A High-Definition Video Accession Facility

DLF Fall Forum 2007

Brian Hoffman

New York University Libraries

2007-11-06

Background

NYU Preservation Repository

- Central Hub for DLTS Collections
- Multimedia
- Preservation Environment
- Publishing Platform
- 2007: Pilot Collections Ingested

NYU DLTS

- A Software & Services IT Group
- No Broadcast Engineers on Staff

Background

Hemispheric Institute Digital Library Project (HIDVL)
- Analog-to-Digital Conversion
- VHS Accession Format
- Digibeta Archive Format
- H.263 Service Copy

NYU Libraries: Special Collections

NYU Libriaries: Preservation

New Facility: Requirements

Accession to File-based Repository

Accession HD (and SD) Content
- 1080p, 1080i, 720p, 720i, NTSC
- 10 bit, 8 bit
- Frame Rates: 24p, 25p, 29.97p, 59.94...

Create Access Copies
...via a 'Mezzanine' copy

New Facility: Requirements

Basic Editing / Post-production

- Head and Tails
- Access Copies
- (Collaboration with Producers / Before-the-fact Preservation)

Laboratory

- Help DLTS create preservation policy for moving images

Scalable

New Facility: Requirements

Improve Operator Environment:

Concerns

Accession Formats

Preservation Formats

What Exactly Are We Preserving?
 - the information on tape media?
 - moving images?

Data Throughput Requirements
-1080i 30 fps 10bit = 160Mbps

Accession Formats:
Hard to Predict

D5, HDCAM SR, HDCAM, P2, DVCPRO HD, HDV, Digibeta, VHS...

Avid DNXHD 100, DNXHD 220, Apple ProRes, DVCPRO 100 (as file)

Wrappers: MXF, OMF, AAF

Initial Accessions Formats

HDCAM
- DCT Compressed
- 3:1:1 Sampling
- 8 Bit Color
- 144 Mbps

Digibeta
- DCT Compressed
- 4:2:2 Sampling
- 10 Bit Color
- 90 Mbps

VHS

1 Representative
for each:
Analog
NTSC Digital
HD Digital

Preservation Formats ?

Uncompressed DV
- huge files (~500GB / hr ?!)
- non-starter for most organizations

JPEG 2000
- dubious open standard (see MXF)
- low adoption
- still a computer science project

MPEG 4
- lossy

Preservation Formats

Uncompressed DV
- huge files (~500GB / hr ?!)
- non-starter for most organizations

JPEG 2000
- dubious
- low adopt
- still a com

MPEG 4
- lossy

We can take this issue offline and explore
- alternative file compression
- HSM
- wait and see

Anyway, our collections are finite now.

Remaining Questions

What to do to the SDI signal?
   - Apply Corrective Techniques?
   - Normalize Sampling Rates?
   - Normalize Bit Rates?

What does it meant to preserve digital tape media?

Remaining Questions

What to do to the SDI signal?
- Apply Corrective Techniques? ***NO***
- Normalize Sampling Rates? ***TBD***
- Normalize Bit Rates? ***TBD***

What does it meant (for us) to preserve digital tape media?

