines or pixels considered in this process; MPEG, for instance, uses a 7-tap filter.

Color-space conversion involves the reduction of color information in the image. As the human eye is not highly sensitive to slight color variations, the impact is not noticeable. Black and white, however, is prioritized, as the human eye is very sensitive at that level.

Scaling addresses the creation of the digital image according to the presentation resolution scale. Rather than digitizing the video signal in large scale, the codec is tuned to the scale of presentation in terms of horizontal and vertical pixels, thereby reducing the amount of data that must be digitized. In consideration of this factor, the aspect ratio must be standardized.

Transforms convert the native two-dimensional video signal into data dimensions. Although beyond the scope of this project, the various approaches include Discrete Cosine Transform (DCT), vector quantization, fractal transform, and wavelet compression.

Quantization and compaction encoding reduce the number of bits required to represent a color pixel. Although again, beyond the scope of this document, compaction techniques include run-length encoding, Huffman coding, and arithmetic coding.

Interframe Compression considers and eliminates redundant information in successive video frames. For instance, the background of a movie scene might not change, although the actors move around the set. While the motion of the actors must be reflected, the background need not be transmitted over the network—it can be compressed out of each frame until such time as background changes must be reflected.

2.1.5 Streaming Video

Streaming video technology is being more widely used on campus. Like most digital video systems, it involves the generation of content, making that into a digital format that can be transmitted over a network, routing the digital data to the intended receivers of the video, and arranging for client software and hardware that will allow the intended users to see and hear the content.

The distinguishing characteristic of streaming video is that the user does not have to download the entire file of video before starting the display. As soon as enough information has been received to make the first image, display starts. Video information is discarded by the receiving system as soon as it is not needed. Since video files can be very large, this makes use easier and more timely. It also affects the requirements that systems should meet in order to operate best.

Streaming video systems from several different vendors are being implemented and used. Each has some proprietary characteristics, but most are attempting to interoperate with the others. For instance, many software clients are able to display output from other vendors' servers. In the next several months, the industry is expected to converge to at most 3 or 4 different output formats with clients that are able to display all of them.

In the process of compiling this list, the campus architecture team looked at the operating characteristics of several streaming digital systems including:

IPTV by Cisco
Netshow by Microsoft
Quicktime by Apple
RealVideo
VideoCharger by IBM

This statement deliberately does not make any comparison of the various technologies, because different technologies may be appropriate for different applications. The choice would be based on content characteristics, number of users, method of delivery, and costs. Nor does this statement address issues of content creation because they are not relevant to issues of delivering streaming video.

Examples of content creation using some of these technologies can be obtained from various sources, including the URL's in the Additional Information section. These can provide more up-to-date and specific information about characteristics and speeds than any "snapshot" that this statement might provide.

Statement of Directions:

Http-streaming

Many users are familiar with a video technology often found on web pages, which this statement will refer to as http streaming. When a link to certain video clips is clicked, the entire video file is downloaded to the user's desktop system, in contrast to true streaming video technologies which never writes or leaves the entire video file (aka the video system asset) on the user system. However, the user's browser starts displaying from an http video clip as soon as enough is downloaded to generate a display, making the operation look like streaming video.

True streaming video is preferable because it provides better management control of the streaming process. It also provides better preservation of the assets. Much video is covered by copyrights, which are easier to enforce if complete copies of the assets are not available to users.

Considerations for true streaming video

In order to operate effectively on a network that carries the traffic of the entire campus, the factors listed below should be considering in planning, designing, implementing, and operating a streaming digital video system. They are divided into several areas of concern to a streaming video system.

***Bandwidth considerations*:**

Since video systems generally require large amounts of bandwidth, extensive use by hundreds if not thousands of users will not only slow the performance of individual users, but will also degrade network performance for the entire campus. Therefore systems designers should minimize the use of bandwidth consistent with reception quality that is adequate for their users' needs.

Designers should consult networking support for advice early in the design of video systems. Planners should be aware that network pricing models may have to change to recover costs of high-bandwidth traffic such as video.

***Client communications*:**

To be effective as visual displays, the user must have communications fast enough to support video without being distracted by jerky or incomplete movement.

Dial-up connections should be at least 28.8 kbps, although in some applications users may not consider the quality acceptable. Faster connections are preferable for improved quality.

Shared LAN connections, e.g. normal Ethernet, may limit delivery of multiple video streams. In such a LAN a fixed total capacity is shared amongst all users, many of whom may be attempting to simultaneously receive a video stream. Switched networking, either layer 3 or switched Ethernet, is preferable to any type of shared LAN.

Desktop software:

Systems should require desktop software that is easily and widely available. There are good user clients that can be downloaded free from web sites. Hence only applications with very unusual requirements should require the use of anything else.

Desktop hardware:

The system specifications below are recommended for minimal video performance. When buying a new system, they should be considered minimums. Older systems that are short of these recommendations may still give adequate performance for some applications and should not necessarily be replaced.

Windows systems:
Pentium-class CPU, 133 MHz or greater
32-bit operating system (not Windows 3.x)
32 Mbytes of RAM
4 Mbytes of VRAM
16-bit sound card

Network communications:

These considerations describe the network resources that should be available to the server portion of the system. A particular system may have special requirements or make an unusual impact on the campus network.

The output network connection should be sized to accommodate the anticipated number of concurrent users.

Switched Ethernet is preferable to shared Ethernet.

Layer 3 switching is preferable to either type of Ethernet.

Multicast operations should be useable because they use less bandwidth than unicast.

Server Parameters:

These considerations describe the capabilities of the server system:

Able to generate multiple output streams from a single stored file

Support RTSP and RTP protocols.

Deliver H.263 formatted output to support dial-up clients.

Audio requirements:

Even though this statement addresses video systems, designers and implementers should recognize that audio is an essential part of most applications. They have to decide what minimum sampling rate to use for the digital audio in order to get good-quality reproduction.

Client software is often limited as to the quality it can reproduce. Furthermore, high audio sampling rates will add to already-high bandwidth requirements.

Therefore systems should be careful not to use higher sampling rates than necessary for adequate reproduction.

2.1.6 Voice and video coding recommendation to be considered for our network

2.1.6.1 Audio Coding Recommendations

The ITU audio coding recommendations are referred to as the "G-series." All H.3xx recommendations (except H.324) mandate the G.711 recommendation for audio. Because of bandwidth limitations, the only audio recommendation mandated by H.324 is G723.1.

G.711 is the oldest compression algorithm. It is mandated by all H.3xx recommendations (except H.324). G.711 codes toll-quality (3KHz analog bandwidth) audio into 48, 56, or 64 Kbps.

