Open Standards Driven Cloud Storage and Preservation

Brian Campanotti
Global Director of Business Development
Oracle Digital Media Solutions
Asset Storage and Preservation Challenges

• The scale of the data can be immense (HD, 4k, J2k, etc.)

• Many different formats used in many ways in many different workflows

• Not another silo! End-to-end integration and content accessibility are key

• Storage technology choices, migration and scalability
Asset Storage and Preservation Challenges

• Redundancy and content protection – we only want to do this once!

• Complex assets should be stored as objects not as files

• How can I integrate into my existing systems, workflows, etc?

• Will I be able to “access” my assets decades from now?
The Complex Challenges We Face
The Real Deal

• Pick the right “encoding” technologies

• Be realistic about technology and storage migrations

• Understand how your valuable assets are actually stored!
The Reality of Storage Technology Migration

10TB in 1998

- Powderhorn with 6,250 slots
- TimberLine 9490EE drives (1.6GB/tape)
- 357 sqft
- 8,200 lbs
- Current media cost of approximately $32,000 (Ebay)
- Approximately 260 hours to write 10TB of data!

10TB in 2014

- A bit more than one single T10000D data tape!
- T10000D technology (8.5TB/tape)
- 0.3 sqft
- 1.2 lbs
- Current media cost of approximately $280
- Less than 10 hours to write 10TB of data!
How is my Data Actually Stored?

- Proprietary storage systems and formats are rampant
  - Silos, data formats, interface protocols
  - High risk of orphaned archives
  - Open source can present unique challenges

- Key user and industry requirements for the “ideal” storage and preservation solution
  - No vendor/technology lock-in
  - Limitless scale (capacity, spanning, etc.)
  - Preservation features (provenance, fixity, etc.)
  - Encapsulation and object-based
  - Storage technology agnostic – not just data tape!
  - Metadata encapsulation
  - Asset updating and versioning support
  - Asset transport (media, streaming and file) capabilities

- What choices exist today?
## Asset Storage and Preservation

<table>
<thead>
<tr>
<th>Key Features</th>
<th>AXF</th>
<th>LTFS</th>
<th>TAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT-centric design applicable to M&amp;E as well as enterprise</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Applicable to cloud-based transfers and storage</td>
<td>✔</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encapsulates (wraps) files and metadata to protect key asset relationships</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Scales to store any number of files of any size and of any type</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Key support for spanning across media (data tape spanning, etc.)</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Universal format regardless of storage technology (disk, flash, data tape, optical, etc.)</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Includes key preservation qualities (provenance, fixity, access control, etc.)</td>
<td>✔</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-Describing Media: Maintains on-media index of all stored objects</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Self-Describing Objects: Maintains in-object index of all stored files</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Each object and file is individually indexed for enhanced recoverability</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Supports any generation of data storage technology (LTO, Oracle T10000X, ODA, Flash, etc.)</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Stand alone applications/drivers allowing for cross-system and cross-platform access</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>File and Object versioning support</td>
<td>✔</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SMPTE AXF Standard

• Society of Motion Picture and Television Engineers recently published standard SMPTE ST 2034-1:2014

• The only open standard for transport, storage and preservation of all file-based assets and collections

• Includes all the functionality of LTFS and TAR while overcoming their limitations and adding significant transport, security and preservation features

• Check out OpenAXF.org for more info
What about Storage and Preservation in the Cloud?

- Philosophical challenges
- Asset security concerns
- Physical data location
- Economic risk
- Long-term accessibility
- Immature and failed solutions
- Proprietary implementation risk
- On and off-boarding challenges
- Deep device integration
- “Simple” storage is simply not enough
CSM: Not just Simple Storage

• Content Storage Management (CSM) is a content-aware, object-based storage abstraction and preservation solution

• Designed for large file and big-data collection management

• Intelligently manages multiple storage tiers based on content, format, etc.

• CSM is an *Active Archive* solution and is *not* HSM

• Deep integration with human workflows, business systems and devices

• Content awareness, deep asset analytics and metadata mining capabilities

• CSM solutions are available on-prem and in private-cloud topologies
CSMaaS Private Cloud Storage and Preservation

- Highly scalable PRIVATE-cloud based CSM as a Service (CSMaaS)
- Not just “simple storage” – built for big-data archiving and long-term preservation
- Geo-predictable locations with physical media “on/off-boarding”
- Highly secure – encrypted links, physically partitioned user space, encrypted storage, offsite duplication, etc.
- Based on high performance data tape which offers increased safety, portability, asset protection and durability
- Open APIs and deep third party integration
- Price aligned with public cloud solutions while offering much higher functionality and increased SLAs
- Archive eXchange Format (AXF) open standard for long-term storage, preservation and accessibility
CSMaaS Active Archive
CSMaaS Simple Web-Based Accessibility
Your CSMaaS Cloud Services Checklist

- Private cloud service
- Content Storage Management (CSM) capabilities
- Redundant copies of all assets
- Physically partitioned user-space
- Automatic storage technology migration
- SMPTE Archive eXchange Format (AXF) open standard based
- Automatic generation of frame accurate proxies
- Web-based MAM front-end for browsing, metadata, etc.
- Direct Integration with leading third party systems

- First byte of data access in seconds
- Content aware features (timecode based restore, etc.)
- Advanced content aware lifecycle policy engine
- Federated view of assets stored on-prem and the cloud
- Data transport and data-at-rest content validation
- Key preservation characteristics (provenance, fixity, etc.)
- Open APIs for third party control
- On and off-boarding by bulk media shipments
- Copies of assets maintained both online and offline
Trusted Solutions for the Worlds Leading Content Owners

• Media and Entertainment
  NBC Universal, Disney, BBC, Sony, Televisa, etc.

• Service Providers
  Encompass, DAMsmart, Omnilab, etc.

• Public Sector
  CBC, LOC, VOA, European Parliament, FEMA, etc.

• Sports Organizations
  MLB, NFL, OLYMPICS, FIFA, NHL, NASCAR, etc.

• Military Agencies
  NGA, CIA, AFRTS, Fort Bragg, etc.

• Non-Governmental
  United Nations, etc.

• Museums and Galleries
  USC Shoah, Smithsonian, Newseum, etc.

• Educational
  NYU, USC, Harvard, University of Virginia, etc.

• National Archives
  Dutch Archives, Danish Archives, LAC, etc.

• Gaming
  Sony PS4 and PlayStation Network