We are preserving the SDI data stream event created by a machine reading a tape

Initial House Formats

Uncompressed DV
- Sampling: TBD
- Color Depth: TBD
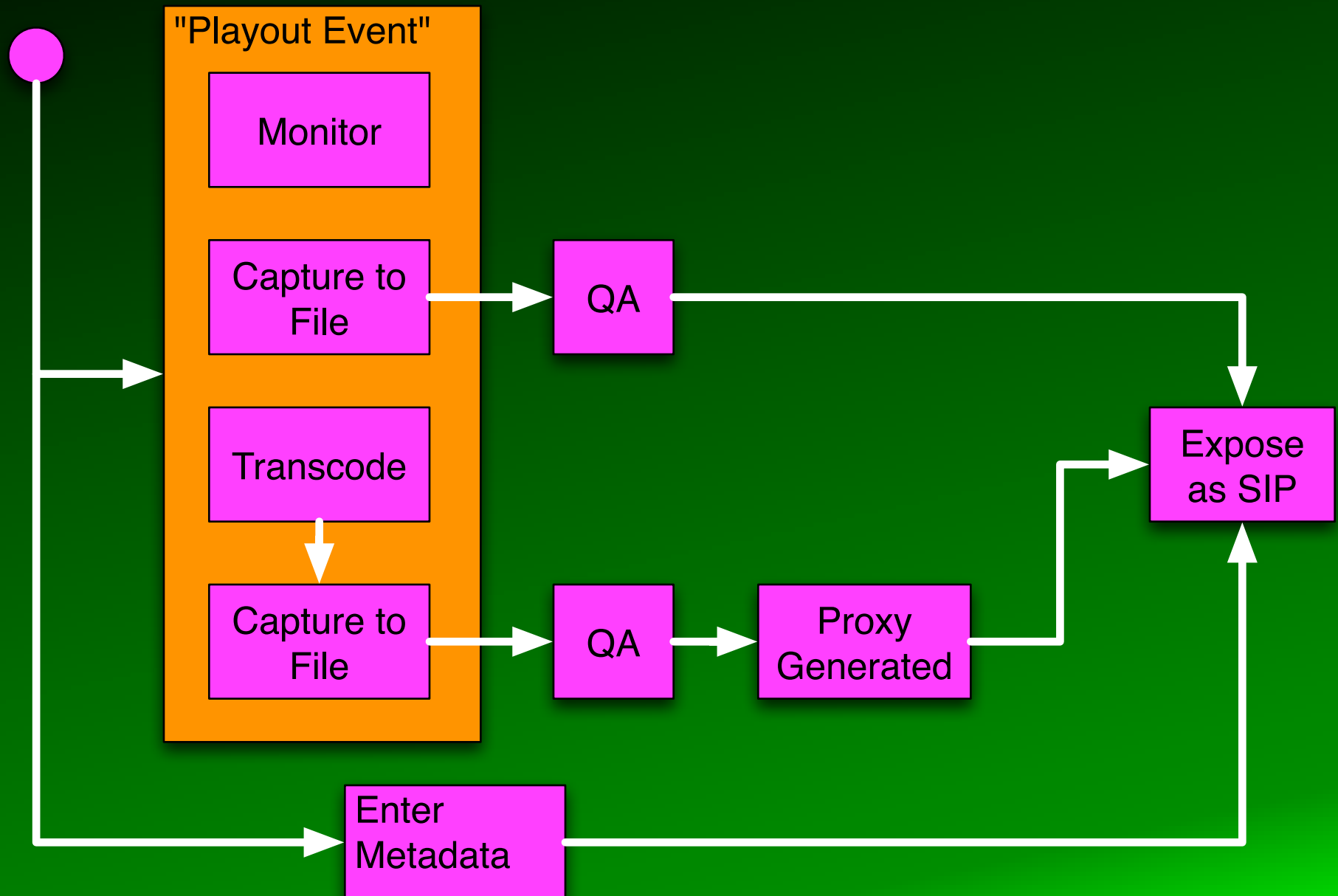- Bitrate = Fn (Resolution, Bit Depth, Sample Rate, Frame Speed)

