A New Jersey Statewide Video Portal Based on Open Source Technologies

Isaiah Beard
Digital Standards & Workflow Manager - SCC
Repository Architects: Ron C. Jantz and Grace Agnew
Rutgers, the State University of New Jersey
Outline & Overview

- Concepts and History
- Architecture and Usage Scenarios
- Challenges unique to NJVid
- Demos/Proof of concept
Beginnings

Concepts & History
Predecessor Technologies

- New Jersey Digital Highway (njdigitalhighway.org)
  - Statewide repository – IMLS grant
  - Established Foundation Platform
- RUcore – Rutgers Community Repository
  - Expansion of NJDH technology
  - Customization, portals
- Born Digital Archiving
Predecessor Technologies

- New Jersey Digital Highway (njdigitalhighway.org)
  - Statewide repository – IMLS grant
  - Established Foundation Platform
- RUcore – Rutgers Community Repository
  - Expansion of NJDH technology
  - Customization, portals
  - Born Digital Archiving
Predecessor Technologies

- **Existing Video Collections**
  - Partner Institutions with video content
    - **New Jersey Digital Highway**
    - **Local Collections – WPUNJ, Montclair, etc.**
Three Major Video Collection Types

- Video Commons
  - Locally Owned/Locally produced
- Commercial Videos – Consortia
- Lectures on Demand
Nuts & Bolts
Architecture & Workflow

Tuesday, April 15, 2008
Reference Architecture

- A Trusted Repository
  - Reliable content delivery
- Standards-based digitization
- Preservation Platform
- Rights Enforcement
- Authentication/Authorization
A Trusted Repository

- Reliable content delivery
- Standards-based digitization
- Preservation Platform

Rights Enforcement

- Authentication/Authorization
### Reference Architecture: Technologies in Use

<table>
<thead>
<tr>
<th>Area</th>
<th>Software Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applications/Web Interface</td>
<td>PHP / MySQL</td>
</tr>
<tr>
<td>Journals</td>
<td>Open Journal System (PKP)</td>
</tr>
<tr>
<td>Search Engine</td>
<td>Amberfish</td>
</tr>
<tr>
<td>Workflow Management</td>
<td>PHP / MySQL (developed in-house)</td>
</tr>
<tr>
<td>Collection Services</td>
<td>PHP / MySQL (developed in-house)</td>
</tr>
<tr>
<td>Handles/Persistent IDs</td>
<td>CNRI Handle Server</td>
</tr>
<tr>
<td>Video Streaming Server</td>
<td>Darwin (Apple)</td>
</tr>
<tr>
<td>Digital Library Framework</td>
<td>FEDORA</td>
</tr>
<tr>
<td>OS/Web Server</td>
<td>Linux/Apache</td>
</tr>
</tbody>
</table>
Reference Architecture

- User Input (Metadata)
- Workflow Management System
- FEDORA Repository Service
- NJVid Commons
- Partner Portals
- Collection/Preservation Services
- Digitized Video Master
- Digital Object Repository (FEDORA)

- Matrox Axio Encoder
- Video

Tuesday, April 15, 2008
Reference Architecture

User Input (Metadata)

Workflow Management System

Matrox Axio Encoder

Video

NJVid Commons

Partner Portals

Collection/Preservation Services

FEDORA Repository Service

Digitized Video Master

<XML>

Digital Object Repository (FEDORA)
Architectural Challenges

- Large Master Video Files
  - Master DV/AVI files @ 20GB+ per hour
    - Ingest, presentation, preservation
- Lectures on Demand – Annotate Video
- Authentication/Authorization Facility
## Architectural Challenges

### Object Architecture

#### Repository ID

<table>
<thead>
<tr>
<th>Metadata</th>
<th>Behaviors (disseminators)</th>
</tr>
</thead>
</table>

#### Data Streams

- SMAP1 - StrMap (TOC)
- FLV1 - Download
- MOV1 - Quicktime Stream
- PDF1 - Transcript
- XACML Policy
- ARCH1 - Archival Master
Architectural Challenges

Object Architecture

Repository ID
- Metadata
- Behaviors (disseminators)

Data Streams
- SMAP1 - StrMap (TOC)
- FLV1 - Download
- MOV1 - Quicktime Stream
- PDF1 - Transcript
- XACML Policy
- ARCH1 - Archival Master

MD Structure
- Descriptive
- Technical
- Source
- Rights
- Digital Provenance

Tuesday, April 15, 2008
Reference Architecture

- Demo
  - Workflow Management System (WMS)
Large Files:
The Digital Video Object

Archival Masters

- **Film transfer/VHS/Betacam/Analog Sources**
  - Uncompressed AVI as Master
- **Digital (DVCAM, miniDV, HDV, etc.)**
  - DV file with native format/encoding
- **DVDs**
  - Disk Image (ISO) kept as archive Masters
Large Files: The Digital Video Object

