Video over the Internet
(Technology and Economic Viability)

Johns Hopkins University
Image Compression and Packet Video, Fall 2002.
Final Project
Ogechi C Wachuku

Introduction:

Video over the Internet has been around for many years. It is a technology that we are very well familiar with. Most of us have used applications such as Real Player and Windows Media player to view video either directly from a website or of clips we have received via email. Video over the internet can be delivered to the consumer in many different ways, such as by streaming, saving the entire clip to the user's hard drive before actual playback, or buffering the clip and playing it back in sections.

The technology has great potential for success in the economic arena. Most of the same sources of revenue that apply to television broadcast can also apply to the transmission of video over the Internet. These include revenue from advertising, pay per view and private broadcast services such as cable TV. Other non-traditional sources of revenue for video over the Internet include video conferencing, and video on demand.

Unfortunately the issue is not determining the types of services that video over the Internet can provide because television has given this new technology a lot to live up to. The issue for video over the Internet is to provide the consumer with the same quality of service that they receive from video transmitted over other media This means that video over the Internet should be able to handle live broadcasts, delay free transmission, error free transmissions, e.t.c. These requirements expose the problems that have plagued Video over the Internet since its inception and they lie at the heart of the issue of its economic viability.

Problems:

The major problem of video over the Internet has to do with the architecture of the Internet. The Internet is not suited for transfer of video because it randomly looses packet and it has unpredictable delays. Other problems include the high data storage requirements for video and the limited bandwidth available to most consumers.

To address the problem of the unreliability of Data Transmission over the Internet most applications choose to use UDP (Unreliable data Protocol) over TCP (Transmission control protocol). The reason for this is that even though TCP is designed to minimize packet loss the delays introduced by this protocol make it unsuitable for Video. UDP is best suited for real-time because it does not retransmit lost packets and especially because it minimizes transmission time. The applications on the receiver side can employ methods such as error concealment to address lost packets.

The current encoding methods available for video compression are very effective in addressing storage requirements, and with the increase in the bandwidth available to the end user, error correction is easier to deliver.

Video compression Standards

The excerpt below is taken from the paper "A review of Video Streaming over the Internet" – Jane Hunter, Varuni Witana, Mark Antoniades. I have included it in its entirety for this section because it does an excellent job of explaining the Video compression standards.

> "The most important video codec standards for streaming video are H.261, H.263, MJPEG, MPEG1, MPEG2 and MPEG4. A brief description of these is given below. Compared to video codecs for CD-ROM or TV broadcast, codecs designed for the Internet require greater scalability, lower computational complexity, greater resiliency to network losses, and lower encode/decode latency for video conferencing. In addition, the codecs must be tightly linked to network delivery software to achieve the highest possible frame rates and picture quality. As one looks at the existing codec standards, it becomes apparent that none are ideal for Internet video. In fact, it is quite clear that over the next few years, we will see a host of new algorithms that are specifically designed for the Internet and are thus more suitable for it. Research is currently underway looking at both new scalable, flexible codecs and ways of scaling existing codecs using transcoding and filters. Section 3 outlines current research in video scalability. New algorithms specifically targeted at Internet video are being developed. Consequently application framework standards such as H323/H.324 for videoconferencing and MPEG4, are being designed that will easily incorporate these new codec innovations into applications being developed today, without significant rework.
>
> **H.261**
>
> H.261 is also known as P*64 where P is an integer number meant to represent multiples of 64kbit/sec. H.261 was targeted at teleconferencing applications and is intended for carrying video over ISDN - in particular for face-to-face videophone applications and for videoconferencing. The actual encoding algorithm is similar to (but incompatible with) that of MPEG. H.261 needs substantially less CPU power for real-time encoding than MPEG. The algorithm includes a mechanism which optimises bandwidth usage by trading picture quality against motion, so that a quickly-changing picture will have a lower quality than a relatively static picture. H.261 used in this way is thus a constant-bit-rate encoding rather than a constant-quality, variable-bit-rate encoding.
>
> **H.263**
>
> H.263 is a draft ITU-T standard designed for low bitrate communication. It is expected that the standard will be used for a wide range of bitrates, not just low bitrate applications. It is expected that H.263 will replace H.261 in many applications. The coding algorithm of H.263 is similar to that used by H.261, however with some improvements and changes to improve performance and error recovery. The differences between the H.261 and H.263 coding algorithms are listed below. Half pixel precision is used for H.263 motion compensation whereas H.261 used full pixel precision and a loop filter. Some parts of the hierarchical structure of the datastream are now optional, so the codec can be configured for a lower datarate or better error recovery. There are now four optional negotiable options included to improve performance: Unrestricted Motion Vectors, Syntax-based arithmetic coding, Advance prediction, and forward and backward frame

prediction similar to MPEG called P-B frames. H.263 supports five resolutions. In addition to QCIF and CIF that were supported by H.261 there is SQCIF, 4CIF, and 16CIF. SQCIF is approximately half the resolution of QCIF. 4CIF and 16CIF are 4 and 16 times the resolution of CIF respectively. The support of 4CIF and 16CIF means the codec could then compete with other higher bitrate video coding standards such as the MPEG standards.

