



Shifts in Storage Landscape... Managing Both Sides of the "Plate"

Nathan Thompson
Chief Executive Officer
Spectra Logic

October 24th, 2019

Current Challenges

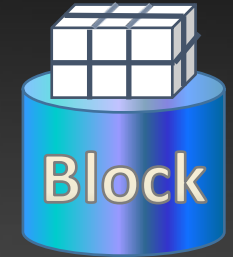
The storage world is complex and dynamic for large storage users...

- The competition is broader, more sophisticated and stronger
- Storage is heavily influenced by IT savvy giants
 - Social Media, Amazon, Google, Twitter
- Users want greater agility & efficiencies
- Want to make clever use of most advanced technologies available in the most dynamic ways
- Want seamlessly automated, dynamic workflows
- Need the ability to quickly change workflows with changing business needs
- Vast amounts of newly generated data/content needs efficient management
- What modern technologies to re-utilize existing assets

Progression of Storage

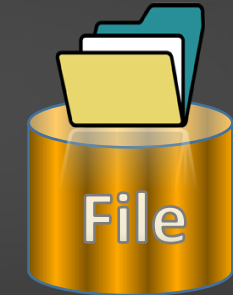
Block storage

Directly attached
Great performance
Complex
\$\$\$



File-based storage

Network attached
Shareable, easy
Limitations on size, # of
files, geographic dist.
\$\$



Object storage

Network based
Shareable Globally
No limit on scale
Friendly to Manage,
Search and Automate
\$



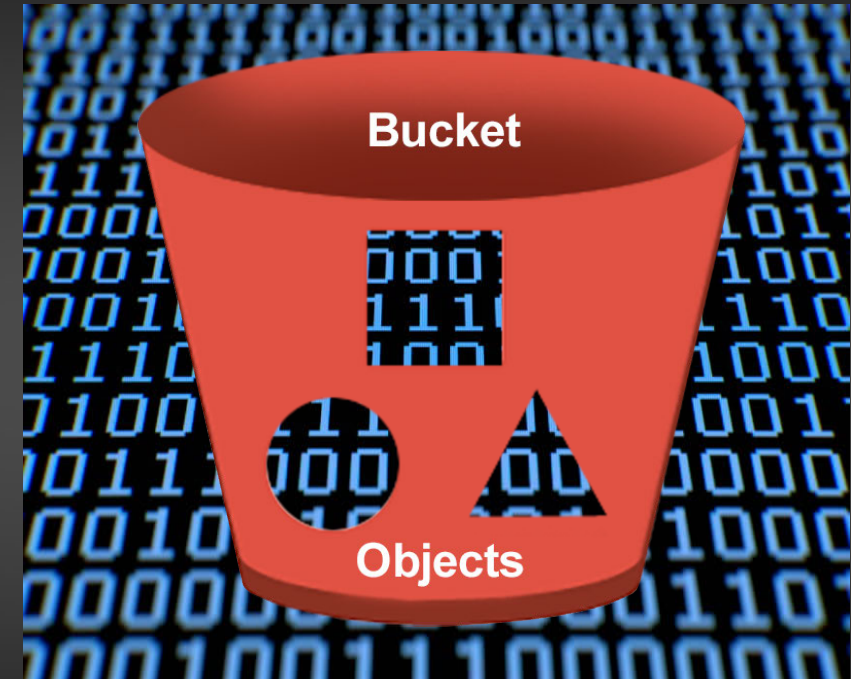
Exploring Object Storage

What is Object Storage?

