



Contribution ID: 86

Type: poster

## Automatic processing of CERN video, audio and photo archives

*Wednesday 5 September 2007 08:00 (20 minutes)*

The digitalization of CERN audio-visual archives, a major task currently in progress, will generate over 40 TB of video, audio and photo files. Storing these files is one issue, but a far more important challenge is to provide long-time coherence of the archive and to make these files available on line with minimum manpower investment.

An infrastructure, based on standard CERN services, has been implemented whereby master files, stored in the CERN Distributed File System (DFS), are discovered and scheduled for encoding into lightweight web formats based on predefined profiles. Changes in master files, conversion profiles or in the metadata database (read from CDS, the CERN Document Server) are automatically detected and the media re-encoded whenever necessary. The encoding processes are run on virtual servers provided on-demand by the CERN Server Self Service Center, so that new servers can be easily configured to adapt to higher load. Finally, the generated files are made available from the CERN standard web servers with streaming implemented using Windows Media Services.

This paper describes the architecture in detail and analyses its advantages and limitations.

**Primary author:** KWIATEK, Michal (CERN)

**Presenter:** KWIATEK, Michal (CERN)

**Session Classification:** Poster 2

**Track Classification:** Collaborative tools