Open Standards Approach for Video and Film Archiving and Preservation

SMPTE Florida

November 19, 2014

Brian Campanotti
CTO
brian.campanotti@fpdigital.com
@BrianCTO

FRONT PORCH DIGITAL

ORACLE | FRONT PORCH DIGITAL

Presentation Topics

- The Importance of Content Storage Management (CSM)
- Preservation Challenges
- Modern Storage Technology Profiles
- The Reality of Storage Migration
- Archive eXchange Format (AXF)
- Clearing up the Confusion – AXF, TAR and LTFS
- AXF Deep Dive, Certified Technologies and Use Cases
- The SMPTE AXF Standard

Brian Campanotti
CTO
Front Porch Digital
@BrianCTO

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

Content Storage Management (CSM)

Content Storage Management
On-Premises or Private Cloud

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

Media Storage and Preservation Challenges

- The scale of the data is immense – HD, 4k, J2k, etc.
- Many different formats used in many ways in many different workflows (archivists, news, post, etc.)
- End-to-end integration and content accessibility are key – Do not build ANOTHER silo!
- Storage technology choices, migration and scalability

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

Media Storage and Preservation Challenges

- Redundancy and content protection - no assets can be lost and we only want to do this once!
- Complex media assets must be stored as “Objects” not as “Files”
- How can I integrate into my existing systems and workflows?
- Will I be able to “access” my assets decades from now?

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

The Complex Challenges We Face

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

The Real Deal

- Pick the “right” encoding technologies including codecs and wrappers
- Be mindful of storage generational and technology migrations
- The most important factor to consider is how your data is actually stored!

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

Data Storage Technology Landscape

	Oracle T10000D	IBM TS1140	IBM/HP LTO-6	Sony ODA
Native Capacity	8.5 TB	4.0 TB	2.5 TB	1.2TB
Native Sustained Data Rate	252 MBps	250 MBps	160 MBps	35/100 MBps (W/R)
End-to-End Data Protection	Yes	Yes	No	Yes
Buffer Size	2 GB	1 GB	512 MB	None
Connectivity Options	16Gbps FC, FICON 10Gbps FCoE	8 Gbps FC, FICON	8 Gbps FC, 6 Gbps SAS	8 Gbps FC
Number of R/W Channels	32	32	16	4

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

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
- Approximately 10 hours to write 10TB of data

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

Storage Technology Migration at Scale

10PB in 1998

x 10

- 1,000 Powderhorn Libraries
- Approximately 6,000,000 slots/tapes
- Close to 8 acres
- Approximately 4,100 tons

10PB in 2014

- One SL3000 Library
- Approximately 1,177 slots/tapes
- 37 sqft
- Approximately 1.5 tons

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

How is my Data Actually Stored?

- Proprietary storage systems and formats are rampant in the industry
  - Silos, data formats, interface protocols
  - High risk of orphaned archives
- Key user and industry requirements for the “ideal” storage and preservation solution
  - No vendor/technology lock-in
  - Limitless scale
  - Support for Media Spanning
  - Encapsulation support for “Complex” Objects
  - Highly resilient format
  - Long-term accessibility – preservation!
  - 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?

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

Tape ARchive Format (TAR)

- Tape ARchive (TAR) format has been around for many decades
- No true universal TAR implementation but many “customized” ones
- Cases exist where TAR packages cannot be recovered due to home-grown code nuances
- Some archive systems do leverage TAR variants but use is very limited today
- TAR does not support many of the identified user requirements

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

Linear Tape File System (LTFS)

- LTFS is a file system for data tape which makes tapes appear as “removable storage”
- LTFS is targeted at the simple transport and exchange of files between systems
- LTFS does not meet many of the user requirements outlined previously
  - Relies on simple folder hierarchies to maintain asset relationships
  - Application specific in terms of implementation
  - No metadata encapsulation
  - No support for spanning (limited scalability)
  - No preservation features
  - Only applicable to data tape
- LTFS is NOT appropriate for long-term storage and preservation

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

Archive eXchange Format (AXF)