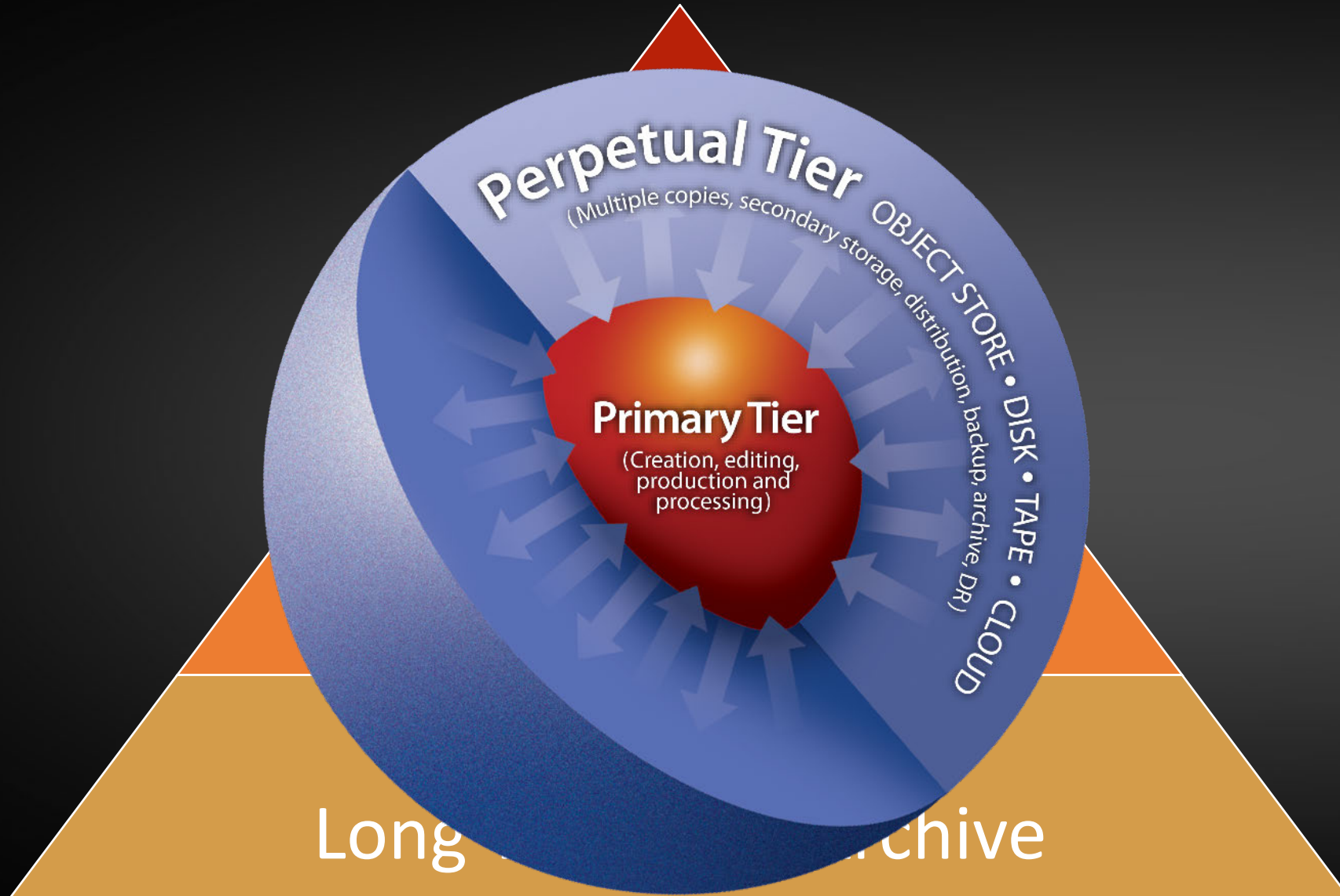
- An alternative to file based systems; ideal for storing large volumes of unstructured data.
- Decouples data from its physical medium or location
- Employs the inclusion of Meta Data, and Universal ID
- Its flat & infinite namespace make large scale storage possible
- Provides a foundation for data longevity techniques

How do you talk to Object Storage?

- RESTful API
- Client server model
- Gateways/Appliances

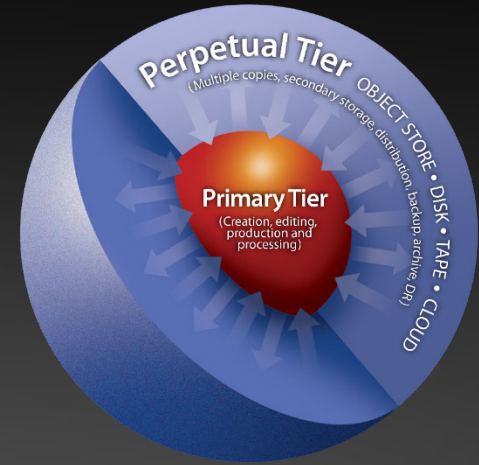


The Original Storage Pyramid



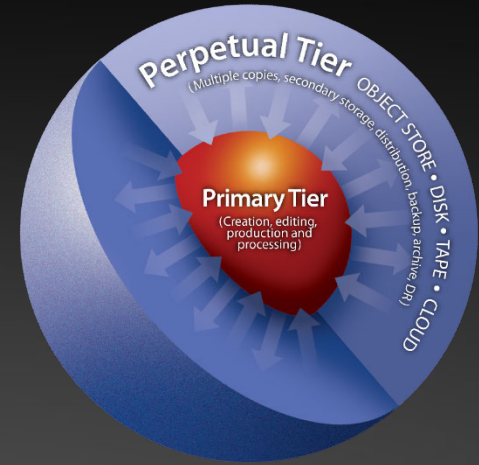
What is the Primary Tier ?