MPEG-2
- 7 Mbps
- Aspect Ratio TBD

H.264 aka MPEG 4
- 480 x 320

SDI-to-File Workflow

Topology

Preservation
Storage

GbE

SAN Storage
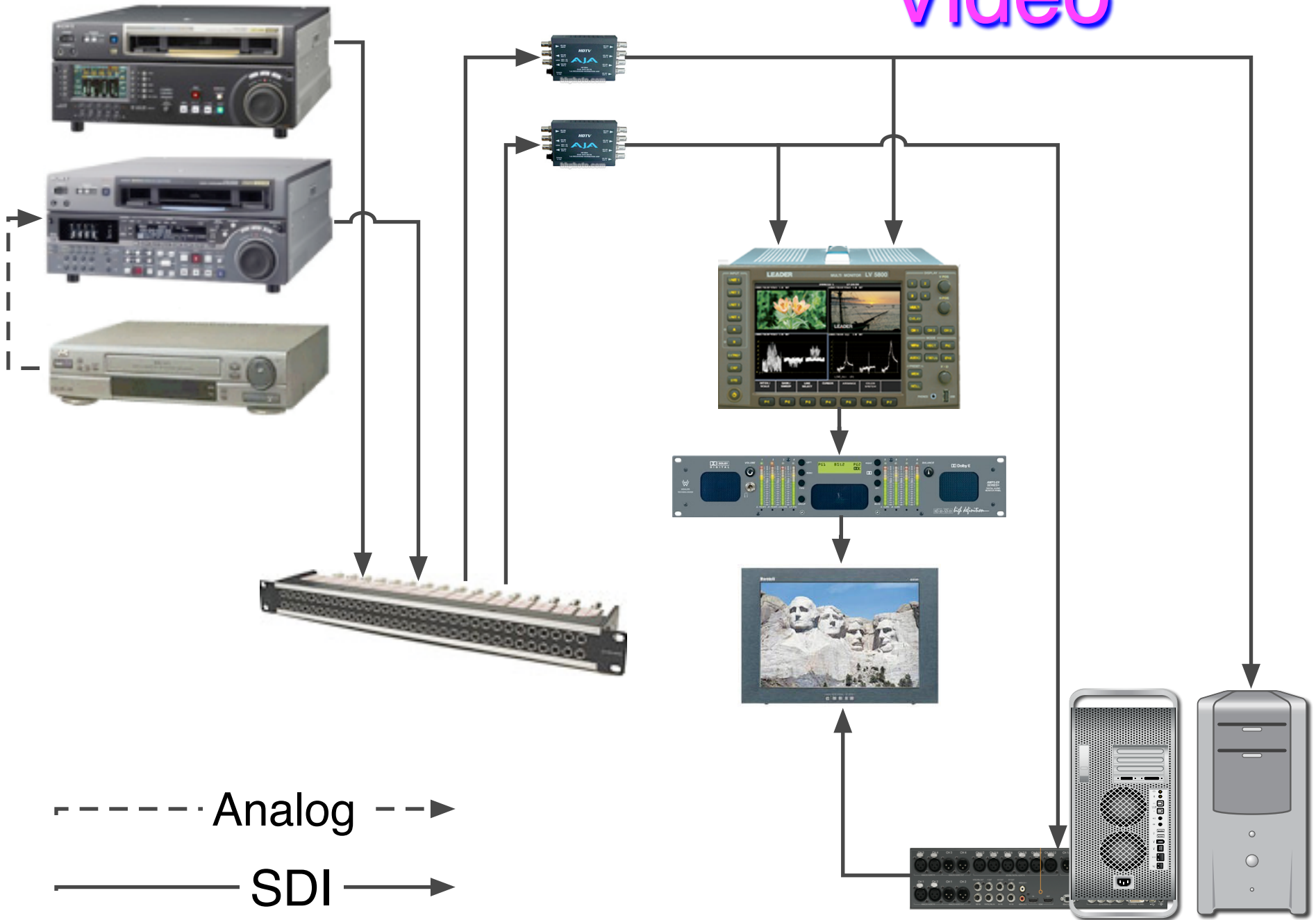
4 Gb Fibre