- AXF is a universal standard for the wrapping (encapsulation), storage, transport and preservation of any type of file assets
- AXF is like an advanced ZIP which encapsulates any number of files of any type, metadata and a universal file system – an “object store”
- AXF is IT-centric and applies to all types and generations of storage technologies
- The first open standard targeting the storage, preservation and transport of file assets
- AXF fully defines (and constrains) implementations for increased interoperability
- AXF includes all of the capabilities of LTFS, overcomes its well-known limitations and adds key long-term accessibility, storage and preservation features

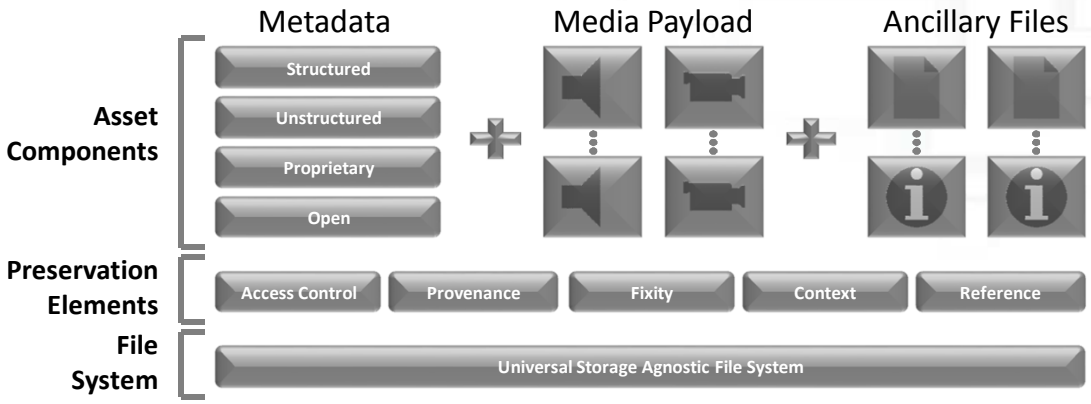
© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

AXF and LTFS Key Feature Comparison

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

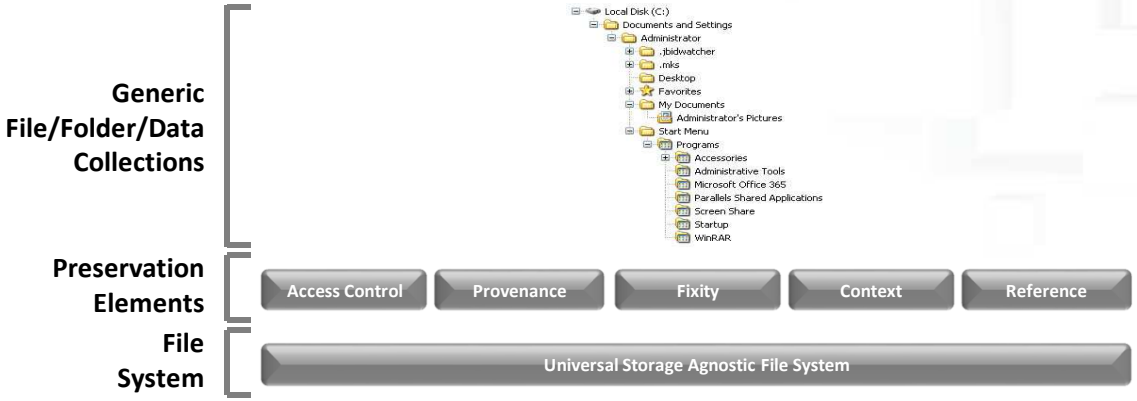
© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

