Library of Congress
Designing Storage Architectures

April 15, 2024

Francis Badzey

Executive Director, Global Archives & Preservation Services
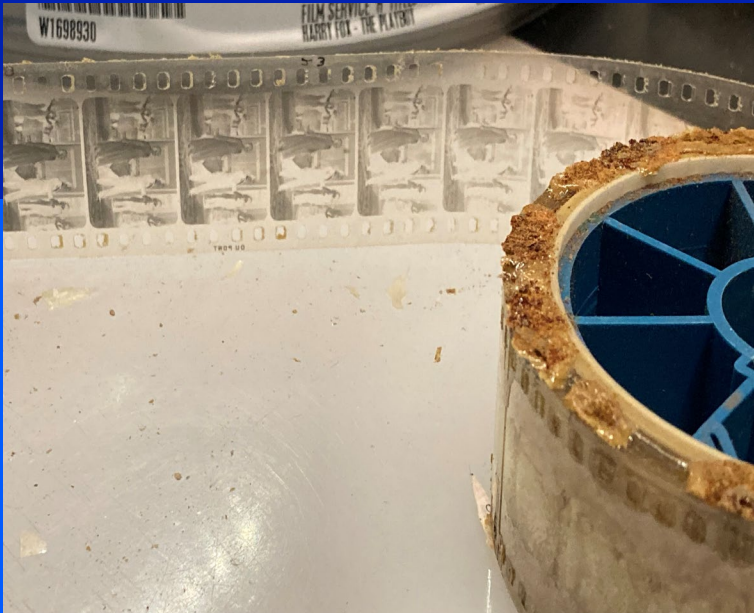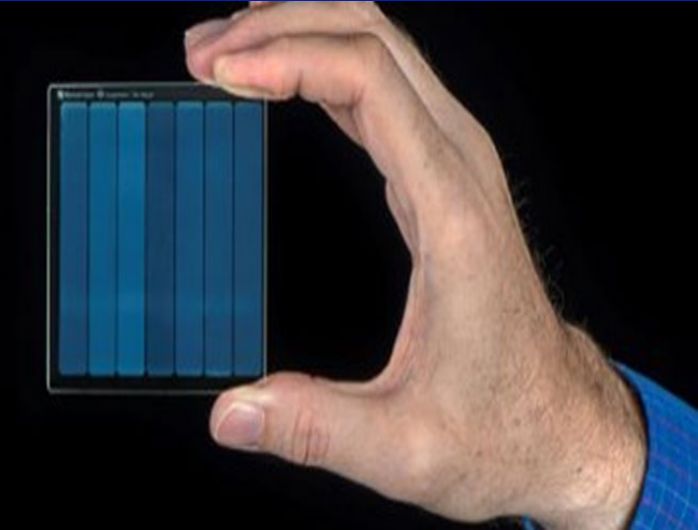
WARNER BROS. DISCOVERY

Agenda

1	A Quick Review
2	The Storage Dilemma
3	Everything is Bigger
4	Data Reduction POC
5	Questions?

WARNER BROS. DISCOVERY

Quick Review

- Physical Storage Challenges

- Digital Storage Challenges

- Goal: Multi-Cloud

- Goal: New Media Format

WARNER BROS. DISCOVERY
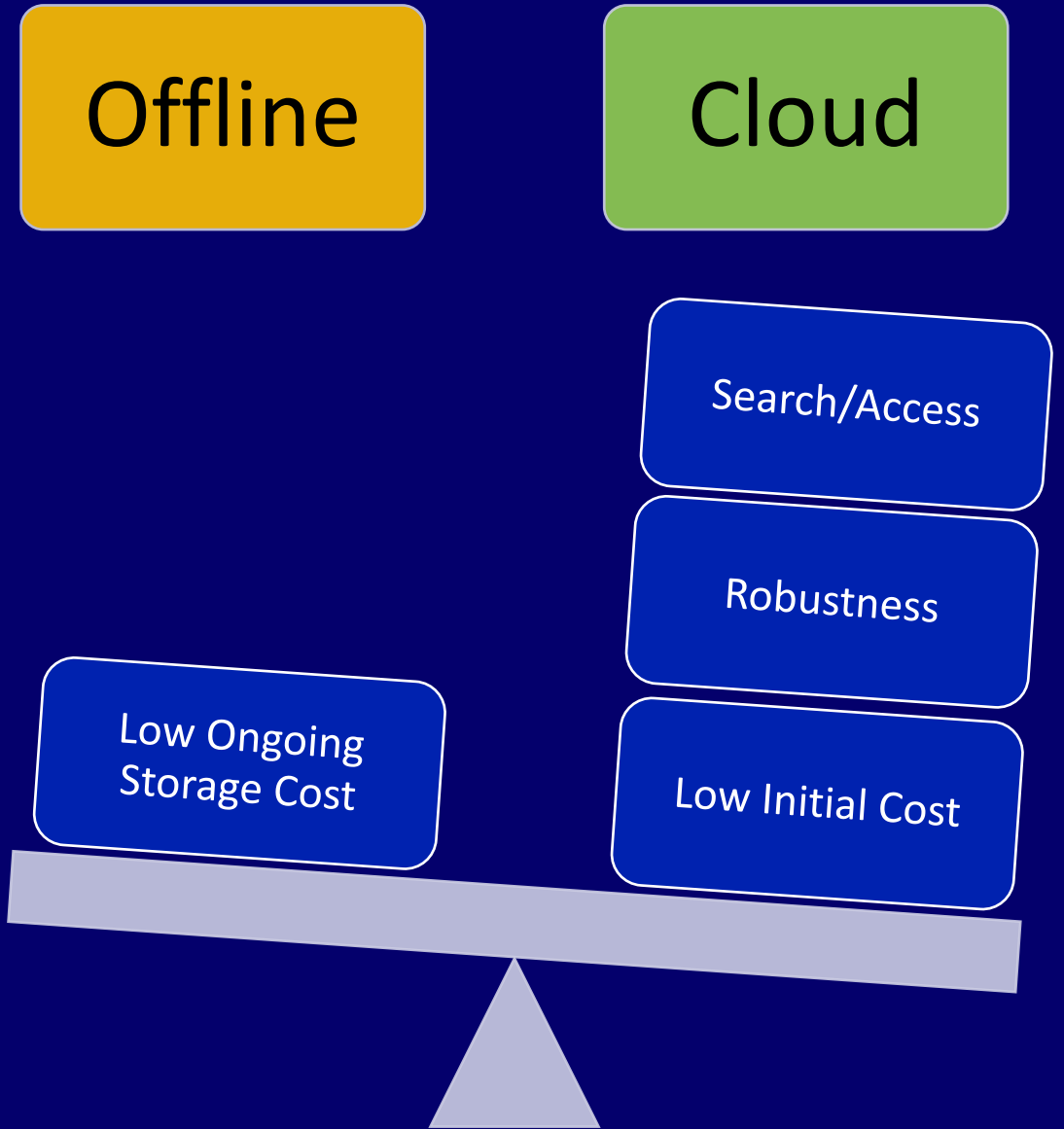
The Storage Dilemma