MJPEG

There is really no such standard as "motion JPEG" or "MJPEG" for video. Various vendors have applied JPEG to individual frames of a video sequence, and have called the result "M-JPEG". JPEG is designed for compressing either full-color or gray-scale images of natural, real-world scenes. It works well on photographs, naturalistic artwork, and similar material; not so well on lettering, simple cartoons, or line drawings. JPEG is a lossy compression algorithm which uses DCT-based encoding. JPEG can typically achieve 10:1 to 20:1 compression without visible loss, 30:1 to 50:1 compression is possible with small to moderate defects, while for very-low-quality purposes such as previews or archive indexes, 100:1 compression is quite feasible. Non-linear video editors are typically used in broadcast TV, commercial post production, and high-end corporate media departments. Low bitrate MPEG-1 quality is unacceptable to these customers, and it is difficult to edit video sequences that use inter-frame compression. Consequently, non-linear editors (e.g., AVID, Matrox, FAST, etc.) will continue to use motion JPEG with low compression factors (e.g., 6:1 to 10:1).

MPEG-1

MPEG 1, 2 and 4 are currently accepted, draft and developing standards respectively, for the bandwidth efficient transmission of video and audio. The MPEG-1 codec targets a bandwidth of 1-1.5 Mbps offering VHS quality video at CIF (352x288) resolution and 30 frames per second. MPEG-1 requires expensive hardware for real-time encoding. While decoding can be done in software, most implementations consume a large fraction of a high-end processor. MPEG-1 does not offer resolution scalability and the video quality is highly susceptible to packet losses, due to the dependencies present in the P (predicted) and B (bi-directionally predicted) frames. The B-frames also introduce latency in the encode process, since encoding frame N needs access to frame N+k, making it less suitable for video conferencing.

MPEG-2

MPEG 2 extends MPEG 1 by including support for higher resolution video and increased audio capabilities. The targeted bit rate for MPEG 2 is 4-15Mbits/s, providing broadcast quality full-screen video. The MPEG 2 draft standard does cater for scalability. Three (3) types of scalability; Signal-to-Noise Ratio (SNR), Spatial and Temporal, and one extension (that can be used to implement scalability) Data Partitioning, have been defined. Compared with MPEG-1, it requires even more expensive hardware to encode and decode. It is also prone to poor video quality in the presence of losses, for the same reasons as MPEG-1. Both MPEG-1 and MPEG-2 are well suited to the purposes for which they were developed. For example, MPEG-1 works very well for playback from CD-ROM, and MPEG-2 is great for high-quality archiving applications and for TV broadcast applications. In the case of satellite broadcasts, MPEG-2 allows >5 digital channels to be encoded using the same bandwidth as used by a single analog channel today, without sacrificing video quality. Given this major advantage, the large encoding costs are really not a factor. However, for existing computer and Internet infrastructures, MPEG-based solutions are too expensive and require too much bandwidth; they were not designed with the Internet in mind.

MPEG-4

The intention of MPEG 4 is to provide a compression scheme suitable for video conferencing, i.e. data rates less 64Kbits/s. MPEG4 will be based on the segmentation of audiovisual scenes into AVOs or "audio/visual objects" which can be multiplexed for transmission over heterogeneous networks. The MPEG-4 framework currently being developed focuses on a language called MSDL (MPEG-4 Syntactic Description Language). MSDL allows applications to construct new codecs by composing more primitive components and providing the ability to dynamically download these components over the Internet. This philosophy is similar to that for the multimedia APIs being developed for Sun Microsystems Java, where it will be possible to dynamically download codec components. This trend is also seen in products from major vendors such as Microsoft and Netscape, where they allow for multiple audio and video codecs to be plugged into their real-time streaming solutions." [1]

Architecture for Video Streaming Systems:

The basic architecture of streaming servers consists of seven components. These seven components are: Video compression, Application QoS control for streaming video, Continuous media distribution services, Streaming servers, Media Synchronization Mechanisms Protocols for streaming Media and Streaming video over wireless IP networks.