AXF in Media and Entertainment



© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

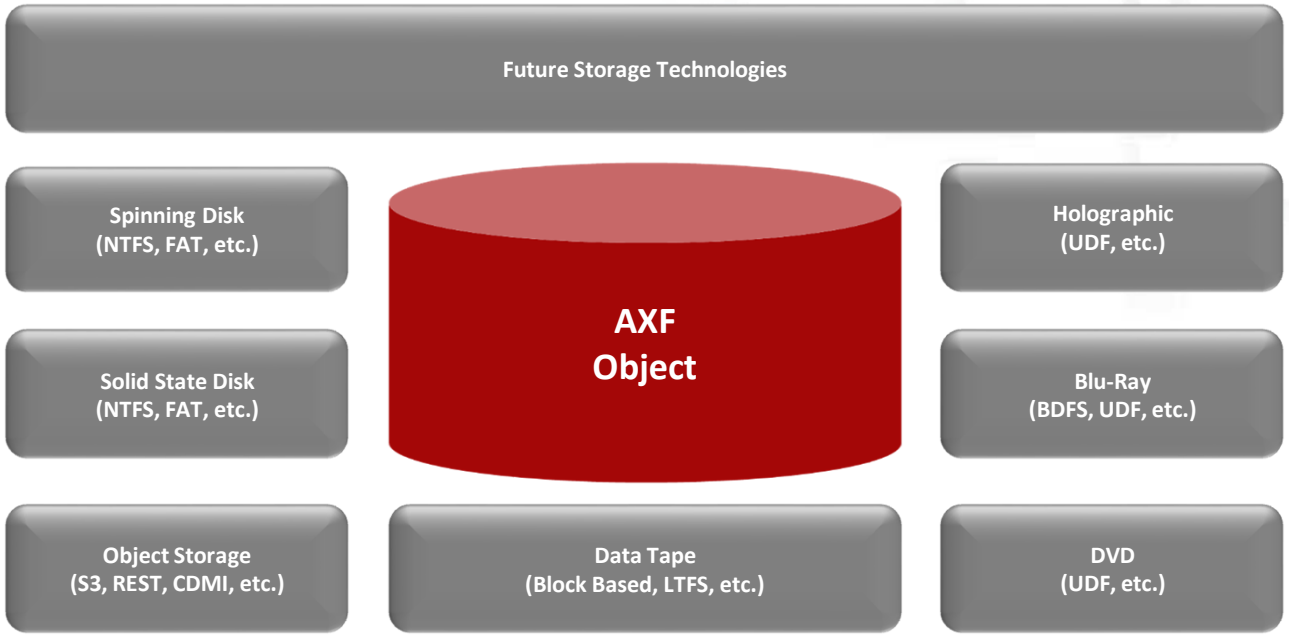
AXF in Big Data Archiving



© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

AXF
Archive eXchange Format

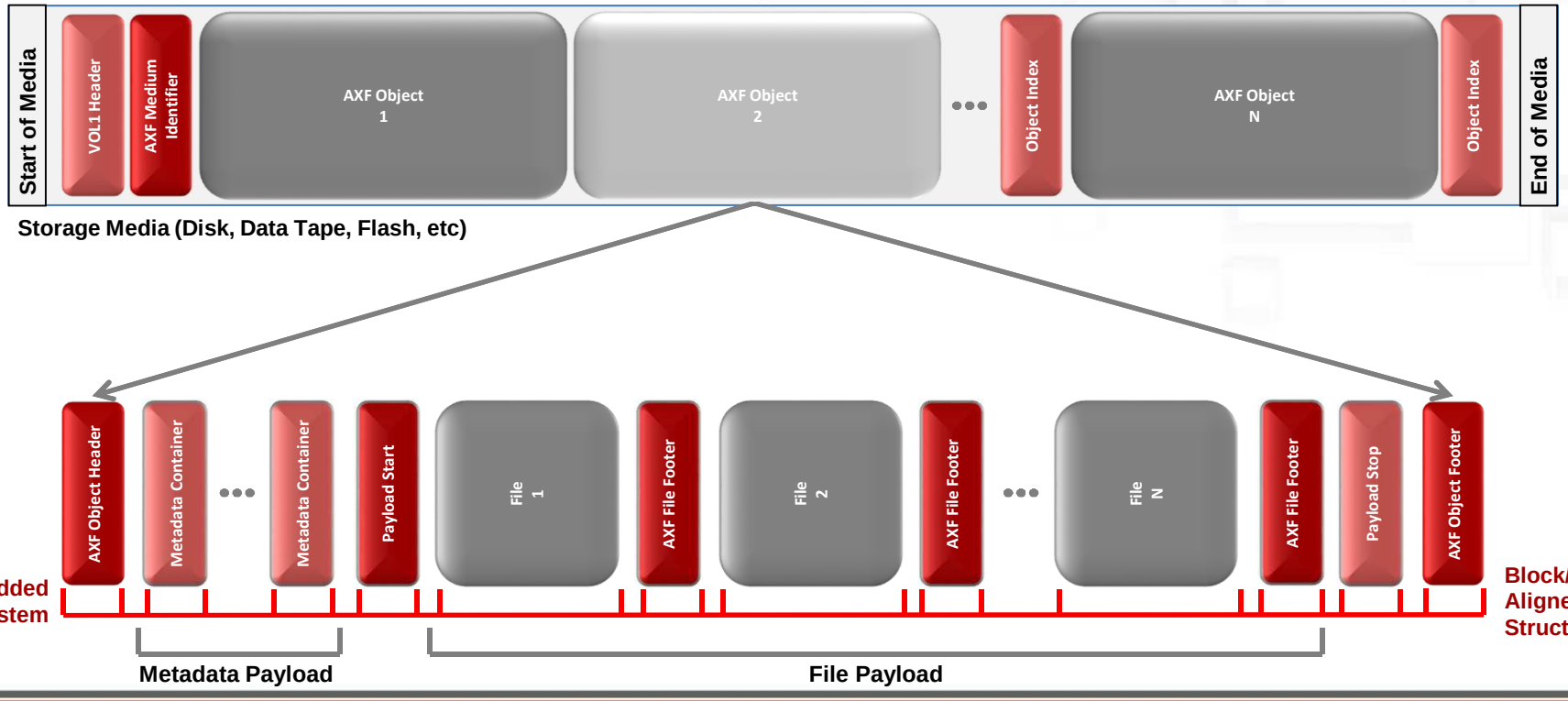
AXF is a Universal Format



© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

