

TERENA TF-VVC Activity G deliverable
Videoconference archiving and streaming state of the art

Author:

András Kovács <akov@niif.hu>, NIIF/HUNGARNET

<http://www.terena.nl>
<http://www.vidkonf.niif.hu>

Version control:

Date	Version	Author	Comment
July 06, 2006	v0.0	András Kovács	Table of Contents, Basics of part, ...
July 11, 2006	v0.1	András Kovács	Update of existing parts, all-in-one devices chapter
July 13, 2006	v0.2	András Kovács	Provider scale archiving, update for many parts
August 22, 2006	v0.3	Cătălin Meiroşu	Typos

Table of Contents

1	Introduction.....	3
2	Related protocols and technologies.....	4
2.1	Standards-based IP videoconference.....	4
2.2	Streaming.....	5
3	Basics of videoconference streaming and archiving.....	6
3.1	Basic archiving and streaming techniques.....	6
3.1.1	Conventional recording.....	6
3.1.2	Recording and streaming with a computer and a capture card.....	7
3.1.3	Built-in endpoint capabilities.....	8
3.2	Overview of built-in MCU features.....	8
3.3	Streaming and archiving content.....	8
4	All-in-one devices.....	10
4.1	All-in-one archiving/streaming boxes.....	10
4.1.1	General requirements.....	10
4.1.2	User control over streaming/recording initialization.....	11
4.1.3	Playback and retrieval possibilities.....	13
4.1.4	H.239 streaming/archiving and playback solutions.....	14
4.1.5	Security considerations.....	15
4.1.6	Basic/potential problems of VCR devices.....	15
4.1.7	Provider scale archiving.....	17
4.1.7.1	A possible service model.....	17
4.1.7.2	Problems of implementation.....	18
4.1.7.3	Proposal for a minimal API functionality.....	19
5	Conclusions.....	20

1 Introduction

Archiving videoconferences is obviously a significant issue in order to keep important video conversations for later playback or for any other possible usage. This particular segment of videoconferencing has never been a clear area with very few out-of-the-box solutions.

This deliverable aims at the detailed investigation of possible methods and currently available equipment for recording and archiving videoconferences. The following main aspects of videoconference archiving and streaming will be covered in this deliverable:

- Overview on related protocols and network technologies
- Basics of videoconference recording
- A short overview of videoconference archiving equipment currently available on the market
- All-in-one devices archiving solutions
- Provider scale archiving

This document is made in the framework of TERENA TF-VVC (Task Force - Video, Voice and Collaboration). The content of this deliverable is entirely based on experiences and related activities at NIIF/HUNGARNET, the Hungarian Academic and Research Network.

The most of the information in this document only applies to IP based videoconference, mainly using H.323 and SIP protocols. Experiences with the SIP protocol are quite limited due to insufficient software support from manufacturers.

2 Related protocols and technologies

This document only covers standards-based videoconference and streaming technologies and does not deal with any proprietary videoconference applications or hardware devices that are not capable of internetworking due to not being compliant with international standards or recommendations. A service provider must always prefer standard base solutions in order to allow using of diverse network and terminal equipment ensuring independence from certain manufacturers.

In addition, this document does not aim to give a detailed overview of protocols and codecs involved in videoconference and streaming technologies as there are many other documents and freely available descriptions dealing with this topic.

2.1 Standards-based IP videoconference

The world of standards-based videoconference is entirely built on two protocols:

- **H.323:** ITU-T recommendation for implementing a multimedia conferencing protocol for IP networks (more generally for packet switched networks) enabling users to exchange realtime audio, video and data both in point to point or multipoint setups. H.323 refers to a huge number of related recommendation documents which all together build up the world of H.323.

H.323 is often referred as a protocol umbrella as ITU-T recommendation H.323 references many related protocols such as H.225.0 (call signaling), H.245 (multimedia control protocol), H.235 (security), H.450 (supplementary services) and many, many others.

The H.323 protocol suite defines basic network elements such as:

- **Terminals:** endpoints, such as professional hardware based terminals, software based conferencing systems, softphones, IVR devices, voice mail servers, etc.
- **Multipoint Control Unit (MCU):** an MCU is capable of establishing multipoint videoconferences allowing 3 or more participants to share the same conference and interact with each other. An MCU device is responsible for handling a number of point-to-point video calls and distributing video data among the participants connected to the same conference. This requires audio and video switching and mixing capabilities supported by mostly purpose built hardware.
- **Gatekeeper:** gatekeepers are optional elements of a H.323 network but they provide many essential network services such as address resolution (i.e. E.164 phone numbers to IP addresses), call routing, call admission control, endpoint authorization and call accounting, etc.
- **Gateway:** a gateway device is capable of protocol translation in order to ensure connectivity between two networks separated by different signaling protocols (e.g. a H.320-H.323 gateway connects ISDN networks to H.323 based VoIP networks).
- **SIP (Session Initiation Protocol):** IETF protocol defined by RFC 3261. SIP protocol was designed for initiating, modifying and terminating interactive sessions carrying multimedia communication data including voice, video, instant messaging, online gaming, etc. The world is obviously moving towards SIP as H.323 protocol is too complicated so it makes difficult to improve products and ensure sufficient interoperability. Although, this transition is being undertaken, SIP videoconference software implementations are still immature and should be carefully treated in most of the cases.

Standards-based videoconference endpoints and applications almost solely use the following video codecs, all specified by ITU-T recommendations:

- **H.261:** similar to MPEG1 codec, mandatory component of all H.323 video endpoints.
- **H.263:** similar to MPEG2 encoding, implemented by most modern videoconference endpoints.
- **H.264:** new codec based on MPEG4 like encoding, providing very good video quality even using low bandwidth network connections. Implemented by most up to date videoconference endpoints.