These areas are illustrated in the figure below:

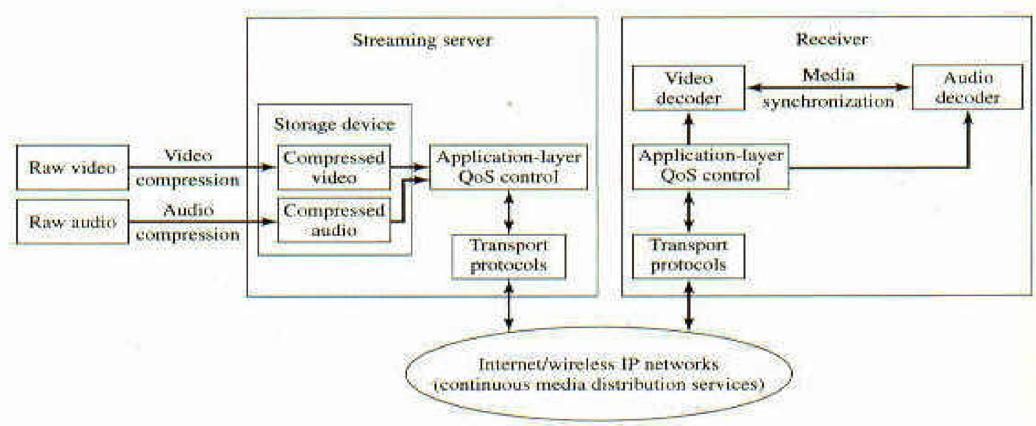


Figure 15.1 Architecture for video streaming.

Figure: Video Processing and Communications, Wang, Ostermann, Zhang. Page 584

1. Video compression: This feature deals with the compression of the raw video to maximize the transmission bandwidth. Compression schemes may be categorized as either scalable or non-scalable.

2. Application Layer QoS control for streaming Video: This deals with providing the user with different presentation quality based on their network conditions. Some of the techniques that have been used are congestion control and error control. Congestion control attempts to prevent packet loss and reduce delay. Error control is employed to improve the data in the presence of packets loss.

3. Continuous media distribution services: These services include network filtering, application level multicast, and content replication. Its goal is to provide QoS and efficiency for streaming video over the Internet.

4. Streaming servers: These are required to process multimedia data under timing constraints and support interactive control operations. These streaming servers also must retrieve media components in a synchronous fashion. They typically consist of a communicator, an operating system, and a storage system.

5. Media Synchronization Mechanisms: These enable the application at the at the receiver side to present media streams in the same way as it was originally

captured. This enables the movement of the speaker's lips to correspond to the audio presented.

6. Protocols for streaming media: These protocols provide services such as network addressing, transport, and session control. They can be classified into three categories (1) network-layer protocol e.g. IP (2) transport protocol e.g. UDP and (3) session control protocol e.g. RSTP.

7. Streaming video over wireless IP networks: An adaptive framework has been proposed to address the challenges posed by the fluctuations of wireless channel conditions. The adaptive framework includes scalable video representations, network-aware video applications, and adaptive services. [4]

Major Video streaming applications:

The major video streaming applications which employ the architecture outlined above in one way or another are Progressive Networks - Real Video, Xing Technology - Stream works, VDONets - VDOlive, and others.

Real Video uses the RSTP protocol on top of UDP. Users can choose between fixed or optimized frame rate encoding in the new Real Video encoder. They can also choose from a number of predefined encoding templates that correspond to the appropriate audio and video formats for a given bandwidth. You also have the choice of having your video stream delivered by TCP, UDP or UDP multicast.

Stream works uses UDP/IP. Content can be encoded at 8.5,24,56 or 112 kbps depending on the bandwidth capabilities of the users. It uses a process called “thinning” which reduces a high-bandwidth stream to improve its transmission over a low bandwidth connection. Streamworks uses MPEG –1 based compression algorithm, which builds an image from a hierarchy of frames. In its compression algorithm it drops B and P frames and transmits only the I-frames. This allows more control over the number and quality of frames that can be delivered over the system each second. By dropping the transmission of B and P frames for low bandwidth connections the error introduced by the loss of that data is minimized providing better picture resolution at those low speeds.

VDOlive uses its own proprietary algorithm centered on wavelet compression. It offers high quality at low bandwidths and scalability. This wavelet algorithm is implemented by dividing the each video frame into multiple layers of which each provide additional detail and image quality. The compression factor determines the number of layers transmitted to the user. The greater the bandwidth the higher the number of layers that will be transmitted to the users machine, therefore the video quality id directly dependent on the bandwidth available to the user. This allows VDOnet to use a high bandwidth file to service differing connection speeds.