Offline

Cloud

"We Can't Treat Everything the Same"

Search/Access

Robustness

Low Ongoing Storage Cost

Low Initial Cost

Our growing Digital footprint

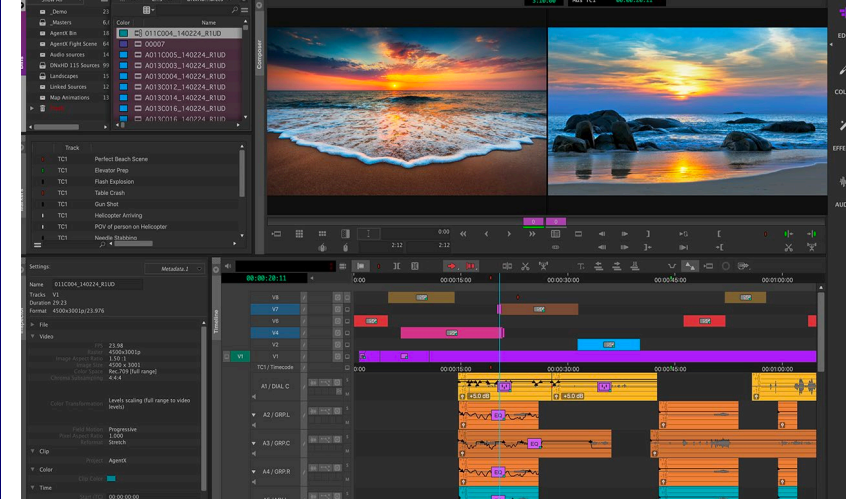WARNER BROS. DISCOVERY

Everything is Bigger

- Bigger Files – Resolutions are getting bigger

- Increased Shooting Ratios

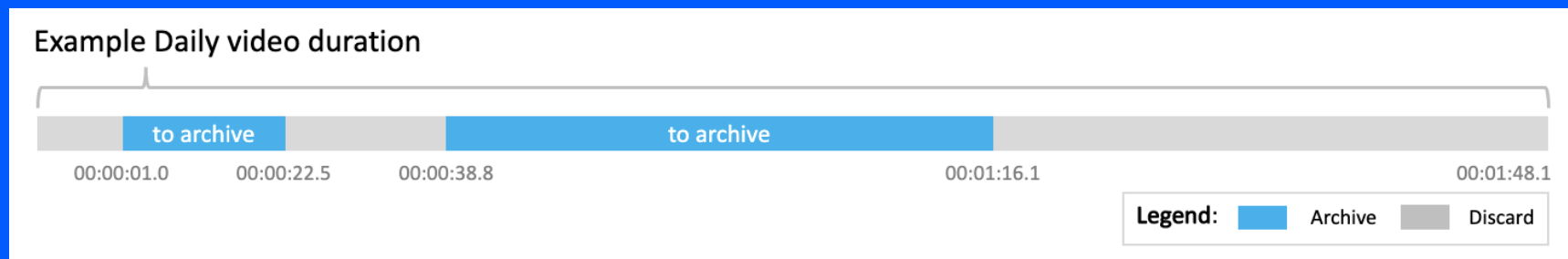- Increased Physical footprint

The Balancing Act

- Some OCF does go in the cloud
  - Select takes (Similar to Cut Negative), EDL's, QT's
  - The essence of program content is readily available
  - Enables egress and re-assembly for future use

- LTO Maintenance and Migration Plan
  - Reduce storage footprint via increase tape density
  - Identify media with longer migration times
  - Discard unneeded data

Data Reduction POC

- Project IOTA – co-engineering collaboration between WBD and Microsoft's Commercial Software Engineering team

  - Machine Learning identified 45% of material on "Digital Dailies" could be discarded

  - Three-pronged approach:
    - Detect Clapper boards
    - Use OCR to extract rolls, scenes and takes
    - "Listen" to audio to detect "Cut!"

Example Daily video duration

to archive		to archive	
00:00:01.0	00:00:22.5	00:00:38.8	00:01:16.1	00:01:48.1

Legend: Archive Discard

- https://devblogs.microsoft.com/ise/archiving-footage-deep-dive/
- https://devblogs.microsoft.com/ise/building-an-action-detection-scoring-pipeline-for-digital-dailies/

Questions?

WARNER BROS.
DISCOVERY