H.264 requires multiple times more processing power from the encoder entity (i.e. software or hardware) due to high complexity of the codec.

2.2 Streaming

Streaming is considered a very important network technology of nowadays in cheap and effective dissemination of knowledge and information by supporting distance learning and teaching, audio/video broadcasting (i.e. Internet radio or TV, etc.), broadcast of a specific event to the rest of the world (i.e. conferences, seminars, lectures, etc.). Streaming technology is about one way transmission of audio, video and possibly other content (so called “rich media”) to an end user having a personal computer and an Internet connection.

When speaking about videoconference archiving, posterior methods of retrieval of archive content must be specified. Obviously, streaming is the most important of all of these methods, as it is the cheapest and most effective way of provisioning of archive recordings. Streaming enables all Internet users to access realtime audio/video content very easily with the help of an ordinary computer, a media player software and a speaker set. Realtime streaming of videoconferences is also having a great importance as a videoconference – requiring extremely expensive H.323 and SIP devices – can easily be expanded with passive participants not owning any videoconference endpoint or application.

Streaming data delivery can be implemented both with unicast or multicast transportation. However, multicast technology needs additional technical expertise from network operators, and in addition interdomain multicast traffic exchange is rarely implemented among service providers. Thus, it is advisable to enable both unicast and multicast transportation at media servers.

Streaming control protocols are also play a key role in accessing the content stored at media servers using streaming capable media players. Control protocols are used mainly for requesting appropriate media stream(s) from the server (i.e. by providing an URL to the server), for negotiating basic network transmission parameters (e.g. such as transport protocols (TCP/UDP), port intervals, etc.) and for providing description of each stream or substream (i.e. several audio and video channels) encoding in order to identify media codecs to be used by the media player. The most often used streaming control protocols are:

- **RTSP (RealTime Streaming Protocol):** IETF's open protocol described by RFC 2326. RTSP is a client-server multimedia control protocol, designed to address the needs for efficient delivery of streamed multimedia over IP networks.
- **HTTP:** often used for TCP based streaming when firewall traversal or proxy usage are important. Media player connects to a HTTP URL and can access the stream through the web server, which is basically a simple file download due to the nature of HTTP protocol.
- **MMS (Microsoft Multimedia Streaming / System):** Microsoft's proprietary streaming protocol used with Windows Media Player and Windows Media Services.

Realtime audio/video content is encoded using different media codecs, depending on the encoder hardware or software being used. Most often used audio/video codecs are:

- MP3
- MPEG-2, MPEG-4
- QuickTime (Apple proprietary)
- RealMedia (RealNetworks proprietary)
- Windows Media Video (Microsoft proprietary)

3 Basics of videoconference streaming and archiving

The main goal of this chapter is to give an overview on the basics of videoconference archiving and streaming. Mainly, systems to use with very limited number of videoconference terminals will be covered constructed from easy-to-get and inexpensive building elements.

3.1 Basic archiving and streaming techniques

In sections below, basic archiving techniques will be introduced, by presenting both conventional and more advanced methods of recording.

3.1.1 Conventional recording

The conventional way of recording videoconferences is using a video recorder device, which is really a cost effective and proven technology of the past. Videoconference equipment manufacturers had been suggesting to connect a VHS video recorder directly to a videoconference endpoint for a very long time as they were unable to offer anything else. This conventional method of archiving can work very well with the following supplements:

- Not scalable, recordings needs to be well organized.
- Analogue technology. Every playback will have effect on the quality of recording.
- VHS recording capacity is around 280-300 horizontal lines, which nearly the same as the CIF (352x288) resolution widely used in videoconference. It is very important to note, that 4CIF videoconference is also possible and emerging. Recording a 4CIF videoconference with e.g. a VHS recorder will have notable quality degradation. However, there are much more advanced video standards that can give much better quality for a slightly higher price (e.g. S-VHS).
- In case the recording is intended to be used in digital format later, it needs to be digitized, which is a manual activity and can take a lot of efforts.
- Many videoconference devices have video input connectors, which allow playing back any video recordings to a remote site.

If choosing the conventional way of videoconference recording, it is better to go for a DVD recorder device, which has been available on the market for the last couple of years. A DVD recorder has several advantages over the legacy analogue video technology:

- Digital. A recording can be played back any number of times without causing damage to the media. In addition, it can be copied as many times as required without any quality degradation.
- A DVD recorder can contain an internal harddisk for storing recordings temporary. Such a device provide editing facilities that enables trimming, partitioning a recording, adding a menu to DVD disc, etc.
- A DVD is able to record the full quality PAL or NTSC video image by capturing over 500 rows of a single frame. This recording capacity is perfect for 4CIF videoconference as well.
- A DVD can be easily edited and converted with freely available editing tools. For example, a more compact and distributable copy (i.e. MPEG4) can be created, or a version needed for a specific streaming server can be produced (e.g. Real, Windows Media, etc.).

Before purchasing any video recorder device, it is advisable to check video input/output connectors of the particular videoconference device. Most often, RCA and S-Video connectors are located on videoconference devices.

In addition to the aforementioned conventional recording methods, there are many other solutions with their own advantages and disadvantages (e.g. DV recorder, beta recorder, etc.). That is for sure if conventional recording is used, it is definitely worth to go for a digital solution. However, the problem of storing, organizing and cataloging recordings still remains to be solved.