AXF
Archive eXchange Format

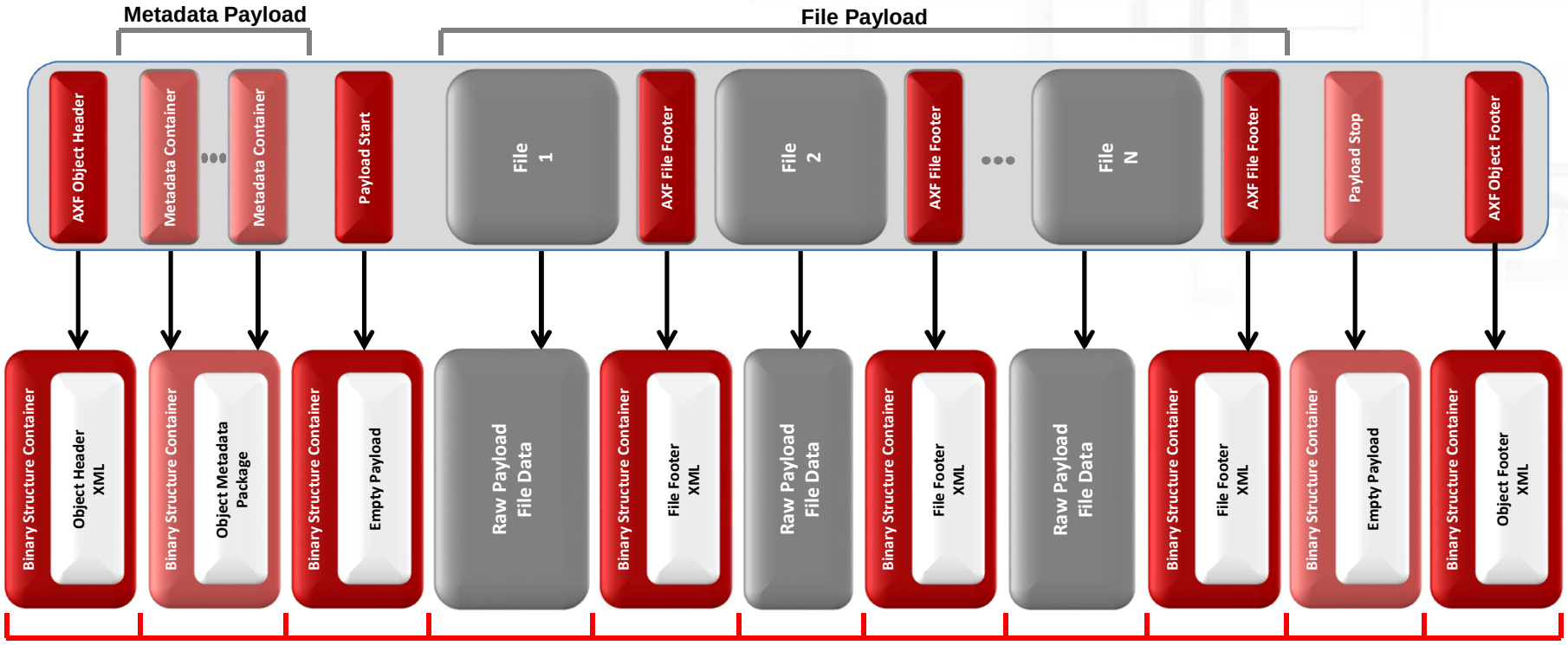
AXF Structural Overview



© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

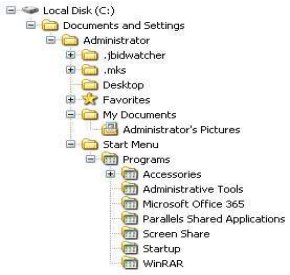
AXF
Archive eXchange Format

Inside an AXF Object



© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

On-the-Fly AXF Object Construction



File Collection