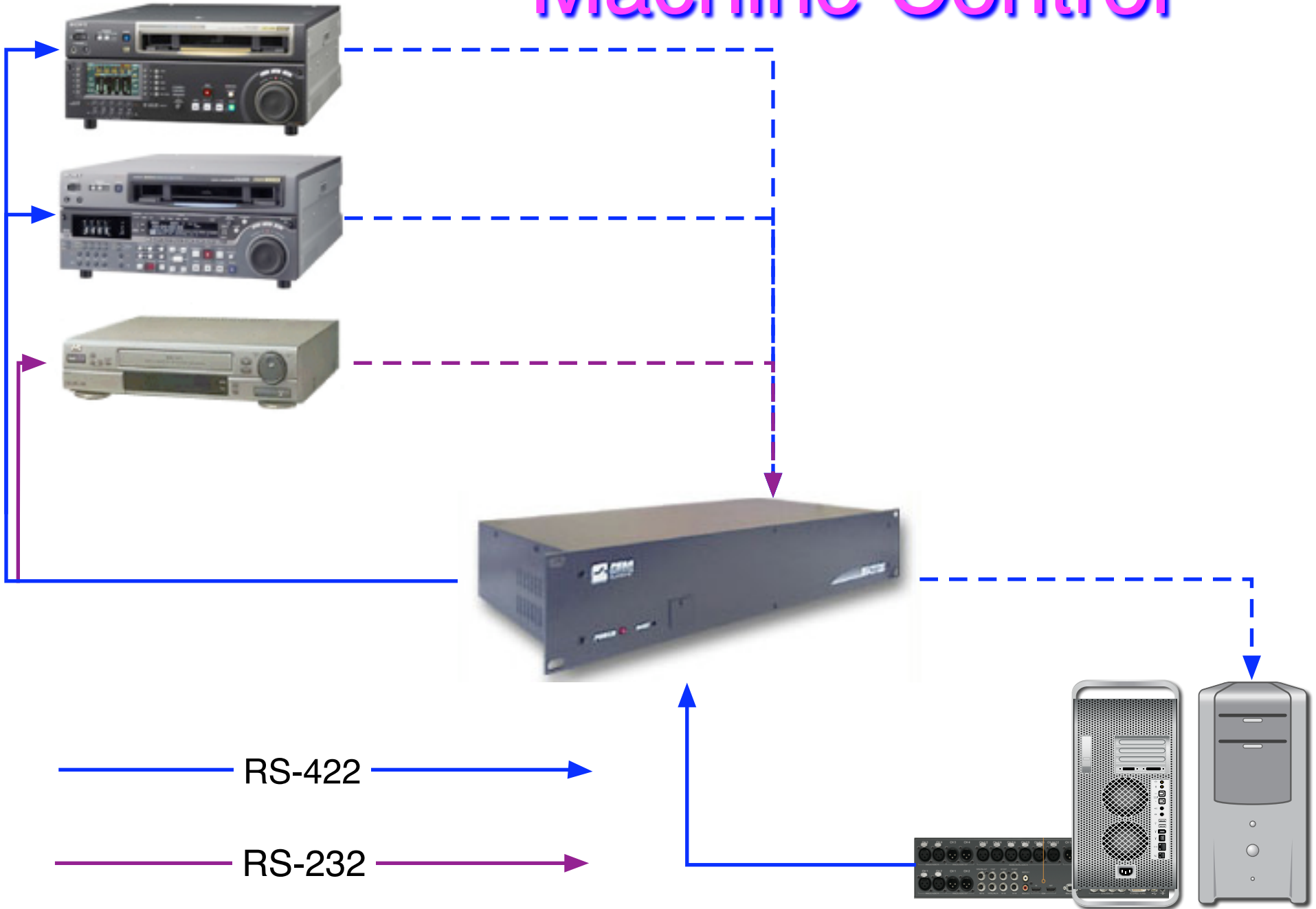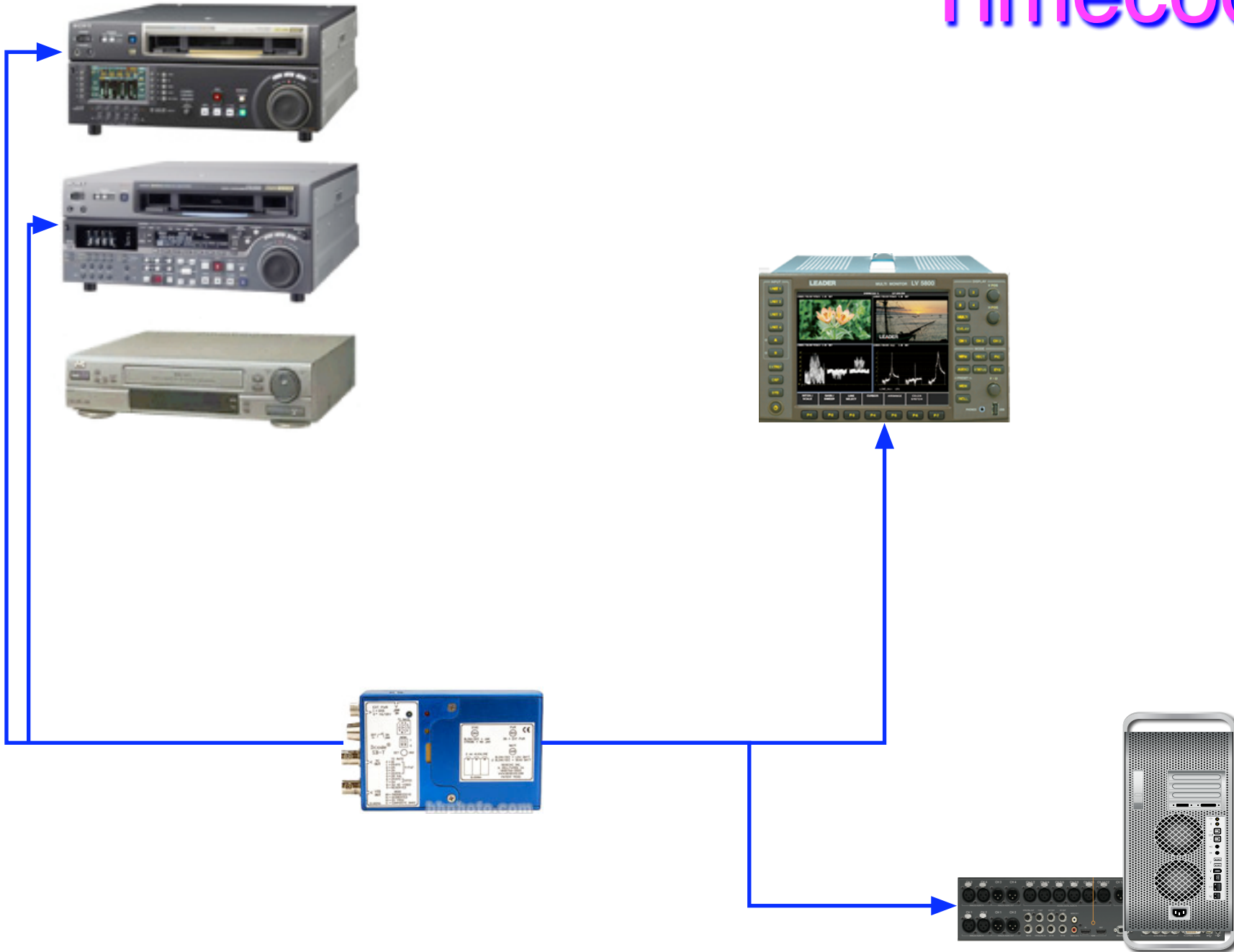
Workflow
Programming