Mandatory for H.324. Optional for all other H.3xx recommendations. G723.1 codes toll-quality (3KHz analog bandwidth) audio into 5.3 or 6.3 Kbps.

Others are optional recommendation such as :

G.722 codes enhanced quality (7 KHz analog bandwidth) audio into 48, 56, or 64 Kbps.

G722.1 codes enhanced quality (7 kHz analog bandwidth) audio into 24 or 32 Kbps. G.722.1 was approved in September 1999.

G728 codes toll-quality (3KHz analog bandwidth) audio into 16 Kbps.

G729 codes toll-quality (3KHz analog bandwidth) audio into 8 Kbps.

The protocols specified by H.323 are the following:

Audio Codec

An audio codec encodes the audio signal from microphone for transmission on transmitting H.323 terminal and decodes the received audio codev that is sent to the speaker on the receiving H.323 terminal. Because audio is the minimum service provided by the h.323 standard, all H.323 terminals must have at least one audio Codev support, as specified in ITU-t G.711 recommendation (audio coding at 64 Kbps). Addition audio codec recommendation such as G.722 (64, 56 and 48 kbps), G.723 (5.3 and 6.3 kbps), and G.729 (8 Kbps) may be also be supported.

Video Codec

A video codec encodes video from camera for transmission on the transmitting H.323 terminal and decodes the received video code that is sent to the video display on the receiving H.323 terminal. Because H.323 specifies support of video as optional, the support of video codecs is optional as well. However any H.323 terminal providing video communication must support video encoding and decoding as specified in the ITU-T H.261 recommendation.

H.261 also refers to the video coding standard and the video format. Specifically, H.261 includes the optional CIF (Common Intermediate Format) at 352 x 288 = 101,376 (i.e. width by height) pixels per frame, at a maximum rate of 30 fps. H.261 also supports the mandatory QCIF (Quarter CIF) at 176 x 144 = 25,344 pixels per frame, which is exactly one-quarter the resolution of full CIF.

H.225 Registration, Admission and Status

Registration, admission and status (RAS) is a protocol between endpoints (terminals and gateways) and gatekeepers. The RAS is used to perform registration, admission control, bandwidth changes, status and disengage procedures between endpoints and gatekeepers. An RAS channel is used to exchange RAS messages. This signalling channel is opened between an endpoint and gatekeeper prior to the establishment of any other channels.

H.225 signalling

The H.225 call signalling is used to establish a connection between two H.323 endpoints. This is achieved by exchanging H.225 protocol messages on the call-signalling channel is opened between two H.323 endpoints or between an endpoint and the gatekeeper.

H.245 Call signalling

H.245 control signalling is used to exchange end to end control messages governing the operation of the H.323 endpoint. These control messages carry information related to the following :

- Capabilities exchange
- Opening and closing of logical channels used to carry media streams
- Flow control messages
- General commands and indications

2.1.6.2 Real Time communication

The real time transport over IP network is based on the following protocols:

Real Time Transport Protocol

Real time transport protocol (RTP) provides end to end delivery services of real-time audio and video. Whereas H.323 is used to transport data over IP-based network, RTP is typically used to transport data via the user datagram protocol (UDP). RTP together with UDP, provides transport-protocol-type functionality. Note UDP is connectionless whether TCP is connection oriented (this is heavy for RTP). Because of its connectionless, UDP cannot ordering packets in arrival time. With datagram concept, each packet is independent from others. That is why RTP is introduced to manage packets. RTP provides payload-type identification, sequence numbering, timestamping and delivery monitoring. UDP provides multiplexing and checksum services. RTP can be used with other transport protocols.

Real Time Control Protocol

Real time Control protocol (RTCP) is the counterpart of RTP that provides control services. The primary function of RTCP is to provide feedback on the quality of data distribution

Resources reSerVation Protocol

Resource reservation Protocol (RSVP) is a protocol that has the main job of reserving all necessary resources required for network transmission. The RSVP work along the network between routers to accommodate enough bandwidth for any real time transmission; It is not a routing protocol, but a protocol used to provide all resources for good QoS when transmission of real time data.

H.323 also defines RTP, RTCP and RSVP.

2.1.6.3 Video coding recommendation

Px64

Px64 is an ITU-T standard designed to support videoconferencing and various levels of bandwidth in increments of 64 Kbps up to a maximum of 2.048 Mbps (E1). (*Note: p* = 1-30 channels of 64 Kbps, with the maximum of 30 bearer channels being supported by E1.) Px64 specifies various frame rates and levels of resolution. Most manufacturers of videoconferencing equipment support Px64 as a lowest common denominator, although they each prefer their own proprietary standards in promotion of their own equipment and unique feature sets. Px64 also is known as H.261, referring to the specific ITU-T video coding standard, as well as the video formats. Those formats include *CIF* (*Common Intermediate Format*), which is optional, and *QCIF* (*Quarter-CIF*), which is mandatory in compliant codecs. H.261 supports 352 pixels per frame, 288 lines per frame and 30 frames per second, although lower frame rates also are supported.

Actually, H.261 is an element of the ITU-T H.320 umbrella standard for videoconferencing that addresses narrowband visual telecommunications systems and terminal equipment. Related ITU-T standards include H.221, which defines a frame structure for support of audiovisual teleservices in 64-Kbps channels; and H.222, which defines the frame structure for such services in an ISDN environment.

JPEG

Motion-JPEG is used in the editing of digital video. JPEG is not appropriate for video transmission, as the compression rate is in the range of only 20:1-30:1. JPEG transmission in support of videoconferencing requires bandwidth in the range of 10 to 240 Mbps.

MPEG

MPEG compression is as high as 200:1 for low-motion video of VHS-quality; broadcast quality can be achieved at 6 Mbps. Audio is supported at rates from 32 Kbps to 384 Kbps for up to 2 stereo channels.

MPEG 1 provides VHS (videotape) quality at 1.544 Mbps and is compatible with single-speed CD-ROM technology. In fact, it was designed to put movies on compact disc. MPEG 1 integrates synchronous and isochronous audio with video, and allows the random access required by interactive multimedia applications. Intended for limited-bandwidth transmission, it provides acceptable quality and output compatible with standard televisions. Current applications include video kiosks, video-on-demand, and training and education. Compression of about 100:1 is supported by MPEG-1. MPEG1 was originally optimized to work at video resolutions of 352x240 pixels at 30 frames/sec (NTSC based) or 352x288 pixels at 25 frames/sec (PAL based), commonly referred to as Source Input Format (SIF) video. MPEG-1 resolution is not only limited to the above sizes, but it in fact may go as high as 4095x4095 at 60 frames/sec.