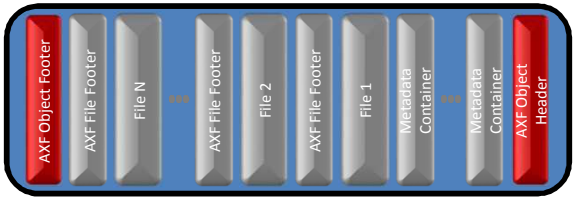
+

Metadata

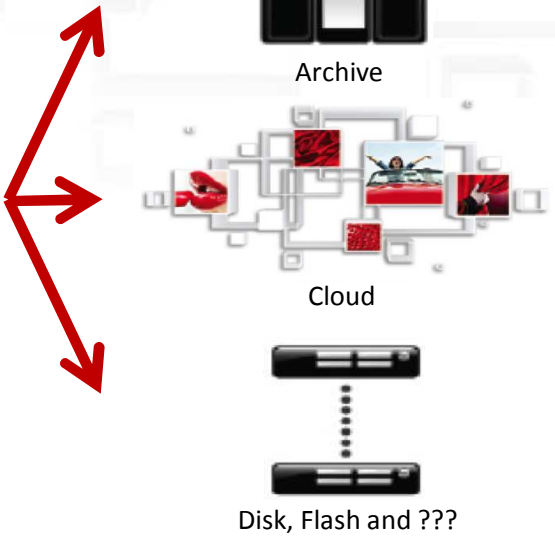
+

Provenance Information

AXF Aware Application



AXF Object



Archive

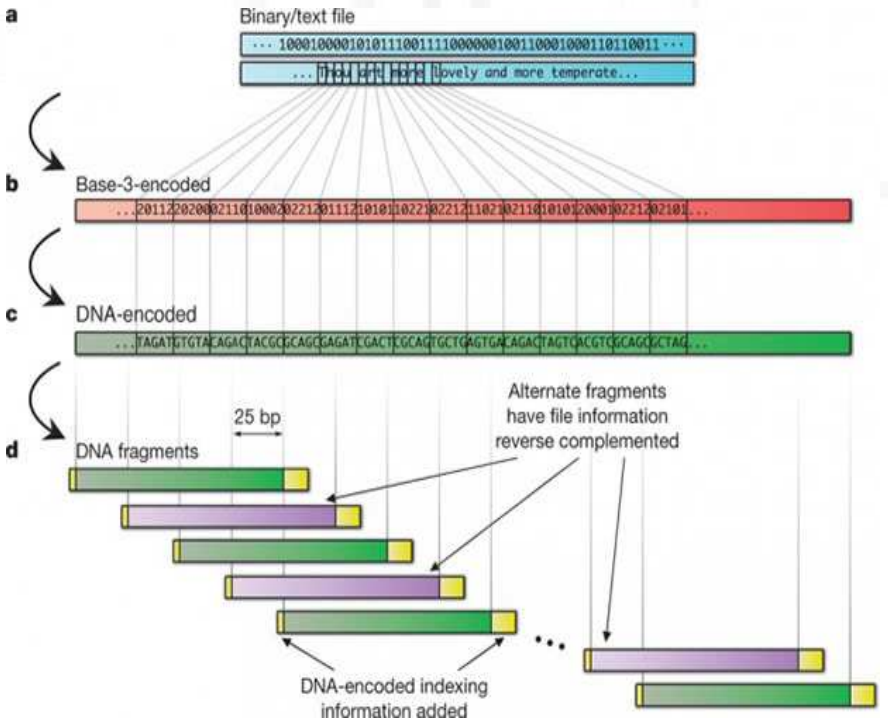
Cloud

Disk, Flash and ???

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

On the Horizon: DNA Based Storage

- Ground breaking research at the European Bioinformatics Institute