- File based
- A high performance platform (Flash or Enterprise HDDs)
- Higher cost storage
- “Work-in-Progress”
 - Data capture/creation
 - Transacting/editing
 - Rendering/modeling
- Content is not available widely and so usually on premise
- *When work is done, no need to keep it here – The longer it stays, the more it costs*

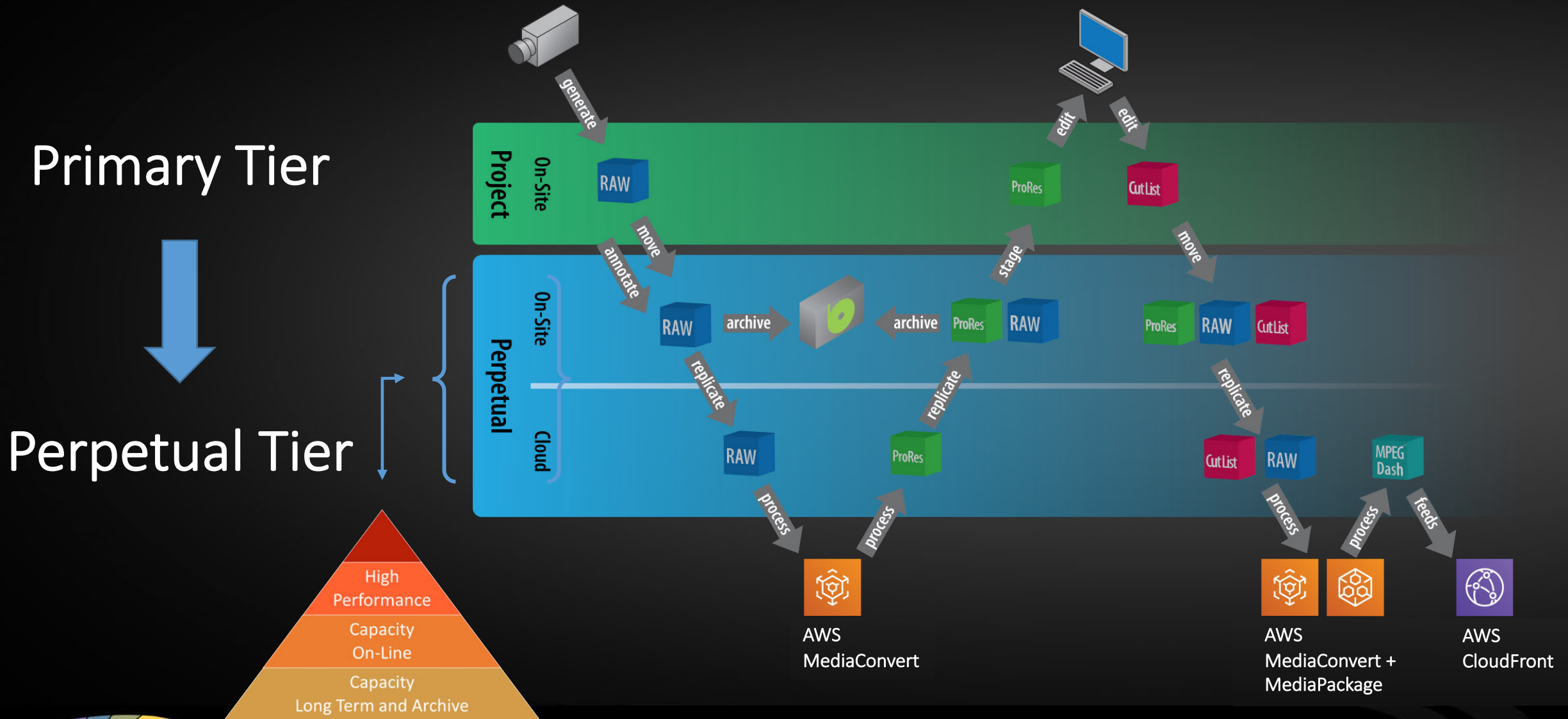


What is the Perpetual Tier ?