MPEG 2 is the proposed standard for digital video at 4 to 100 Mbps over transmission facilities capable of such support (fiber optics, hybrid fiber/coax and satellite). While MPEG-2 requires much more bandwidth than MPEG-1, it provides much better resolution and image quality at much greater speed. While there remain technical issues associated with the transmission of MPEG 2 video over ATM, those issues are being addressed. MPEG 2 already has found application in Direct Broadcast Satellite (DBS) services, also known as Direct Satellite Systems (DSS). Such services employ VSAT dishes in competition with CATV, running MPEG at rates of about 3 and 7.5 Mbps. In a convergence scenario, MPEG-2 is the standard of choice, supporting compression rates of about 200:1. *MPEG 3*, designed for HDTV application, was folded into MPEG-2 in 1992. MPEG-2 consists of *profiles* and *levels*. The profile defines the bitstream scalability and the colorspace resolution, while the level defines the image resolution and the maximum bit-rate per profile. Probably the most common descriptor in use currently is Main Profile, Main Level (MP@ML) which refers to 720x480 resolution video at 30 frames/sec, at bit-rates up to 15 Mb/sec for NTSC video

MPEG-4 is a low bit-rate version intended for application in videophones and other small-screen devices

2.1.7 Audio-visual application over IP network

2.1.7.1 Computer Telephony

Application of audio under computer network include computer telephony .This application will mainly be based on Voice over IP. Voice over IP works like that : by digitizing voice over data packets, sending them on the network and reconverting them in voice at destination.

TCP/IP network are amde of IP packects containing a header (to control communication) and payload to transport data : VoIP uses it to go across the network and then come to destination

The purpose of setting up a VoIP communication requires the followings items :

First the Analog to Digital converter (ADC) convert analog voice from microphone to digital signals (bits)

Now the bits have to be compressed in good format for transmission.There is a number of protocol to perform that (MPEG-audio, etc..)

Here we have to insert our voice packets in data packets using RTP (typically over RTP over UDP over IP).

We need a signalling protocol to call users. H.323 do that.

At the receiver, we have to dissemble packets, extract dat, the convert them into anolog voice signals and send them to sound card (or phone)

All that must be done in real time fashion cause we cannot waiting for too long for a vocal answer . Here the Qualiy of Service is considered.

Base Architecture

Voice))-----ADC-----Compression algorithm-----Assembling RTP in TCP/IP

------→

←-----

Voice((-----DAC-----Decompression Algorithm---Disassembling RTP from TCP/IP

Based on that architecture, we have to analyse each part of the process.

Analog to digital conversion

This is made by hardware, typically card integrated ADC

Today very good sound cards allow you to convert with 16 bit a band of 22050 Hz (for sampling it, you nee a frequency of 44100 Hz by Nyquist Principle) obtaining a throughput of 2 bytes(16 bits)*44100 (samples per second)= 88200 bytes/s. For stereo stream, 4 bytes*44100= 176.4 bytes/s

For VoIP, we need'nt 22 Khz bandwidth (and also we need'nt 16 bits). we will analyse other coding used for VoIP.

Techniques for VoIP

PCM (Pulse Code Modulation), standard ITU-T G.711

Voice bandwidth is 4 Khz , so sampling bandwidth has to be 8 khz by Nyquist's Principle.

We represent each sample with 8 bits (having 256 possible values)

Throughput is 8000 Hz * 8bits= 64 kbit/s as typical digital phone line

In real application, μ-law (north America) and a-law (Europe) variants are used which code signal on logarithmic scale using 12 or 13 bits instead of 8 bits. Both laws are G.711 recommendation and both use non-uniform quantization step sizes which increase logarithmically with signal level.

How VoIP Processes a Typical Telephone Call over our network

Before configuring VoIP on a router, it helps to understand what happens at an application level when you place a call using VoIP. The general flow of a two-party voice call using VoIP is as follows:

1. The user picks up the handset; this signals an off-hook condition to the signaling application part of VoIP in the router.

2. The session application part of VoIP issues a dial tone and waits for the user to dial a telephone number.

3. The user dials the telephone number; those numbers are accumulated and stored by the session application.

4. After enough digits are accumulated to match a configured destination pattern, the telephone number is mapped to an IP host via the dial plan mapper. The IP host has a direct connection to either the destination telephone number or a PBX that is responsible for completing the call to the configured destination pattern.

5. The session application then runs the H.323 session protocol to establish a transmission and a reception channel for each direction over the IP network. If the call is being handled by a Private Branch Exchange (PBX), the PBX forwards the call to the destination telephone. If Resource Reservation Protocol (RSVP) has been configured, the RSVP reservations are put into effect to achieve the desired QoS over the IP network.

6. The coder-decoder compression schemes (CODECs) are enabled for both ends of the connection and the conversation proceeds using Real-Time Transport Protocol/User Datagram Protocol/Internet Protocol (RTP/UDP/IP) as the protocol stack.

7. Any call-progress indications (or other signals that can be carried inband) are cut through the voice path as soon as end-to-end audio channel is established. Signaling that can be detected by the voice ports (for example, inband dual-tone multifrequency (DTMF) digits after the call setup is complete) is

also trapped by the session application at either end of the connection and carried over the IP network encapsulated in Real-Time Transport Control Protocol (RTCP) using the RTCP application-defined (APP) extension mechanism.

8. When either end of the call hangs up, the RSVP reservations are torn down (if RSVP is used) and the session ends. Each end becomes idle, waiting for the next off-hook condition to trigger another call setup.

In our case, we limit with the use Quicknet phoneJack and LineJack connected directly to a telephone line instead of PBX, to perfom calls on other type of network.

2.1.7.2 Videoconference interactive meeting

Videoconferencing is the interactive conferencing of multiple individuals on a visual, as well as an audio, basis. Videoconferencing systems, therefore, consist of cameras and monitors, as well as microphones and speakers. The IP videoconference must be supported over a IP network, which can take the form of either a Local Area Network (LAN) or a Wide Area Network (WAN). The overall systems can range from room systems to rollabout systems to PC-based, or desktop, systems. Where participants from more than two sites are conferenced, Multipoint Conference Units (MCUs) also must be involved.

Technical specification

As stated above, the basic dimensions of the video aspect of the communication include frame rate, resolution, chrominance and luminance. In accomplishing this basic process, the codec naturally compresses the signal as it encodes it, with the decoding process resulting in a reasonable approximation of the original signal.