In case a computer application is used for videoconferencing, it can be very difficult or even almost impossible

to record a videoconference session unless the application itself supports it, which is rather rare. Converting computer VGA signal to video signal can be one possible solution, although not recommended due to bad quality of VGA-video conversion techniques.

3.1.2 Recording and streaming with a computer and a capture card

A conventional recording device can be easily substituted with a computer, a video capture card and appropriate streaming software. Initialization of recording is still done manually and requires more expertise from operating staff than an ordinary recorder. However, a sophisticated setup can be used with the help of a custom made web application controlling the streaming application in the background, see figure.

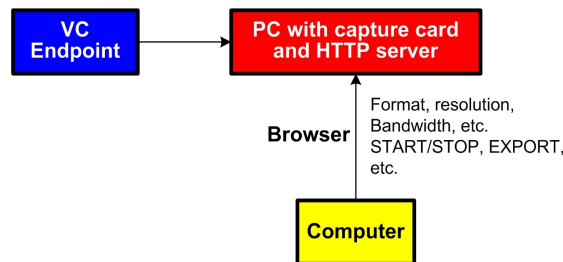


Illustration 1: PC based archiving with HTTP control

The main advantages of a computer based solution are:

- Recording is immediately available in a digital format, in form of a video file. It makes distributing much more easy than in case of DVD recorder based archiving.
- Audio/video can be directly streamed to media players, or pushed to a media server for further distribution. Resulting file can be directly uploaded to a media server for archive provisioning.
- Type of used audio/video codecs can be chosen, thus quality and resulting size is under the control of operator.

The following parameters of recording should be carefully chosen before starting an archiving session:

- Audio and video codecs used, as availability of media player codecs can be an issue.
- Bandwidth depending on the average network access speed of your user community. Multiple bandwidth streams matching more than one network access speed can be delivered as well.
- Audio sampling rate and bandwidth, video resolution and framerate according to the application.

When choosing an adequate video capture card, be aware of the followings:

- Videoconference endpoints usually has RCA or S-Video video output connectors.
- More advanced – thus more expensive – capture cards have realtime MPEG1, MPEG2, etc. encoding capabilities. Most streaming software won't be able to handle these special features, so central CPU will be used instead of the capture card's DSP

Several free quality software tool is available for most of the operating systems for capturing and encoding a video session and some of them even with built-in streaming server capabilities. The list is far from being complete:

- FFmpeg
- mencoder (mplayer encoder)
- VideoLAN
- Windows Media Encoder

3.1.3 Built-in endpoint capabilities

Videoconference endpoints or applications could also have recording facilities, but this feature is extremely rare among current products on market. As it was mentioned earlier, manufacturers advise to achieve this with additional external devices.

In the market segment of purpose built videoconference endpoints only Polycom's iPower series used to support this feature, but it is no more available. iPower is a PC and Windows operating system based endpoint with special hardware and software components with all drawbacks of a personal computer, even so it was capable of saving a videoconference to the local harddisk together with content presented in a conference.

In case archiving is needed, it is strongly suggested to not to build upon these solutions as they often work with proprietary/closed formats (e.g. as content saved and synchronized with audio and video) or very difficult to convert these archived materials to a usable form, which could mean publishing recording on a HTTP server or even converting it to a form that fits to a currently existing video on demand archive.

Streaming facility included in a videoconference endpoint is a more common feature and is available in most of the modern videoconference endpoints. Usually, H.261 and H.263 video codecs are supported with QuickTime or RealPlayer media player clients. Most products only support IP multicast technology as videoconference terminals are not destined to serve as a unicast streaming server as it takes processing capacity according to the number of media players connected.

Even though, IP multicast technology is a perfect medium to use for streaming, one should be aware that it needs special support from network devices and in addition multicast traffic exchange between service providers is not established on the interdomain level. For sake of extending this sort of videoconference streaming, it is advised to relay IP multicast multimedia traffic to unicast with the help of a streaming server. Although, this document does not intend to cover this topic, there are a plenty of documents available on the Internet dealing with this issue. Apples's freely available Darwin Streaming Server (DSS) can be a good streaming server application to do the job.

3.2 Overview of built-in MCU features

Advantages of implementing streaming and recording functionality in a Multipoint Control Unit (MCU) videoconference server is pretty obvious. A single device serving as an MCU for establishing multipoint conferences seems to be the best possible device for distributing separate video streams for media players. In a multipoint conference all the participants are connected to a single MCU device in order to switch, relay, transcode and mix media accordingly. In spite of its obviousness MCUs rarely support direct streaming of videoconferences and even recording due to the following reasons:

- Acting as an unicast streaming server needs resources proportional to the number of clients connected.
- Recording facility needs local storage capacity and additional computing resources.
- Streaming and recording could require many value added services (see section on all-in-one devices), so that it is better to have a single device to implement these together with a separate management interface offering these facilities.

It is important to note that point-to-point conferences only require two videoconference terminals, thus media is not traversing the particular MCU device. In case a point-to-point conference is desired to be streamed by an MCU, it must be set up through the MCU itself.

3.3 Streaming and archiving content

Streaming and archiving content of a videoconference and even retrieving it afterwards can be very difficult and can require sophisticated ideas and custom built environment. There are several ways of transmitting content in parallel to a videoconference (e.g. T.120, VNC, shared desktop, web based collaboration application, etc.), but this document will only cover H.239 as it is a standards-based solution and has been widely available in the last two years. The other standards-based way of sharing content is using the T.120 protocol, which is very barely

used and considered as a deprecated technology. This ideas will also apply to Polycom's People and Content and Tandberg's DuoVideo as H.239 emerged from these two technologies.