- Object based (on-premise, remote and/or public cloud)
 - Massively scalable
 - No size limitations
 - No limitations on number of files
 - *Includes metadata so easily searchable*
- A less performant platform – but sufficient for most operations
- Lowest cost storage (Glacier, Azure Cold or local tape looking like Glacier)
- Standard base (HTTPS, RESTful, LTFS, etc..), Open
- Web based model - breaks the barriers of on-premise and cloud
 - Content is available widely
 - Can be local and/or geographically distributed
 - Multi-tenant - Content can be shared with different people and different applications via secure credentials
- Keeps content protected & available via policies that define expenditure and workflow requirements

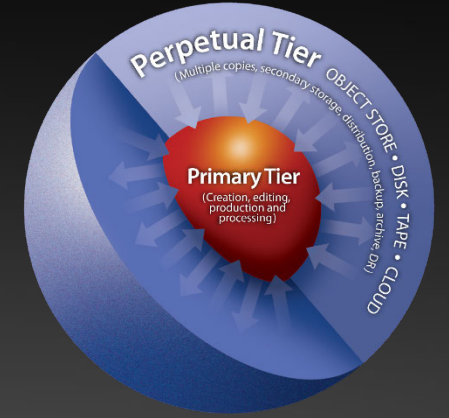


What Does the New 2-Tiered Model Look Like ?



Perpetual Tier Benefits

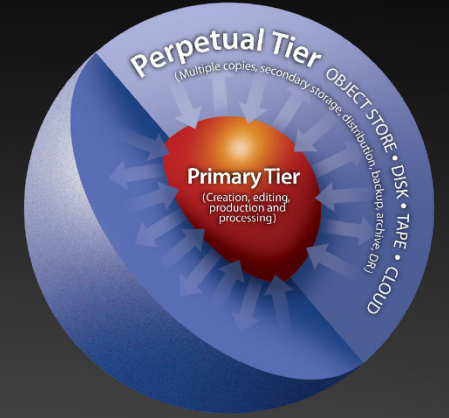
- Actionable Intelligence built-in
 - Health and integrity checking of content
 - Self healing of assets when necessary
 - Automated migration to new technologies offering lower cost/GB
 - Removes pain of migrations
- Supports performance optimization relative to workflows
- AI friendly
- Breaks barriers between on-premise & cloud storage
 - Seamless cloud integrations
 - Hybrid cloud workflows
 - Business insurance – DR
 - Provides leverage against cloud runaway costs



Perpetual Tier Benefits

- Mechanisms to control cloud egress costs
- Supports faster collaboration and sharing
- Supports metadata driven workflows
- Flexible with great affinity for automation
 - Operational & cost efficiencies
- *Some Production Tier functions are moving to Perpetual Tier*
 - Transcoding
 - Streaming proxies

Only functions requiring high performance are left on Primary Tier





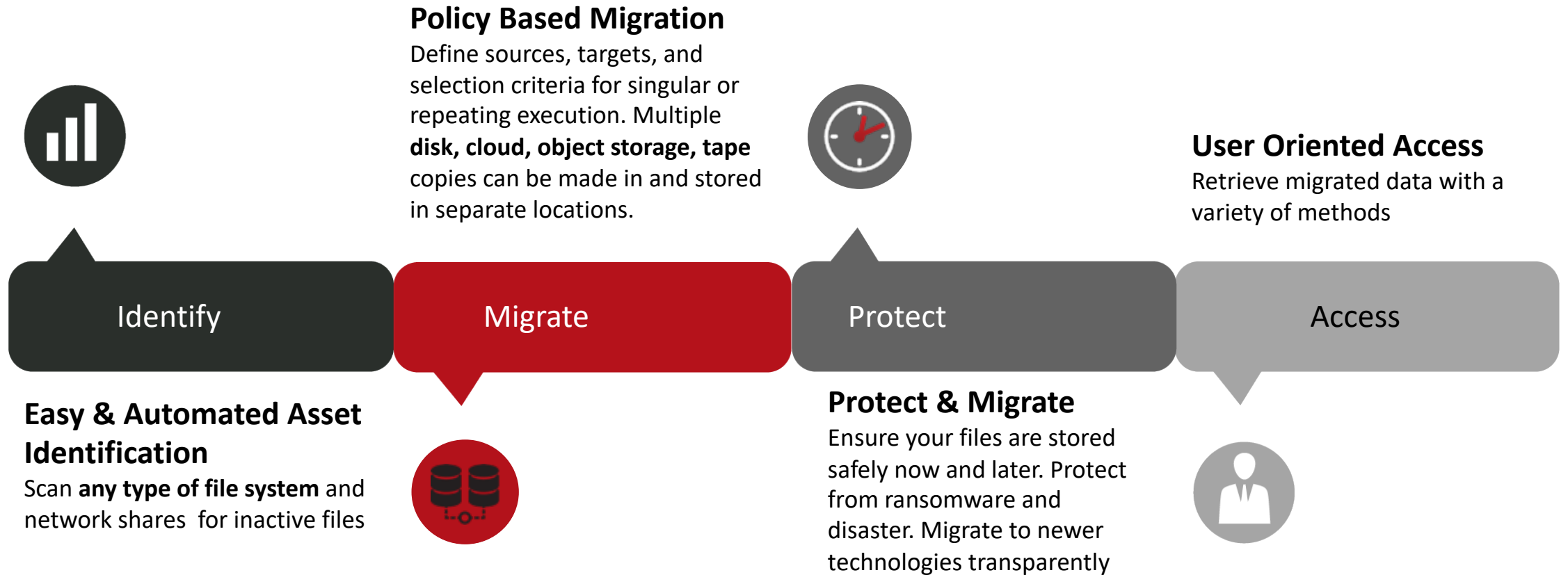
Implementing a Two-Tiered Storage Strategy