The coding function is associated with the video camera and microphone on the transmit side of the equation. The native information stream comprises a continuous stream of reflected light waves, which vary infinitely in terms of their color variations. The native signal is captured by the analog camera and is sampled in areas, or blocks, known as pixels (picture elements).

Each block is then encoded into digital format, with its approximate color expressed in digital code. The set of digitally expressed binary values associated with the image are then transmitted in serial mode (i.e., one-at-time) across the network.

On the receive side, a codec associated with a presentation device (i.e., TV set or computer monitor, and speaker) reverses the process, presenting the pixels for viewing and reconstructing the analog voice signal.

Note that the audio signal takes priority over the video signal - audio without video is more valuable than is video without audio.

The ITU-T (International Telecommunications Union-Telecommunications Standardization Sector) developed a set of international videoconferencing standards in 1990. The H.320 standard, developed for N-ISDN (Narrowband ISDN), supports networked video and multimedia communications at bit rates ranging from 64 Kbps to 1.920 Mbps, in increments of 64 Kbps.

Video compression, in the H.320 domain, makes use of a standard known as H.261, also known as Px64, where P denotes the number of Bearer channels and 64 denotes 64 Kbps per channel. This standard is rooted in N-ISDN, where clear channel communications of 64 Kbps is supported, and where the number of channels peaks at 30, which is consistent with the international version of ISDN PRI.

H.261 also refers to the video coding standard and the video format. Specifically, H.261 includes the optional CIF (Common Intermediate Format) at 352 x 288 = 101,376 (i.e. width by height) pixels per frame, at a maximum rate of 30 fps. H.261 also supports the mandatory QCIF (Quarter CIF) at 176 x 144 = 25,344 pixels per frame, which is exactly one-quarter the resolution of full CIF.

2.1.7.3 IP TV

TV over IP is broad streaming solution that includes several application, all of which can be implemented on digital broadband network such Ethernet, fibre or Wireless LAN/WAN. TV over IP is being utilised in the following application:

TV in living room (instead of cable TV)

Time-Shifted or Personal Video Recoder (PVR)

Interactive TV

TV on desktop

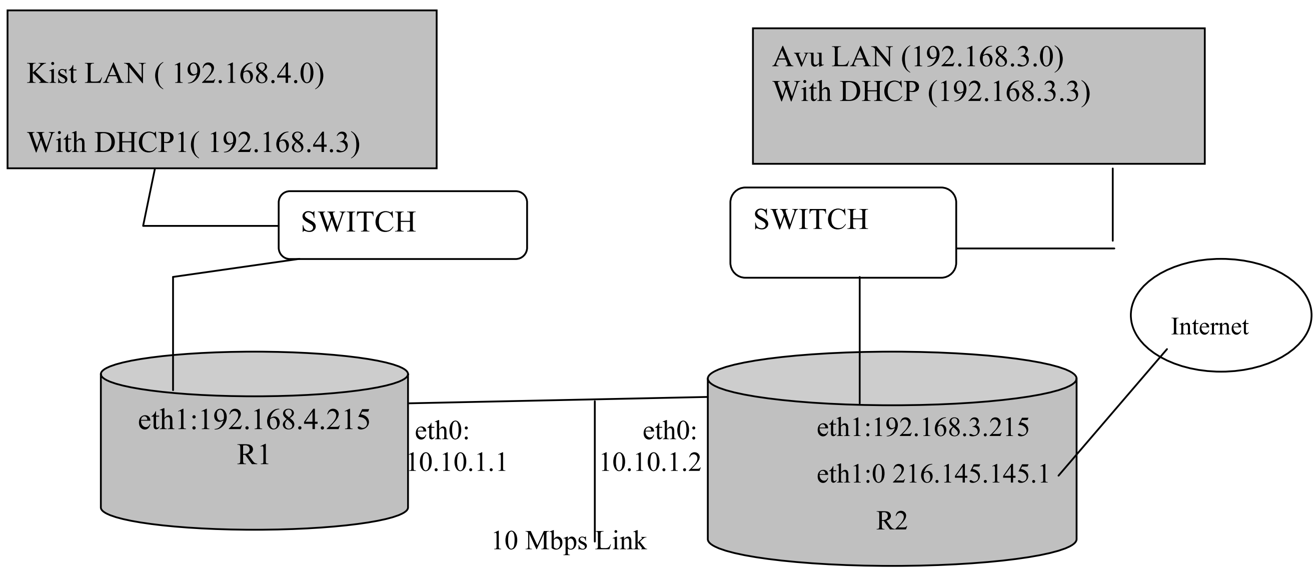
With TV over IP , operators can offer greater level of service to their customers. The fact that the customers receive convergerd services on single pipe an interface with a single provider for all communication needs results in technical maintenance, streamlined billing and hence provide customer services. KIST as ISP, can expand their services through also IP TV broadcasting.

III. NETWORK DESIGN AND REQUIREMENT PLAN

The plan of the network, as well as the design have the main objective to have a platform that will work as model for voice, video and data communication on the single IP. The design focuses on Px64 standards,likely included in H.323 standard for VoIP communication, and audio and video conference. MPEG standars will be used for designing video broadcasting, specifically IP TV.

3.1.Evaluation of the current configuration

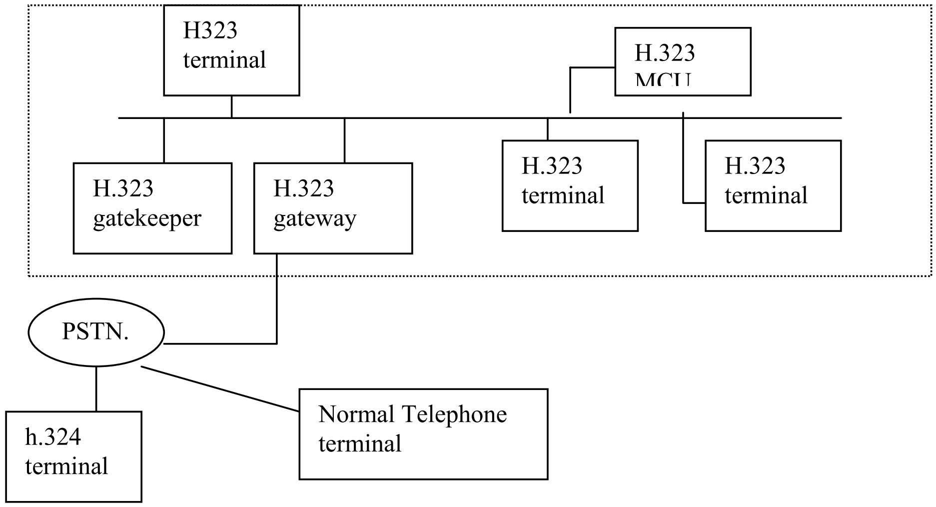
During this project, the network plan is referred to ITU-T recommendation called H.32x recommendation that provides multimedia communication over packet network, notably IP network. According to the set objective, the network will permit VoIP communication inside KIST Campus. This Campus is mainly made of two subnets, Kist Lan Network (192.168.2.0) and AVU LAN network (192.168.3.0). The two subnets are connected through two routers (R1 and R2), one on each side.Another router connect both subnets to Internet. The figure below shows the general architecture of the network.