Two possible method of extracting content from a videoconference:

- VGA-video conversion: any VGA signal can be converted to an ordinary video signal with notable loss of quality due to major structural differences between these two signals and to downsampling to PAL/NTSC video resolution. A video signal is easier to handle with cheap capture cards and video grabber/recording devices. However, a quality VGA-video converter is extremely expensive. After having a video signal, the following options are there:
 - Capturing still images of graphics. This requires manual intervention when a still image is needed to be stored.
 - Content can be recorded as a video.
 - Content can be mixed to live video with the help of a video mixer. This solution provide perfect synchronization to live video, looks good, but does not allow handling, editing of content after the recorded event.
- VGA capture device: consumer VGA capture devices have been on the market in the last 1-2 years. Before this time period, only professional VGA grabber cards had been available, mainly for scientific applications. A VGA capture device can receive a VGA signal and connects to the PCI, USB or FireWire port of a computer. This method allows grabbing of computer graphic without any loss of image quality and resolution.

A VGA capture device can be used for still image grabbing and high resolution content video recording, similarly to the above described scenarios.

4 All-in-one devices

All-in-one streaming and archiving devices are exactly designed and destined to stream and record videoconferences digitally through IP networks, in an efficient and conveniently manageable way. Videoconference service providers, organizations and larger institutions/companies are targeted by these devices, depending on the capacity and scalability of the particular product. In case of an individual videoconference terminal, this way of archiving/streaming videoconferences can lead to perhaps high and unjustified expenses.

At the time of writing, the following all-in-one archiving/streaming network equipments are available on the market:

- Codian IP VCR 2200
- Polycom ReadRecorder (audio only)
- Polycom RSS 2000 (not yet available, only announced)
- Starbak VCG
- Tandberg Content Server

However, this document does not intend to deal with listing the feature set of each device and nor does it have any preference over any of the above products.

4.1 All-in-one archiving/streaming boxes

Across the following sections VCR, the (VideoConference Recorder) abbreviation, will be used to refer to all-in-one videoconference streaming and archiving devices.

4.1.1 General requirements

In general, a VCR device can be capable of the following facilities and important feature sets:

- Support H.323 and SIP protocols to connect to videoconferences and act as a terminal, thus is able to receive audio, video and content implicitly.
- Support standards-based streaming protocols and codecs as RTSP, HTTP and H.261, H.263, etc.
- Capacity properties can be: max. number of conferences recorded in parallel, max. number of unicast streaming clients served, size of internal harddisk, etc. There can be differently priced licenses according to the capacity.
- Management interface for both administrators and general users with appropriate authentication. Most VCR devices provide a web based management interface.
- Automatic activation of recording to keep archiving easy and convenient to use.
- Facilities for organizing recordings stored locally (i.e. delete, edit properties, organizing in folders, etc.).
- Ability to use remote filesystems like: ftpfs, NFS, Samba, etc. to store/provision (see later) recordings for a long time. In order to connect to a storage device, there can be an option for a FiberChannel adapter for example.
- HTTP/FTP server to let users to download their recordings.
- Archive/stream H.239 content.
- H.323 dial in interface to see list of recordings and to play them back using a videoconference terminal.
- Protection of recordings (password, PIN code, etc.).
- Export audio/video and content in acceptable, post-processable format.

In order to give a clear picture on the implementation of the above functionalities, next sections will describe each main function sets. It is important to know, that these devices haven't been on the market for a long time, thus are or can be immature in some of the above listed areas. The author of this document thinks that there are some specific areas, which need more intention and support from manufacturers to allow seamless integration with other videoconference network equipment and services based on standards and open systems. This document will also raise some questions related to immature implementation of each functionality and will propose ideas to resolve these issues.

4.1.2 User control over streaming/recording initialization

A videoconference archiving and streaming device is designed to serve many videoconference users, it is important to make this service convenient and easy to use without any technical knowledge. To implement such a service, most of the operations on archives should be delegated to users. The most important of these are the way a user can initiate a recording:

- **Web based:** most VCR devices provide a web based management interface for both administrators and restricted users. If this interface is opened for general users, they can have the chance to get control over their recordings by initiating them with the help of a web browser, modifying their properties, deleting them, etc.

Recording initialization is usually implemented on a dial out and record basis, that requires an E.164 destination number or an IP address to be given. The VCR dials out to this particular number and starts recording of received streams immediately. This way of starting a recording session is mostly suitable for MCU conferences.

- **Recording prefix for P2P calls:** E.164 recording prefix based initiation of the archival process at the VCR is the easiest way of starting a recording from the user point of view. The basic idea is to provide a short prefix to be attached before the E.164 number of the destination endpoint if a user would like to have the conference archived. Basically, all the calls matching this special prefix - let's say this prefix is "123" (see below figure) – are routed through the VCR device by the gatekeepers as it would have been a simple gateway. After the call reached the VCR it strips the additional prefix ("123") and dials the destination endpoint in a regular way. In fact, the VCR device maintain two separate calls by on one hand relaying media (audio, video and content) to the other endpoint and on the other hand recording it to the local harddisk or external storage.