All of the systems mentioned above offer scalability. It can either be dynamic, dependent on current network conditions, or it may require the user to choose from a range of encoding formats which correspond to a range of bandwidths. [1][3]

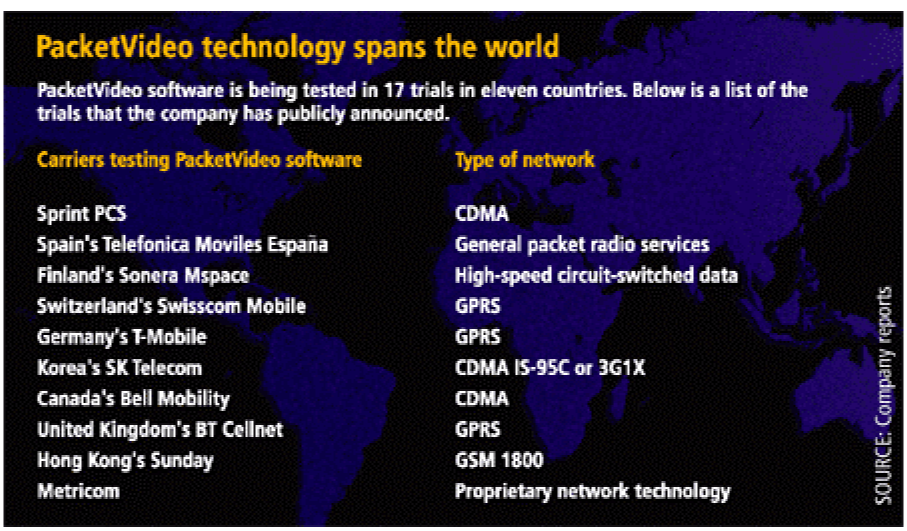
Business Models of Video over Internet:

The applications described above are based on the model that uses proprietary encoding algorithms to provide consumer with software that can create, transmit and receive video over the Internet. These companies make money by licensing the software that performs these operations to other businesses and end users. This is at lowest level of the revenue stream for the technology. Building on top of that some companies offer training courses via Internet using the functionality provided by the lower level software. Also websites such as LAUNCHCAST (launch.yahoo.com) use video technology to provide music videos to consumers and make money from advertising that is also offered via video stream. Another site that offers video streams is IFILM (www.ifilm.com) The business model of this website is to offer short films as content for consumers but the cost to the consumer is to watch a commercial before that can see the video content. Other business models using video over the Internet include remote security surveillance using web cams that provide constant video streams from a particular location. These business models show that Video over the Internet is indeed an economically viable technology. As the technology matures there will be more opportunities especially with the transmission of video over wireless.

Future Directions:

Much of research into future technology involves offering Video over wireless networks. High bandwidth fluctuations and high bit error rates require the video over wireless applications to be network aware. Other specific design requirements for wireless devices, such as low power requirements to preserve battery life need specialized applications that take those requirements into consideration.

PacketVideo a San Diego, CA, company that uses a standard based MPEG4 video compression technology delivers video over wireless at Data rates of 14.4kbs and higher. The company plans to make money by licensing its software to equipment makers. Device manufacturers and wireless carriers have embraced the technology and it is currently undergoing trials all over the world. It also recently obtained a patent for software that maximizes battery life.



PacketVideo technology spans the world

PacketVideo software is being tested in 17 trials in eleven countries. Below is a list of the trials that the company has publicly announced.

Carriers testing PacketVideo software	Type of network
Sprint PCS	CDMA
Spain's Telefonica Moviles España	General packet radio services
Finland's Sonera Mspace	High-speed circuit-switched data
Switzerland's Swisscom Mobile	GPRS
Germany's T-Mobile	GPRS
Korea's SK Telecom	CDMA IS-95C or 3G1X
Canada's Bell Mobility	CDMA
United Kingdom's BT Cellnet	GPRS
Hong Kong's Sunday	GSM 1800
Metricom	Proprietary network technology

SOURCE: Company reports

[2]

In trials being conducted in Finland with Sonera Mspace the technology attracted so much interest that the Sonera had to deny some people the service because it was a limited trial.

If the technology catches on it will be a whole new revenue stream for wireless carriers and equipment makers.

Conclusion:

Video over Internet has definitely caught on; there are many revenue streams that have been based on the technology. The ones mentioned here, software licensing, advertising, and online training are some of the most popular in the market. The inherent problems that exist due to the architecture of the Internet are being addressed through the release of better video standards such as MPEG4. The technology is definitely economically viable its future looks bright.

References:

[1] A review of Video Streaming over the Internet:

http://archive.dstc.edu.au/RDU/staff/jane-hunter/video-streaming.html

[2] Streaming Video Attempts new Heights:

http://www.wirelessinternetmagazine.com/news/0106/0106_ideas_streaming.htm

[3] Video over the Internet:

http://www.rad.com/networks/1996/video.htm

[4]. Wang, Y., Ostermann, J. and Zhang, Y. *Video Processing and Communications*

Upper Saddle River, NJ: Prentice Hall, 2002, 520-522