On this network, mail, HTTP, FTP services, classified as non-real time services are being perform inside the network and on the Internet. The project now aims to give details on various steps to add various multimedia services. This also require configuration on bandwidth and Quality of services needed for real time application over transmitting units of the network. General design of the networks upgradable services will look as follows:

3.2 H.323-VoIP network design

fig:



The H.323 signalling protocol,for example used by Microsoft NetMeeting to make VoIP calls allow a variety of elements talking each other:

- Terminals, clients that initialize VoIP connection. Although terminals could talk together without anyone else, we need some additional elements for a scalable vision.
- Gatekeepers, that essentially operate:
- address translation service, to use names instead IP addresses
- admission control, to allow or deny some hosts or some users
- bandwidth management
- Gateways, points of reference for conversion TCP/IP - PSTN.
- Multipoint Control Units (MCUs) to provide conference.
- Proxies Server also are used.
- h323 allows not only VoIP but also video and data communications. Concerning VoIP, h323 can carry audio codecs G.711, G.722, G.723, G.728 and G.729 while for video it supports h261 and h263.

N.B. :

The gatekeepers, gateways and MCUs are logically separate components of the H.323 standards but can be implemented as single physical device.

3.1.1 Network requirement

Hardware and software requirement

To create a little VoIP system we use the following hardware:

PC 386 or more

Sound card, full duplex capable

a network card to allow communication between 2 PCs over LAN (ex: Kist LAN)

For this project, we use special cards with hardware accelerating capability such as Quicknet PhoneJack wich is a sound card that can use standard algorithms to compress audio stream like G723.1 down to 4.1 Kbps rate. With the Internet PhoneJACK , we use familiar telephone (including your cordless phone) to make and receive Internet phone calls**.** User can plug a standard analog telephone, fax machine or headset into the Internet PhoneJACK and keep phone calls private. The Internet PhoneJACK allows phone to ring when a call arrives and automatically launches the favorite Internet phone application when user picks up the handset. It has a ISA or PCI connector bus.

For software, we can choose what O.S. to use:

- Win9x
- Linux

For end user client, we will use the Win9x because most of terminal on KIST and AVU LAN network are under Windows operating system. Under Win9x client PC, we have applications like Microsoft Netmeeting, Internet Phone, DialPad or others in case of normal full duplex sound card or Internet Switchboard for Internet PhoneJack cards.

To perfom calls with other networks via a VoIP gateway, we need the following hardware and software :

Quicknet LineJack can be connected to a PSTN line allowing VoIP gateway feature

PC 486 and more.

To manage gateway feature (join TCP/IP VoIP to PSTN lines) we need some kind of software like this:

Internet SwitchBoard for Windows systems also acting as a h323 terminal;

PSTNGw for Linux and Windows systems , downloaded from OpenH323.org

To administrative purposes , we can use also Gatekeeper software (from openh323.org) and because the gateway is upgraded on R1, then PSTNGw for Linux is considered.

3.3 IP videoconference design

2.3.1 Network requirements

Hardware and software requirement

To create a simple desktop videoconference,we need the tools:

PC486 and more

video card

Monitor resolution of 640*480

Speaker

Microphone

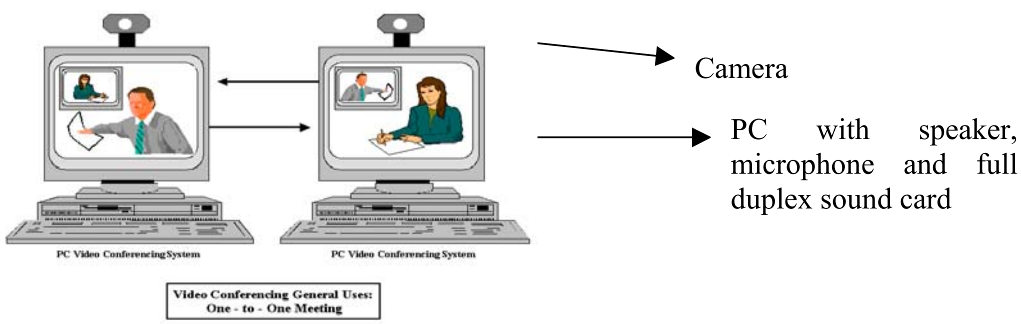
Camera

Network card to communicate over LAN

Microsoft netmeeting or other software to perfom desktop videoconference.

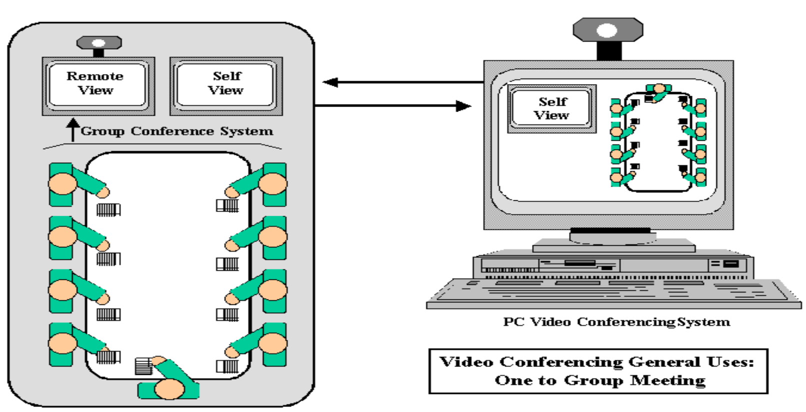
Fig: One to one desktop conferencing

To create one to group conferencing we need to allows routers to enable MPEG2 CODEC to allow broadcast over the network as Video on Demand servers and then also enable routers to multicast and