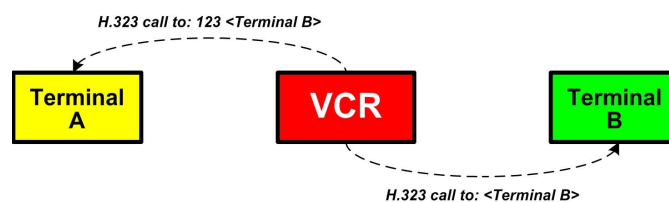


Illustration 2: E.164 P2P recording prefix mechanism

This is a transparent mode of archiving a point-to-point videoconference as users would not even notice any difference to a normal call. However, as both terminals are sending their own video, it is important to define the archived video layout as it wouldn't make sense to record them separately. Thus both video streams can be recorded with the help of a single video layout. Usually, choices for example can include the followings:

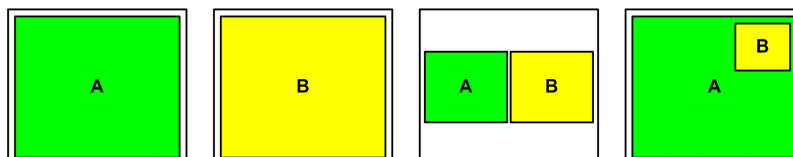


Illustration 3: Example layouts for P2P conference recording

This method also enables one to record an MCU conference by dialing conference dial in number together with the special archival prefix ("123" in the example). Although, in an MCU conference

usually the video stream received from the MCU has importance from the point of archival, so layout should be chosen accordingly.

- **E.164 dial in based:** most VCR devices offer a single E.164 number to call in order to start recording immediately. This is a convenient way of automatic initialization of a recording session. This feature can be used mostly with an MCU, where from the specific conference this E.164 number is dialed out to start archival of the multipoint conference.

In most of the cases videoconference provider entities deploying a VCR device are already having an MCU in order to support multipoint videoconferences. In this particular case, it is obvious to implement a videoconference archiving and streaming service through an existing MCU booking system that is being used for MCU resource reservation. This setup is being used at NIIF/HUNGARNET as well.

A generic layout and components of such a system is depicted below. However, exact implementation details can vary by each MCU product, mainly based on the API it provides for booking resources.

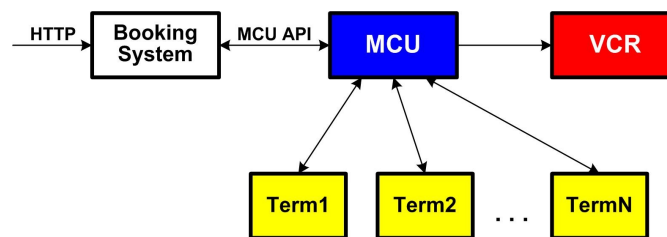


Illustration 4: MCU booking system controlled provisioning with VCR as a predefined participant

In the very first step, a user books MCU resources in advance by using the booking system, which is usually a web based application with a sufficient authentication infrastructure. The web application negotiates on conference parameters and the reservation is made by the MCU device with the VCR as a predefined participant. At the time of conference start the MCU dials out to the VCR device and as well as to the terminals (**Term1**, **Term2**, etc. on the diagram) or it waits for other users to dial in. Thus, the conference will be archived from the very beginning to the end.

Great advantage of this setup, that controlling of VCR resources are tied to authentication through the booking system, so access can be very well restricted, logged and controlled accordingly. One of the major drawbacks of this setup is that the MCU conference will be archived to its end regardless of the effective length of the videoconference, unless the conference is not terminated manually. Thus, very long blank video can be archived in this case, consuming a relatively large storage capacity.

4.1.3 Playback and retrieval possibilities

There are several ways to retrieve an archived videoconference session from the VCR device:

- **Streaming:** streaming of videoconferences are one of the most important features of VCR devices. The VCR should support streaming of live videoconferences without a significant latency (max. a few seconds) and as well as recorded videoconferences.

In order to support a broad range of media players and codecs, VCR should support at least two out of three, most widely used media players and related codecs (e.g. QuickTime, RealPlayer, Windows Media). Support of several control protocols for content accession also have key importance. RTSP and HTTP should be supported as a minimum, and possibly MMS. To best adapt to different networking environments, transmission quality, firewall filtering, etc. the VCR should also support some protocol fallback mechanisms. For example, in case of a firewall, fallback to TCP and port 80 can be very useful to pass the stream to the end user.

Below figure shows a complete VCR setup with a H.323 videoconference that is being archived and with several streaming clients connected using different protocols. A firewall traversal with the help of HTTP protocol is also shown on the picture.

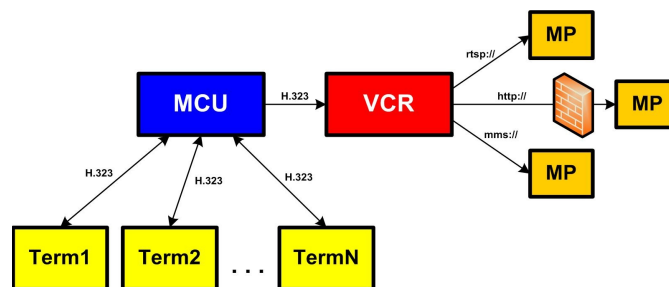


Illustration 5: An MCU conference being archived and streamed by a VCR device

In most of the cases, streaming broadcast is shown in a web page embedded media player on the HTTP interface of the VCR. Apart from this facility, direct URLs to VCR streaming server (e.g. rtsp://, http://, etc.) can be provisioned to users or can be presented on a different web page and embedded media player that connects to the VCR streaming server directly. The latter solution allows any design to be applied to the particular webpage.