- Data encoded using DNA's bases A, T, C and G
- Projected shelf life in the 10,000+ year range
- 2PB of data storage per gram of DNA material
- First trials were able to independently encode and decode several file-based payloads
- AXF will work here too!



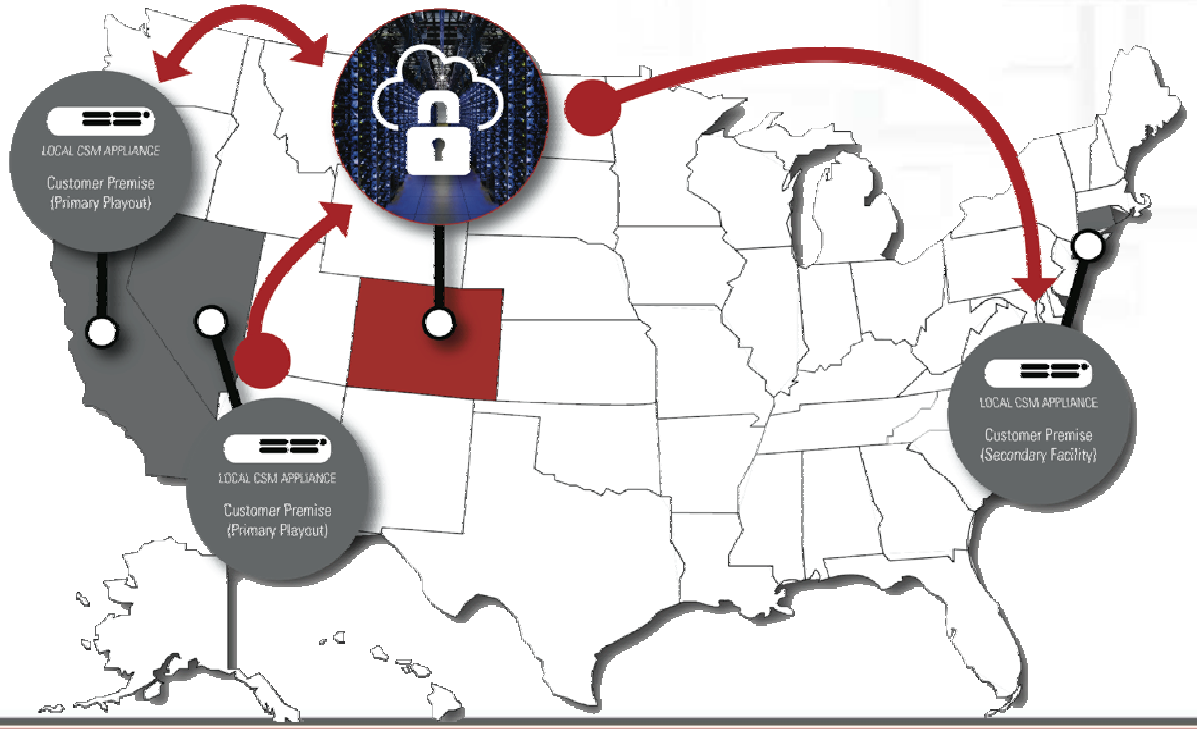
© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