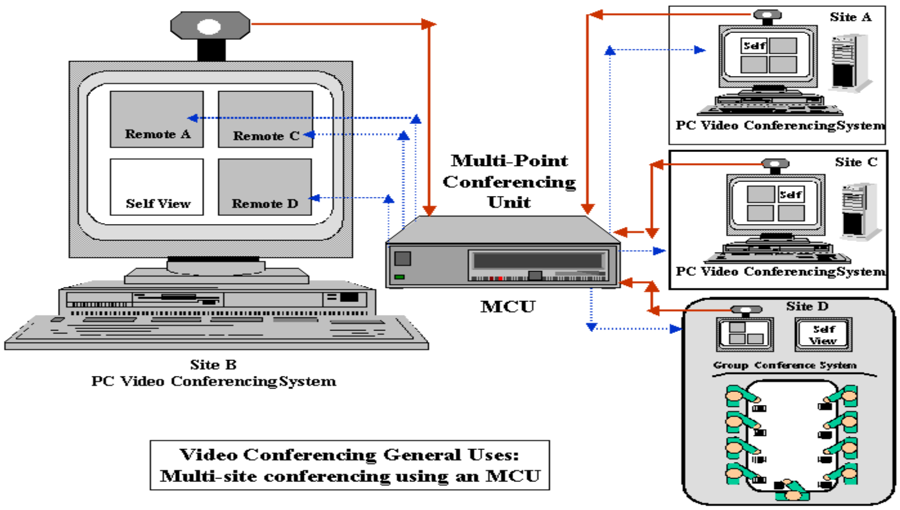
configure, desktop to join the multicast group by sending IGMP request to the routers.

Fig: one to group



To create group to group video conferencing , we use MCU unit. For this project, the MCU unit is a PC486 or more running H.232 MCU application (downlaoded from openh323.org). Gateway is a need in case of conversion from H.323 (TCP/IP) to H.320 (ISDN) and visa versa. The PSTNGw (downlaoded from openh323.org) over linux gateway is used in this project to offer communication over PSTN via ISDN connection.

Fig:



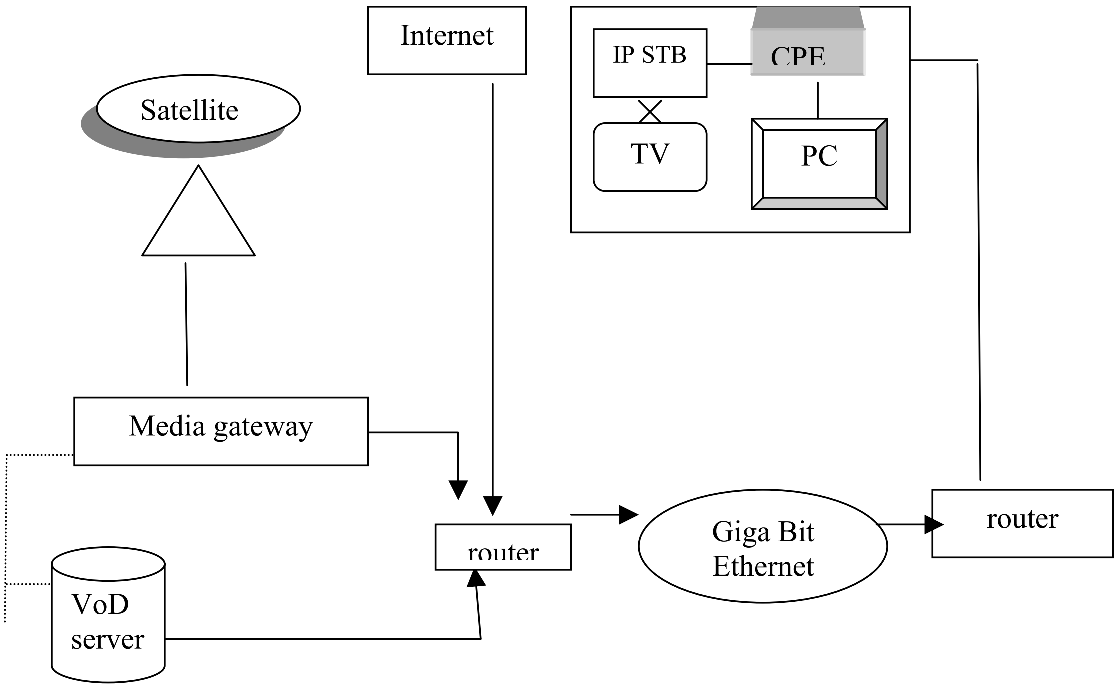
NB:

desktop participating in conference must work as H.323 teminals by installing desktop conferencing software like Microsoft Netmeeting or OpenH323 client software to be compatible with the gatekeeper.

Hardware-assisted Desktop Clients can list with an analog camera, video capture card, software, microphone, and headset.

3.4 IP TV Desing

FIG:



<u>Hardware and software requirement</u>

<u>Media gateway</u>

The media gateway can be a hardware part of the network or software that can installed on router level. The software is caracterised by the number of live streams in real time and pre-encoded streams that are stored on Video Server. The Media gateway transmit the stream to the switch or routers which transfer them on the backbone to central or remote offices.

<u>Video Server</u>

For store and forwards transmission, they store digitally encoded content and stream it through the router via operator's networking infrastructure. It receives also encoded digital signal thti is uploaded from the media gateway.

<u>Router</u>

Router that supports multicast transmission. The routers resides on both sides of the network. On router send and another router receives data into end user Ethernet based network.Video broadcasting compression techniques have to be enabled on routers such as MPEG 2 or MPEG 4 to support wireless networks.

<u>CPE (customer point equipment)</u>

This equipment located at the end point receives the TV/IP stream. The CPE receive stream from router and transfers it directly to PC for display on the desktop or to the IP STB.

<u>IP Set-To-Box</u>

IP STP receives the IP stream, and decodes the stream (with MPEG decoder technique) for display on the TV. Typically, the IP STB receives the stream from CPE. Some STP can receive directly from router.

<u>NB</u>

- The CPE has to transfer the MPEG-2 or MPEG-4 to IP STP which decodes the stream on a TV Set.

MPEG 2 and 4 can be also be played on PCs using the Microsoft Media Player.

The project limits on video on demand server storing entertainment movies and broadcast using MPEG coder software.