Presentation Streams

- **Streaming Protocol:** MPEG-4 H.264 (Quicktime)
  - 15–20 frames/sec, 640x480
  - Data rate: 384–512 kbps
- **Progressive Download**
  - Flash Video (FLV with SWF wrapper)
  - 15 fps, 640 x 480 – 256–384 kbps
- Intended as an alternate presentation mode

Tuesday, April 15, 2008
Large Files: The Digital Video Object

Presentation Streams

- **Streaming Protocol:** MPEG-4 H.264 (Quicktime)
  - 15–20 frames/sec, 640x480
  - Data rate: 384–512 kbps

- **Progressive Download**
  - Flash Video (FLV with SWF wrapper)
  - 15 fps, 640 x 480 – 256–384 kbps
  - Intended as an alternate presentation mode
Large Files: The Digital Video Object

Presentation Streams

- **Streaming Protocol:** MPEG-4 H.264 (Quicktime)
  - 15–20 frames/sec, 640x480
  - Data rate: 384–512 kbps

- **Progressive Download**
  - Flash Video (FLV with SWF wrapper)
  - 15 fps, 640 x 480 – 256–384 kbps
  - Intended as an alternate presentation mode
Large Files: The Digital Video Object

Presentation Streams

- Codec Selection Criteria
  - Easy to Use/Install
  - Supports Multiple Platforms
  - Built on established industry standards
  - Excellent quality/bandwidth ratio
  - Widely Used (Akamai, YouTube, etc)
Large Files:
The Digital Video Object

Content Demo

- Demo Video Clip from Existing Collections
Large Files: The Digital Video Object

Presentation Streams

- Open Source-able architecture – Darwin Streaming Server (DSS)
  - Can run on all major OS platforms
  - Free, open source license (APSL)
Large Files: The Digital Video Object

Presentation Streams

- Open Source–able architecture – Darwin Streaming Server (DSS)
- Can run on all major OS platforms
- Free, open source license (APSL)

Tuesday, April 15, 2008
Large Files: The Digital Video Object

Presentation Streams

- **Transcripts**
  - When available
  - Searchable & Indexable
    - Text/XML
  - Will require labor
    - (transcribing & data entry)
Application Architecture

Partner/Custom Portals

Middleware
(Host Server/Shibboleth)

Digital Object Repository (FEDORA)

Backend Storage Subsystem

Archival Master Files

- External Storage
- Checksum Integrity Validation

Darwin Streaming Server
(Video Content Delivery)
Application Architecture

**Partner/Custom Portals**

**Middleware**  
(Host Server/Shibboleth)

**Digital Object Repository (FEDORA)**

**Progressive Download**  
(Flash Video)

**Retrieve Streams in FEDORA via NFS**

**Darwin Streaming Server**  
(Video Content Delivery)

**Backend Storage Subsystem**

**Archival Master Files**

- External Storage
- Checksum Integrity Validation

**Quicktime/H.264 Stream**
Usage Scenarios
The Front-End Interface
End User Scenarios

- Faculty Annotation for Lectures on Demand
  - Specific clips, segments for dissemination
- General Public Access – NJVid Commons
- Student Access to Licensed Videos
  - A/A required for selective access
End User Scenarios

- Faculty Annotation for Lectures on Demand
  - Specific clips, segments for dissemination
- General Public Access – NJVid Commons
- Student Access to Licensed Videos
  - A/A required for selective access
Lectures on Demand: Annotated Video Clip

- Workflow
  - Instructor previews video and selects segments for students
Lectures on Demand: Annotated Video Clip

- **Workflow**
  - Instructor previews video and selects segments for students
Lectures on Demand: Annotated Video Clip

- **Workflow**
  - Instructor previews video and selects segments for students
  - Instructor authenticates, and creates annotation objects for the video, specifying in and out points
Lectures on Demand: Annotated Video Clip

- **Workflow**
  - Instructor previews video and selects segments for students
  - Instructor authenticates, and creates annotation objects for the video, specifying in and out points

![Image of Make Multiclip window]
Workflow

A structure map with appropriate time segments and persistent ID of the source video is created. All pieces fall into place.
Workflow

- A structure map with appropriate time segments and persistent ID of the source video is created. All pieces fall into place.
Development Needs for NJVID

- Finalization of Video Architecture (Summer 2008)
- Implement Shibboleth for Authentication and Authorization (Fall 2008)
- Fine-Grained object access via XACML (Fall 2008)
- Video Table of Contents (early to mid 2009)
- Multiple Structure Map Annotations for Complex Segments (mid 2009)
- NJVID Custom Installation Configuration
Isaiah Beard
Digital Standards & Workflow Manager
Scholarly Communication Center
Rutgers University Libraries
isaiah.beard@rutgers.edu
Watch This Space!

View our blog, sample videos and progress reports at:

www.njvid.org

William Paterson University

Institute of Museum and Library Services
NJEDGE.Net

Internet2 Spring Member Meeting
Arlington, VA - April 23, 2008