- **HTTP/FTP download:** FTP or HTTP download facilities are very important in delivering archive files to the end users before removing them from VCR's harddisk. Implementation of this function can be tricky, as several VCR devices directly save streams received through a H.323/SIP videoconference, so they need conversion before or during download as general users are not able to handle these raw formats. Possible formats can include MPEG1, H.261/H.263 or any other that is supported by a wide range of media players and codecs that are freely available.
- **H.323 dial in interface and IVR control:** an innovative method of provisioning archive recordings by the VCR device can be a H.323 dial in interface controlled through far end camera control (FECC) and Interactive Voice Response (IVR, with DTMF audio signaling). This feature is rarely implemented by manufacturers, but according to experiences it can be extremely useful and convenient way of retrieval as the recordings are directly delivered onto the screen of a videoconference terminal.

The VCR device allows the administrator to configure a specific E.164 number to dial into the VCR and access a special menu where recordings are listed and can be chosen for viewing. The way of controlling the menu can be far end camera control for navigation and DTMF signaling for other inputs as entering PIN code of each recording for example. When the recording is being played back directly to a videoconference terminal, other useful functions can be implemented as pause, rewind, forward, etc. with the help of FECC or IVR. This way the VCR is used exactly as an ordinary video cassette recorder.

Some of the VCR products enables the system administrator to associate a specific E.164 number to each of the recordings. By dialing this number takes the user immediately to the particular recording.

This convenient way of retrieval enables provisioning of archive recordings for a longer period of time. Users just dial a special E.164 number in order to reach their archives and they can play them back exactly as it would be a live videoconference. This service element requires external storage capacity as recordings should be kept for at least months or even a year or so.

- **Web based dial out and play:** in case this is supported and users are enabled to access management interface of the VCR device, they are allowed to dial out a H.323/SIP videoconference terminal and start playing back a recording. This looks very much the same as the H.323 dial in interface except the retrieval is initiated from the VCR device.

4.1.4 H.239 streaming/archiving and playback solutions

H.239 is obviously an emerging technology for sharing computer graphics in a videoconference. Apart from T.120, which is considered as a legacy technology, H.239 is very easy to use as usually terminals offer a VGA input connector to connect for example a laptop or any other VGA output to it. H.239 transmits computer graphics in VGA/SVGA/XGA/SXGA resolution to keep maximum detail for remote viewing and projection. According to H.239's usability and strong support from dominant manufacturers, it is going to be widely deployed and used, hence if purchasing a VCR device it is worth to chose a product that also supports archiving of H.239 content.

H.239 is received by the VCR equipment as it would be an ordinary videoconference terminal except that H.239 stream is saved just like the audio and video streams. The following methods are known in connection with playing back H.239 content using a VCR device:

- **H.323 interface:** when H.323 dial in/out playback interfaces are used, H.239 is played back just like in a normal videoconference, as the VCR delivers the H.239 stream to the particular videoconference terminal.
- **Streaming:** archived H.239 graphics should be also played back when the archive is retrieved as streaming. When content is streamed beside audio and video streams we speak about "rich media", which is considered as a very important technology, but can be implemented in several ways. Streaming products that support this feature often use some proprietary solution to do this (e.g. timing track in video with file/URL references, as it is in QuickTime and Windows Media).

In case of VCRs, either these proprietary media features can be supported or some other sophisticated solutions might be implemented. These include:

- Special Java Applet or JavaScript (or any other) code to follow video timings and play content accordingly.
- Second media player window for playing back H.239.
- **Export/download:** H.239 should be kept with and somehow attached to the exported video file that are available after direct download through HTTP or FTP services. The way of this should rely on open technologies and should allow easy post-processing if some modifications are needed before incorporating the session into a custom video on demand archive that supports rich media.

The best solution is to allow H.239 exporting as individual still images together with timing information relative to the start of the video. Timing information should be made available in text or any other well-known data format (i.e. XML).

H.239 exporting could also work in video format. However, it would make synchronization easier but would also make any processing more difficult and require two media players for playback.

4.1.5 Security considerations

Security issues should be seriously considered when deploying archiving services implemented by a VCR device. As these devices are not mature enough, they often miss some of the most important security facilities to protect recorded sessions substantially. This could require higher level of intention or even frequent manual interventions on behalf of the system administrator. In order to protect each access interface of a VCR device, below security features are implemented or can be implemented by different products:

- **H.323 reachable content:** H.323 interfaces can be protected in several ways. ITU-T recommendation H.235 specifies a security framework for H.3xx series systems. H.235 can provide simple password or digital certificate based authentication method or a combination of this two. Unfortunately, H.235 is rarely implemented in any H.323 products and it seems that this will not likely change in the near future.

An other way of applying protection to a H.323 interface is the usage of IVR passwords, asking for DTMF codes as entry passwords. This is a widely used simple authentication scheme used in call centers, voice mails, etc. Most VCRs implement or should implement this approach.

- **Streaming content:** authentication of streaming interfaces should be compatible with most of the media players that are freely available, so any proprietary solution should be strictly avoided. Unfortunately, an open and well-defined method for authentication of streaming clients does not yet exist.
 - **RTSP/HTTP authentication:** Although, RTSP and HTTP protocols share the very same authentication scheme, it is not implemented in most of the media players and streaming servers. In case this feature is implemented and can ensure certain levels of interoperability with most of the streaming clients, it is advised to use.
 - **PIN code or password:** a PIN code or a pass phrase can be specified for each of the recordings. The easiest way to provide a basic security to include this PIN code or password phrase in the URL (i.e. RTSP, HTTP or MMS URL) of a particular recording. This ties accessing of a particular streaming content to the knowledge of the PIN code or the password. PIN code specified for the H.323 interface could also be reused for protecting streaming content.
- **HTTP/FTP download:** downloading of different recordings is obviously must be protected with appropriate username and password pairs. It is very important to only allow and offer content(s) for download for a single user that was/were recorded by this particular user.

