

Windows Media Transcoding using CXP Archive Service Data

Fred Videon

*Dept. of Computer Science and Engineering
University of Washington
(fred@cs.washington.edu)*

1. Problem Statement

The ConferenceXP Archive Service supports the capture of conference and presentation data into a SQL Server database, and the playback of a conference to a CXP venue, however many of the users who would derive value from replaying archived conferences use networks or devices which are not capable of running ConferenceXP. These may be clients running lower powered systems, they may be located behind corporate firewalls, on low-bandwidth links, or they may be portable devices which are frequently running disconnected. Further, the Archive Server architecture is not designed to scale to large numbers of concurrent clients playing distinct conferences. To support these usage scenarios, we would like to be able to convert archived conference and presentation data into Windows Media format, which may be streamed in a scalable fashion using Windows Media Server, through firewalls, or on low bandwidth networks, or which may be downloaded to mobile devices for off-line viewing.

The ConferenceXP to Windows Media Gateway¹ transcodes a live CXP conference to Windows Media format in real-time. While it is currently possible to play back a conference to a CXP venue from the Archive Server while archiving with Windows Media Gateway, this process cannot go faster than real-time. In contrast, the speed of a transcoding process with direct access to the Archive Server database would be limited by the CPU of the encoding system.

An additional function of the Windows Media Gateway is to archive presentation data including slide navigation and ink from Classroom Presenter². On the client systems, the CXP WebViewer³ is used to play back live or archived Windows Media streams which contain this presentation data. We would like the Windows Media archives produced from Archive Server data to integrate

¹ <http://www.cs.washington.edu/education/dl/confxp/wmgateway.html>

² <http://www.cs.washington.edu/education/dl/presenter/>

³ <http://www.cs.washington.edu/education/dl/confxp/webviewer.html>

well with these applications. Presentation data resident on Archive Server should be convertible to a format which is compatible with WebViewer, and is synchronized with the Windows Media file.

Currently the process of creating a Windows Media archive with presentation and ink with Windows Media Gateway is a multi-step process which involves several systems, several files and some manual post-production work, including a process of trial and error to synchronize the presentation stream. By producing the Windows Media archive from conference and presentation data resident in an Archive Server database, and by designing the appropriate UI and logic, this process can be simplified and streamlined.

2. Expected Outcomes

The tangible output of the project will a software application (hereafter referred to as 'the application') freely available for download from a public web page in binary and source form. The application will be written primarily in C#, and use the .Net Framework.

2.1 Audio/Video Transcoding

The core function of the application will be to convert a set of audio and video streams stored in an Archive server database into a Windows Media stream with optional presentation data. The details of the processing will differ depending upon the streams and options chosen by the user.

In most circumstances it will be necessary to decompress existing AV streams prior to encoding to the target Windows Media profile, however if the user chooses to retain the native compression and bandwidth used by the CXP streams, it will be possible to perform the transformation to Windows Media format without recompression. Decoding techniques including those using DirectShow and using the native Windows Media Format SDK will be evaluated for efficiency and range of functionality. If multiple audio streams are selected by the user, they will be mixed into a single audio stream.

The Windows Media Format SDK will be used to perform the encoding. The SDK's "two-pass encoding" feature will be evaluated to see if it can be used to produce a higher quality output with CXP video sources, and if so, will be included as a user option.

2.2 Presentation Data Integration

For the initial version of the application, Classroom Presenter will be used to generate the presentation data streams. The application will perform a transformation on the presentation data stream to optimize it for the archive playback scenario. The native data stream produced by

Classroom Presenter is designed to accommodate late-joining clients in the real-time scenario. For example, Classroom Presenter produces a set of heartbeat messages which fully define the current presentation state. In the archive playback scenario, the entire presentation dataset is available to the client as soon as the media is opened, and the client application may at any time seek forwards or backwards in this data in response to user navigation, therefore the data needs only contain navigation transition events and ink. Much of the heartbeat information is redundant, and can be filtered out. The result of the transformation will be a much smaller dataset formatted as XML, and compatible with CXP WebViewer.

2.3 User Options / UI

The Application UI will permit the user to query the Archive Server database to find available conferences. After selecting a conference, the user will be allowed to choose from the streams in that conference. It will be possible to choose multiple audio streams which will be mixed together. In the initial design, the output Windows Media file will contain only one video stream; however it will be possible to change between video sources at user-specified times in the conference. Likewise, only one concurrent presentation stream will be allowed, but it will be possible to change the presentation source in the middle of an encoding session.

Since in practice Presentation data often uses a CXP Venue which is distinct from the conference audio/video venue, the user will be permitted to choose from any available presentation which is concurrent with the AV conference.

The user will be allowed to specify time points in the conference at which the encoding should begin and end. It will also be possible to specify multiple time spans to be merged into the output. For instance if a class took a 10 minute break in the middle, it would be desirable to remove the break from the encoded output.

The user will choose Windows Media codecs and other encoding parameters in the form of a Windows Media Profile. Profiles will be fully customizable with support from the Windows Media tools.

Data formatting may vary slightly depending upon whether the target clients are on-line, streaming from a Windows Media Server, or are downloading for off-line use. The application will support formatting and packaging for both scenarios.

When all parameters are set, an encode button will initiate the job.

The application will support a command-line mode for batch processing. The command-line mode will accept transcoding parameters as command-line arguments, or in the form of a XML batch file.

2.4 Missing Data and Packet Loss

Since conference archiving depends on real-world networks and systems, it is inevitable that data which would be used by the application in an encoding process will occasionally be unavailable in the Archive Service database. Common causes of missing data are CXP nodes which leave the venue temporarily, short term network outages and network packet loss. In these cases it will be necessary for the application to work-around the problems as well as possible. There may be some scenarios where the application should skip past a bad section. In other cases it may be preferable for the application to insert null data (copied frames or silent audio) to smooth over the missing data. Experiments will be done with real-world conference data to identify the typical problems and the most appropriate solutions will be implemented.

2.5 Further Enhancements

There are a variety of additional features which will be considered based on available time and perceived demand.

Audio/Video Preview. Having the ability to play back sets of streams from a conference would be helpful to the user in locating start and end times. The user should be permitted to choose a time point anywhere in the conference at which to begin the preview.

Support for other presentation formats. Some users may not use Classroom Presenter, but may use the Presentation tool built-in to CXP. The application may be able to convert archived data from either RTDocuments or from Classroom Presenter format into WebViewer archive format.

Enhanced query options. When selecting streams for an encoding job, it may be an aid to the user to be able to query the archive database in other ways, for example by participant, by date/time, or by conference title keywords.

Audio enhancements. An audio level adjustment feature would be useful in the case of a conference conducted with source audio gains set at disparate levels. The resulting archive would be improved by allowing the user to manually boost or attenuate levels of individual audio streams. Audio limiting, noise canceling or other filters could also be implemented.