Media gateway as software takes configuration on routers to allow streaming video. The software supports

IV. NETWORK IMPLEMENTATION

The network that is manly to be implemented in this project refers to the current data communication in KIST LAN network and AVU LAN. Coding and

4.1 Network addressing

.The network IP adresssing in both network are done under DHCP configuration. The simple Linux configuration of DHCP1 on KIST LAN and DHCP2 on AVU LAN may look as follows:

ON DHCP1:

Configure the network IP address of the DHCP1: #ifconfig eth0 192.168.4.3 netmask 255.255.255.0 broadcast 192.168.4.255

download Dhcp server module on ftp://ftp.isc.org/isc/dhcp/

After download we unpack it. After we do cd into the distribution directory and type: **./configure**

After it is done type: **make** and **make install.** Next step is to add route for 255.255.255.255. In order for dhcpd to work correctly with picky DHCP clients (e.g., Windows 95), it must be able to send packets with an IP destination address of 255.255.255.255. Linux insists on changing 255.255.255.255 into the local subnet broadcast address (here, that's 192.168.4.255). This results in a DHCP protocol violation, and while many DHCP clients don't notice the problem, some (e.g., all Microsoft DHCP clients) do. We type: **route add -host 255.255.255.255 dev eth0.** With the message "255.255.255.255: Unknown host", we're adding the following entry to your /etc/hosts file: 255.255.255.255 all-ones .Then, try:

route add -host all-ones dev eth0 or route add 255.255.255.0 dev eth0

To configure DHCPd, we create or edit /etc/dhcpd.conf and perform this configuration:

```
# Sample /etc/dhcpd.conf
# (add your comments here)
default-lease-time 600;
max-lease-time 7200;
option subnet-mask 255.255.255.0;
option broadcast-address 192.168.4.255;
option routers 192.168.4.215;
option domain-name-servers 192.168.4.1, 192.168.4.2;
option domain-name "kist.local";
subnet 192.168.4.0 netmask 255.255.255.0 {
  range 192.168.4.10 192.168.4.100;
  range 192.168.4.150 192.168.4.200;
}
```

This will result in DHCP server giving a client an IP address from the range 192.168.4.10-192.168.4.100 or 192.168.4.150-192.168.4.200. It will lease an IP address for 600 seconds if the client doesn't ask for specific time frame. Otherwise the maximum (allowed) lease will be 7200 seconds. The server will also "advise" the client that it should use 255.255.255.0 as its subnet mask, 192.168.4.255 as its broadcast address, 192.168.4.215 as the router/gateway and 192.168.4.1 and 192.168.4.2 as its DNS servers.

The same configuration is made on DHCP2 with eth0:192.168.3.3, netmask 255.255.255.0, broadcast 192.168.3.255. The option router is 192.168.3.215 and 192.168.4.1 and 192.168.4.2 as DNS servers. The domain is avu.local.

Routers are configured as follows:

On router R1

Interface eth1:

#iifconfig eth1 192.168.4.215 netmask 255.255.255.0

#route add –net 10.10.1.0 eth1

Interface eth0 :

#ifconfig eth0 10.10.1.1 netmask 255.255.255.0.

#route add –net 10.10.1.0 eth0

#route add –net 192.168.3.0 eth1

#route add default gw 216.147.145.1

On router R2

Interface eth1:

#iifconfig eth1 192.168.3.215 netmask 255.255.255.0

#route add –net 10.10.1.0 eth1

#ifconfig eth1:0 216.147.145.1

#route add default gw 216.147.145.1

Interface eth0 :

#ifconfig eth0 10.10.1.2 netmask 255.255.255.0.

#route add –net 10.10.1.0

#route add –net 192.168.4.0 eth1

#route add default gw 216.147.145.1

After this addressing evaluation, the data application sharing is performed over the both subnets. The Voice communication can be corporate on this by performing additional hardware setup and network configuration to offer also voice communication, and future accomodation of video and data conferencing. The DHCP is mainly used to easy IP addressing management in case of the network growth like network offering audio visual services with easy IP addressing.

4.2 VoIP network implementation

4.2.1 **Client PC setup**

As we saw, Quicknet Phonejack is a sound card with VoIP accelerating capability on Client PC. It supports:

G.711 normal and mu/A-law, G.728-9, G.723.1 (TrueSpeech) and LPC10.

Phone connector (to allow calling directly from your phone) or

Mic & speaker jacks.

Quicknet PhoneJack is a ISA (or PCI) card to install into your Pc box. It can work without an IRQ.

Under Windows we have to install:

Card driver

Internet Switchboard application

When we pick up the phone Internet Switchboard wakes up and waits for the calling number (directly entered from the phone), as follows:

enter an asterisk, then type an IP number (with asterisks in place of dot) with a # in the end

type directly a PSTN phone number (with international prefix) to call a classic phone user. In this case the registration to a gateway manager to which pay for time is needed ex: Denwa.com.

enter directly a quick dial number (up to 2 digits) you have previously stored which make a call (IP or PSTN).

Internet Swichboard is h323 compatible, so if you can use, for example, Microsoft Netmeeting at the other end to talk.

In place of Internet Switchboard you can use openh323 application openphone (using GUI) or ohphone (command line).

4.2.2Gateway setup

The setup is made on the Linux or Windows PC gateway that can be incorporated oN AVU LAN (192.168.3.0). First of all, we setup the quicknet Linejack that allows gateways features with PSTN over ISDN line

4.2.3 **Call configuration**

Simple communication is made as follows:

A (Win9x+Sound card) - - - B (Win9x+Sound card)

192.168.4.15 - - - 192.168.4.20

192.168.4.15 calls 192.168.4.20.

A and B should:

have Microsoft Netmeeting (or other software) installed and properly configured.

have a network card or other kind of TCP/IP interface to talk each other.

In this kind of view A can make a H323 call to B (if B has Netmeeting active) using B IP address. Then B can answer to it if it wants. After accepting call, VoIP data packets start to pass.

Using names

If you use Microsoft Windows in a lan you can call the other side using NetBIOS name. NetBIOS is a protocol that can work (stand over) with NetBEUI low level protocol and also with TCP/IP. It is only need to call the "computer name" on the other side to make a connection.

```
       A          - - -          B
192.168.4.15     - - -     192.168.4.20
       Peter         - - -          Maggy
```

Peter calls Maggy.

This is possible cause Peter call request to Alice is converted to IP calling by the NetBIOS protocol.

The above 2 examples are very easy to implement but aren't scalable.

In a more big view such as Internet it is impossible to use direct calling cause, usually, the callers don't know the destination IP address. Furthermore NetBIOS naming feature cannot work cause it uses broadcast messages, which typically don't pass ISP routers .