4.1.6 Basic/potential problems of VCR devices

The following list contains all the possible problems of VCR devices that are not yet resolved and can make integration of a VCR rather painful by limiting desirable features and convenient usage or even raise security issues:

- **Auto prefix for dial in and immediate archive:** most of the VCR devices leave an E.164 number to dial in to start a recording session immediately as it was discussed earlier. When making E.164 prefix based archiving part of a videoconference service, security issues should be thoroughly considered. If this special prefix is open to use for anyone, the VCR device can be a subject of a denial of service (DoS) attack by using up its full capacity by creating many parallel video calls to the VCR or by filling VCR's harddisk with long dummy recordings to eat up room for any other valid conference. According to these, this type of service should be carefully protected by access lists or whatever at the gatekeepers or at the VCR device itself.

This also applies to special recording prefix for P2P calls.

- **Auto file name generation problem:** when automated initiation is used, a VCR device generates a filename automatically. This filename usually include date and time in order to avoid collision with other session's name. As it will be detailed in the next section, the automatic session name generation doesn't make it possible to predict archive/live reachability URLs without and API. If the

VCR doesn't have an API interface that enables querying, it is impossible to provide these URLs to users automatically.

If a relatively high number of recording is created in this way, it is also very difficult – even for the system administrator (!) – to find out which recording belongs to a specific conference, endpoint or user. To support this “research”, the VCR should record as much information as possible such as calling E.164 number, H.323 ID of the remote endpoint, etc. or a separate call detail record database will also help if available. From any point of view it is extremely painful both for the users and for system administrators.

- **PIN code issues:** a recording should be protected by PIN code, for preventing unauthorized users to see an archive. In case automated recording initiation is used a random PIN code should be immediately assigned when an archiving session is started. If a PIN is not assigned, the recording will be world readable. In case the VCR is controlled through an API, the API should allow assigning a PIN code when a specific booking is created.
- **Linking a user to each recording (e.g. E.164 prefix):** when a recording is started automatically using a videoconference terminal (E.164 based initialization), a file will be created according to a generated name as detailed above. As a consequence, this recording is impossible to associate with any VCR users as no information was available at the time the recording session was started: someone just dialed in from some videoconference terminal. In this case all the recordings should be seen by a VCR user or if automatic PIN code assignment is used none of the new recordings will be seen by any users. In the latter case, intervention from system administrator is needed as the desired content in PIN code protected.
- **Lack of Application Programming Interface (API):** current VCR products lack any API interface to enable remote controlling of the device, which makes seamless integration to existing videoconference service infrastructure impossible. See next section for details.

4.1.7 Provider scale archiving

This section will give a detailed overview of a possible setup for provider scale videoconference archiving and streaming by describing a service model, necessary components and their interconnection. This section was inspired by efforts being carried out at NIIF/HUNGARNET in order to integrate a VCR device (Codian IP VCR 2200) to the existing networking infrastructure for provisioning transparent videoconference recording and streaming facilities to the research and education community.

The following service scenario description assumes the existence of a basic H.323/SIP videoconference infrastructure: gatekeepers/registrars, MCU and a booking system to allow users to make their reservations without any manual intervention. These assumptions are likely to be given in most of the European National Research and Education Networks (NREN) that runs a national level videoconference service (see figure below). Apart from this statement, below scenario is not carved into stone, any parts can be removed and modified accordingly as your local videoconference service requires.

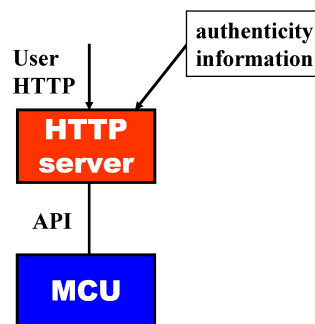


Illustration 6: Current deployments

Basic parameters of provider scale archiving can be measured in the following way:

- Digital archiving through the network.
- Support for both H.323 and SIP videoconference with E.164 based dialing.
- High capacity recording capabilities (many parallel recording sessions are possible, storage capacity for hundreds of hours of recorded video).
- Support of several control protocols and codecs.
- High capacity streaming capabilities (hundreds of unicast streaming clients can be served, multicast support).
- H.239 archiving, streaming and exporting capabilities.
- All operations are automated, not requiring manual intervention.

4.1.7.1 A possible service model

The following list details access points, methods of controlling the videoconference archiving/streaming service from the user point of view (service model):

- **Single point of user access for controlling the service.**

This single point of access in this example is the MCU booking system that is deployed and used by a large user community. Opening of VCR's web interface for general users (if this is possible) is luring, but on a provider scale it is impossible to store and manage authenticity information at two locations.

A user should be able to book MCU conferences using this interface with recording and streaming enabled.

- **Automated provisioning of archive/live reachability.**

In addition to basic MCU conference details, an MCU booking system should present the reachability information of the future H.323 archive (if H.323 reachability is provided), live streaming, archive streaming, and download URLs. It is advised to send notification e-mail messages with these information for further dissemination among conference participants.

In optimal case a user will receive the following: an E.164 number where the archive is reachable afterwards by a H.323 terminal and related PIN code, an URL for live streaming feed (e.g. rtsp://), an URL for archive streaming feed (if differs from the other) and a HTTP/FTP URL to download the recording from.