AXF
Archive eXchange Format

AXF Certified Technologies

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

AXF for Content Transport and Cloud Storage



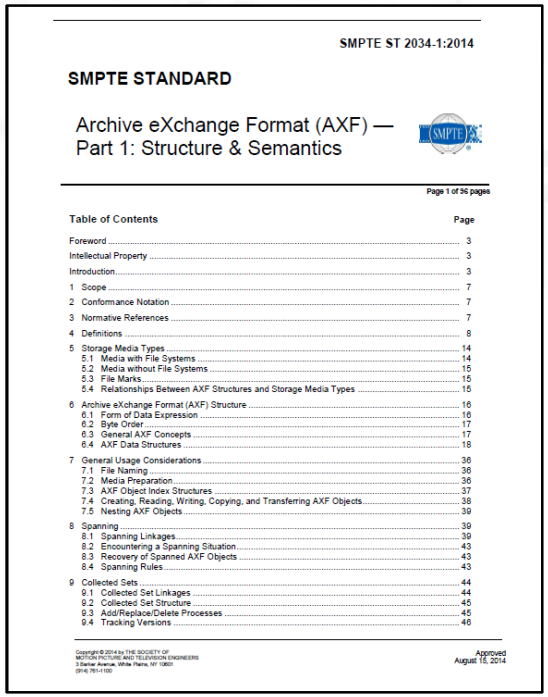
© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

SMPTE AXF Standard

- Society of Motion Picture and Television Engineers recently published the AXF Part 1 Standard **SMPTE ST 2034-1:2014**

- AXF has been deployed in the industry for more than three years with more than 25 organizations around the world using it today!

- Visit OpenAXF.org for more info



SMPTE ST 2034-1:2014

SMPTE STANDARD

Archive eXchange Format (AXF) — Part 1: Structure & Semantics

Page 1 of 36 pages

Table of Contents Page

Foreword ... 3
Intellectual Property ... 3
Introduction ... 3
1 Scope ... 7
2 Conformance Notation ... 7
3 Normative References ... 7
4 Definitions ... 8
5 Storage Media Types ... 14
5.1 Media with File Systems ... 14
5.2 Media without File Systems ... 15
5.3 File Marks ... 15
5.4 Relationships Between AXF Structures and Storage Media Types ... 15
6 Archive eXchange Format (AXF) Structure ... 16
6.1 Form of Data Expression ... 16
6.2 Byte Order ... 17
6.3 General AXF Concepts ... 17
6.4 AXF Data Structures ... 18
7 General Usage Considerations ... 36
7.1 File Naming ... 36
7.2 Media Preparation ... 36
7.3 AXF Object Index Structures ... 37
7.4 Creating, Reading, Writing, Copying, and Transferring AXF Objects ... 38
7.5 Nesting AXF Objects ... 39
8 Spanning ... 39
8.1 Spanning Linkages ... 39
8.2 Encountering a Spanning Situation ... 43
8.3 Recovery of Spanned AXF Objects ... 43
8.4 Spanning Rules ... 43
9 Collected Sets ... 44
9.1 Collected Set Linkages ... 44
9.2 Collected Set Structure ... 45
9.3 Add/Replace/Delete Processes ... 45
9.4 Tracking Versions ... 46

Copyright © 2014 by THE SOCIETY OF MOTION PICTURE AND TELEVISION ENGINEERS

Approved August 15, 2014

© 2014 Oracle | Front Porch Digital
All Rights Reserved
Unauthorized Reproduction or Distribution Strictly Prohibited

DEFINE THE
MARKET
Brian Campanotti
Chief Technology Officer
Front Porch Digital
brian.campanotti@fpdigital.com
@BrianCTO
FRONT PORCH DIGITAL
fpdigital.com