Metadata

QA

Video Ingest

Video

Analog

SDI

Machine Control

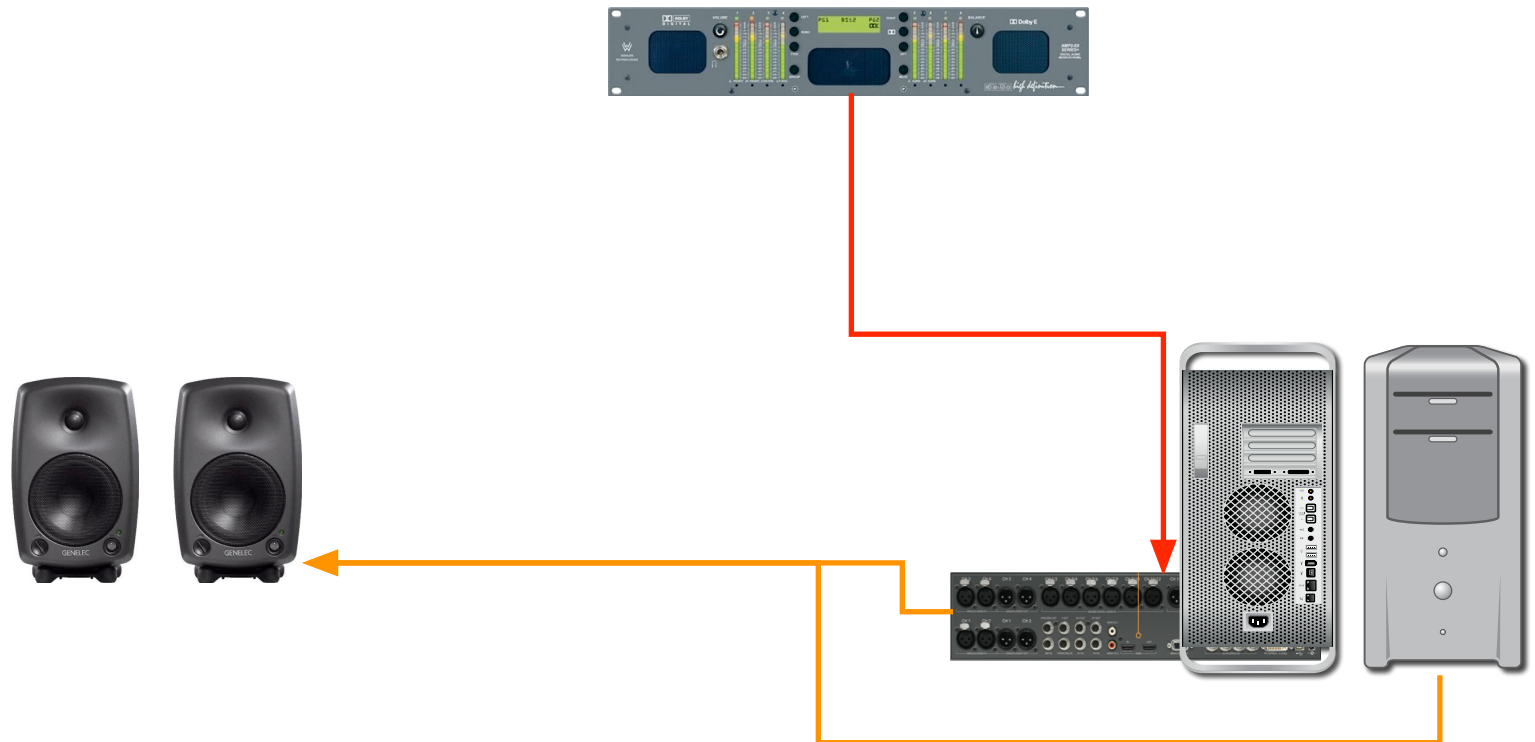RS-422

RS-232

Timecode

Audio

ANALOG →

AES →

Operator Environment

Old:

Operator Environment

New:

Future Work

Install / Wire Equipment

Testing

Ingest a pilot collection

Develop integrated, non-sequential
   metadata workflow

Experiment with file-based proxy
   generation

Experiment with file-based archival
   compression

Expand Accession Format Array

Centralized Video Core?

Acknowledgements

Andrew W. Mellon Foundation
National Endowment for the Humanities
Library of Congress

NYU
   - Dean Carol Mandel
   - Dr. David Ackerman
   - Roddy Austin
   - James Bullen

Contact

Brian Hoffman
brianjhoffman@nyu.edu