- **Automated initialization of recording.**

A recording session should start exactly when the MCU conference is initialized according to the conference start time. One way of doing it is specifying the VCR device as a predefined participant in the particular MCU conference. This implies that the MCU will dial out automatically to the VCR device and the recording/streaming is started without any further intervention. However, the MCU booking system should allow to stop and restart the archiving session as a minimum.

- **Other user operations.**

User should be allowed to delete his/her recordings.

4.1.7.2 Problems of implementation

The following diagram shows all of the service building blocks and their interconnection, together with service access points.

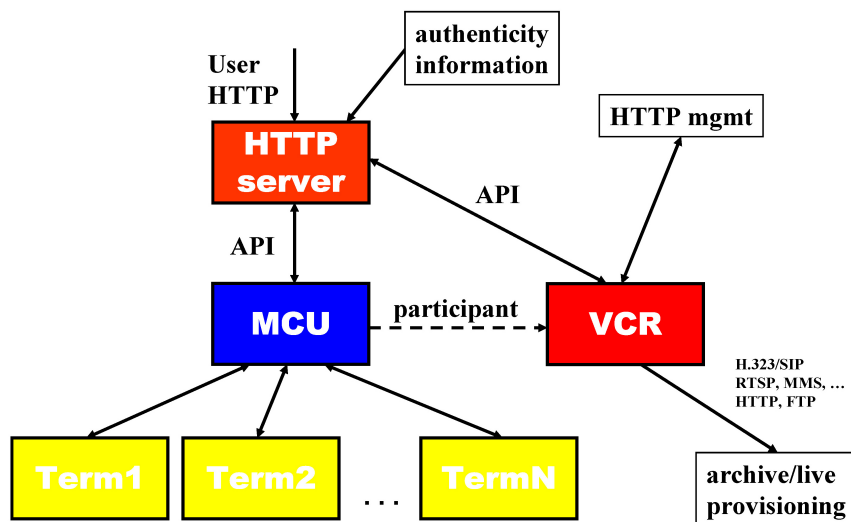


Illustration 7: A possible provider scale videoconference archiving setup

The MCU booking system is hosted on a HTTP server allowing a user to easily create and modify reservation with the help of a web browser. The booking system communicates with the MCU using a special API for creating and managing resource reservation at the MCU device. The VCR is dialed out from the MCU when the conference is initiated, so recording session automatically starts. VCR is managed through the web by the system administrator, user authenticity information is stored at the HTTP server or a backend database connected to it. The diagram also shows possible retrieval options (i.e. H.323, streaming and download). In a perfect world, the booking system communicates with the VCR using the VCR API interface in order to implement association between the booked conference at the MCU and the recording to be made by the VCR. In an optimal case it would include preassigning PIN code to a recording and getting of archive/live reachability from the VCR in order to inform user in advance.

As any VCR products on the market do not include an Application Programming Interface, the connection between the booking system and the VCR device doesn't exist at all. Let's summarize (see previous sections) all

the problems implied by this issue:

- MCU dials out to the VCR, a recording is created and named according to current date and time or anything else that can ensure avoidance of future filename collisions.
- If a PIN code is not assigned automatically to the recording, any user is able to play it back using a simple media player. If for example a random PIN code is assigned, manual intervention from system administrator is needed later.
- As there is no chance for acquiring URL of streaming and HTTP/FTP download, provisioning of this information could also require manual intervention from system administrator. Dissemination of this information should be done by automatic e-mail addresses.
- As usually nothing refers to the origin of a recording, the system administrator needs to manually find the right MCU conference and the appropriate user who made the booking.

4.1.7.3 Proposal for a minimal API functionality

Author of this document contacted manufacturers and inquired about the problems above. It was rather surprising that manufacturers are not aware of these as their biggest market share is in enterprise segment, where these problems are irrelevant or even do not show up. However, a large sized company running a complex videoconference service is very likely to take the same directions introduced in previous sections.

In the following paragraphs, we propose a possible method for providing a minimal API functionality that is needed for integrating a VCR device into a usual environment depicted in this chapter. This method would also allow seamless integration of multi vendor products:

- When the booking is made at the MCU, the booking system turns to the VCR device and ask for the reservation of an archiving session for a given time frame.
- The VCR API returns a specific E.164 dial in number for this reservation, together with a random generated PIN code, a URL for streaming and an other URL for HTTP/FTP download. This information can be populated using automated e-mail messages.
- The E.164 dial in number reserved for an archiving session should be dialed out from the MCU in an automatic manner.
- The API should enable deleting of reserved archiving sessions both before and after the recording is done.
- In case the archiving is stopped and restarted, more than one file could be generated by the VCR. The device should be aware of associating these files with the appropriate recording session and provision URLs and related information according to this (i.e. more URLs is listed for one session, etc.). However, automatic continuation can be visioned once a specific E.164 number is there for a single recording session.

Fulfillment of the aboves would resolve most of the problems outlined in this documentation regarding integration of VCR devices.

5 Conclusions

This document introduced several methods of videoconference archiving and streaming that are available today for different scale of usage from recording videoconferences with the help of a single endpoint to provider scale archiving services. However, not all of the possible solutions and setups were described, the document can be a sufficient starting point in selecting appropriate devices and infrastructure enabling videoconference archival.

Apart from the basic information, capabilities and problems of purpose built videoconference recorder devices were discussed to give a state of the art view of what can be done using this VCR devices. It can also help one to choose a VCR device for purchase depending on current requirements set up by existing network infrastructure and claims of local user communities. The document pointed out that VCR devices are immature from certain points of view, but it is pretty obvious that professional applications will surely require this modern approach. We shouldn't forget that these devices have been brought to the market just recently, so there is definitely a way to go.