The following voice communication can be also being perfomed the two subnet (192.168.4.x and 192.168.3.x)throught routers R1 and R2 with QoS configured

```
       A          - - -          D
192.168.4.15     - - -     192.168.3.20
       Peter         - - -          Mary
```

Peter calls Mary.

Internet calling using a WINS server

The NetBIOS name calling idea can be implemented also in a Internet environment, using a WINS server: NetBIOS clients can be configured to use a WINS server to resolve names.

PCs using the same WINS server will be able to make direct calling between them.

```
A (WINS Server is  S) - - - - I  - - - -  B (WINS Server is S)
                              N
                              T
                              R  - - - - -   S (WINS Server: 192.168.4.1)
C (WINS Server is S) - - - -  A
                              N
                              E  - - - -  D (WINS Server is S )
                              T
```

Intranet communication

A, B, C and D are in different subnets, but they can call each other in a NetBIOS name calling fashion. The needed is that all are using S as WINS Server.

Note: WINS server hasn't very high performance cause it use NetBIOS feature and should only be used for joining few subnets.

Internet calls

A problem of few real IPs is commonly to communicate can be solved using the so called masquering (also NAT, network address translation): there is only 1 IP public address,215.147.145.1 (that Internet can directly "see"), the others machines are "masqueraded" using all this IP.

In the example A,B and C can navigate, pinging, using mail and news services with Internet people, but they CANNOT make a VoIP call. This because H323 protocol send IP address at application level, so the answer will never arrive to source (that is using a private IP address).

A (192.168.4.15- - -

B (192.168.3.15) - - - Router with NAT(216.147.1451) - - - Internet

C (192.168.3.50)- - -

The call doesn't work

there is a Linux module that modifies H323 packets avoiding this problem. First, we download the module from linuxdoc.org and install the module on Routers. To install it we first copy it to source directory specified, modify Makefile and go compiling and installing module with "modprobe ip_masq_h323". Unfortunately this module cannot work with ohphone software . the configuration is as follows:

put the file ip_masq_h323.c in the directory /usr/src/linux/net/ipv4

edit the file /usr/src/linux/ipv4/Makefile

modify the line like this:

M_OBJ+= ip_masq_ftp.o ip_masq_irc.o

Ip_masq_raudio.o ip_masq_quake.o ip_ masq_h323.o

-Be sure to enable the module in the Kernel

-Issue “make modules modules_install” from /usr/src/linux

-Try to “ insmod ip-masq_h323 and look the log file “/var/log/messages

A - - - Router with NAT

B - - - + - - - Internet

C - - - ip_masq_h323 module

This works

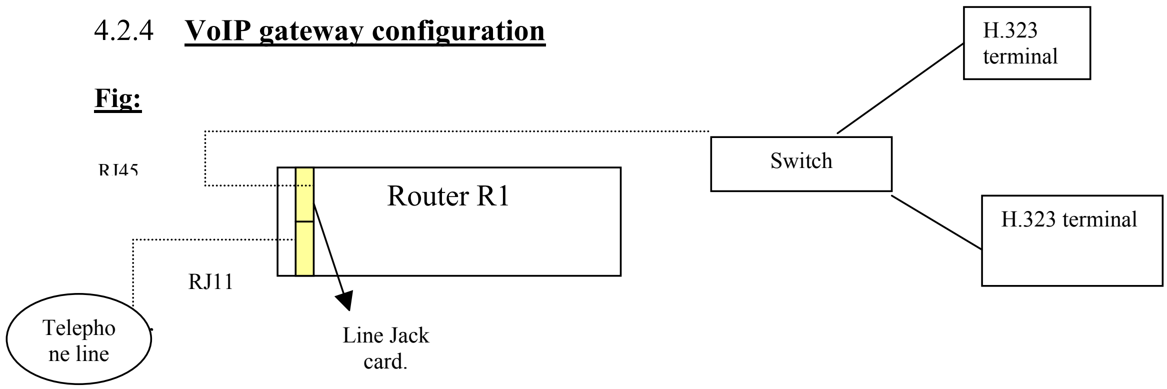
There is a application program that also solves this problem:

B - - - PhonePatch - - - Internet

This works

4.2.4 **VoIP gateway configuration**

Fig:



Router R1 can work as Voice over IP gateway by installing the Line jack card under Linux OS and configure it as follows:

Installing the kernel

1. Download the linux-2.2.13 kernel source tarfile and install it in /usr/src/linux as usual.

2. Download the pre-patch-2.2.14-16 patch file and apply it to the kernel.

3. Configure your kernel with telephony and Quicknet support as modules. Then execute:

    make clean dep bzlilo modules modules_install

or your favorite command to compile and install the kernel.

Reboot and make sure the new kernel is running OK

Installing the modules

5. We should now be able to install kernel modules for the card. - now we need to specify two IO ports as following for the LineJack:

/sbin/insmod phonedev
/sbin/insmod ixj dspio=0x300 xio=0x310

We should hear the relay on the LineJack go "click" when the module installed correctly, and the lights on the back of the card should light up. If not, something is wrong.

Installing the Linux SDK

6. Download and unpack the Quicknet Linux SDK from:

ftp://ftp.quicknet.net/Developer/Linux/Docs/linux-ixj-0.3.4-sdk.tgz

It contains lots of useful utilities that are very useful in testing installed cards. We ignore the support document during installation

7. Move or symlink the file /usr/src/linux/drivers/telephony/ixjuser.h to /usr/include/linux/ixjuser.h. This file has been installed in the wrong place

8. Search through all of the .c files in the SDK directory and change all occurences of:

TELEPHONY_EXCEPTION to union telephony_exception

9. Search through all of the .c files in the SDK directory and change all occurences of:

PHONE_CAPABILITY to struct phone_capability

10. We should be able to compile the SDK programs by typing make. Then try running a few like ring or playtone to check that the card is working.

The next is to install the PSTNGw software and configure it

V. SUMMARY AND CONCLUSION

The audio visual computer network developed under this project illustrates a set of adaptation that mainly require streaming audio and video using standards like MPEG and PX64 for video and voice and H.323 to support real time transmission. For this project, the information stream is encoded by G.723 coder for telephony specified by H.323 standards using PhoneJack and LineJack sound card. Video information is encoded by H.261 and H.263 for videoconference and MPEG2 for audio and video broadcast.

In this study, the focus was on deployment of a network that could support audio and video services to our campus over IP network already performing non-real data communication. I investigate on the main applications of audio visual network including telephony, audio and video conference and TV over IP network. I have designed a prototype of those applications by specifying requirement to implement each system. The implementation focuses on Voice over IP network deployment between two subnets of KIST campus.