Jeff Braunstein
Director Of Product Management

Traditional Approach to Address Data Tiering: Hierarchical Storage Managers (HSMs) HPSS, DMF...

- Work well in some environments
- Expensive, complex, time consuming to manage
- File centric – not data centric
- Supports organizations' data migration workflows
- Requires savvy users
- Doesn't support modern storage architectures
- Proprietary data formats
- File stubs to tape not intuitive, don't work well in many organizations

A Modern Approach to Tiering – Data Management

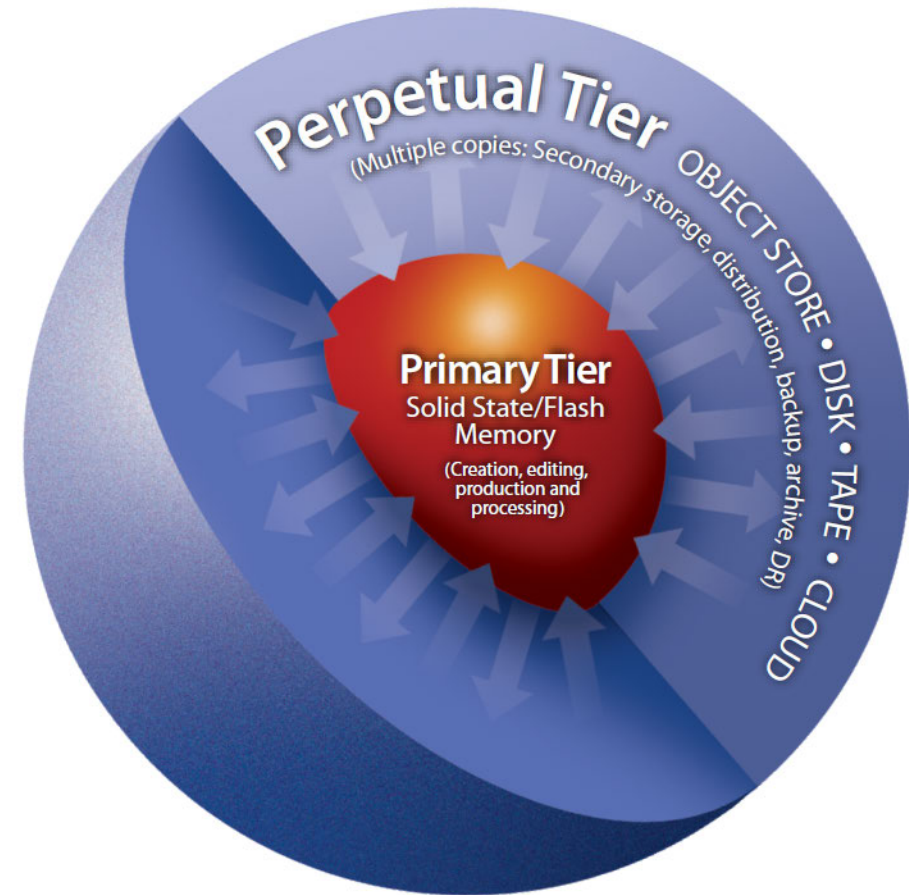


What to Look for in a Data Management Solution

- Low price relative to target storage
- Not priced primarily on capacity
- Agnostic to type of storage
- Allows both migrating and copying of data (DR protection)
- Strong focus on protecting migrated data
- Tight integration to tape and other high-latency storage targets (more later)

Imagine Always Having Your Data in the Right Place

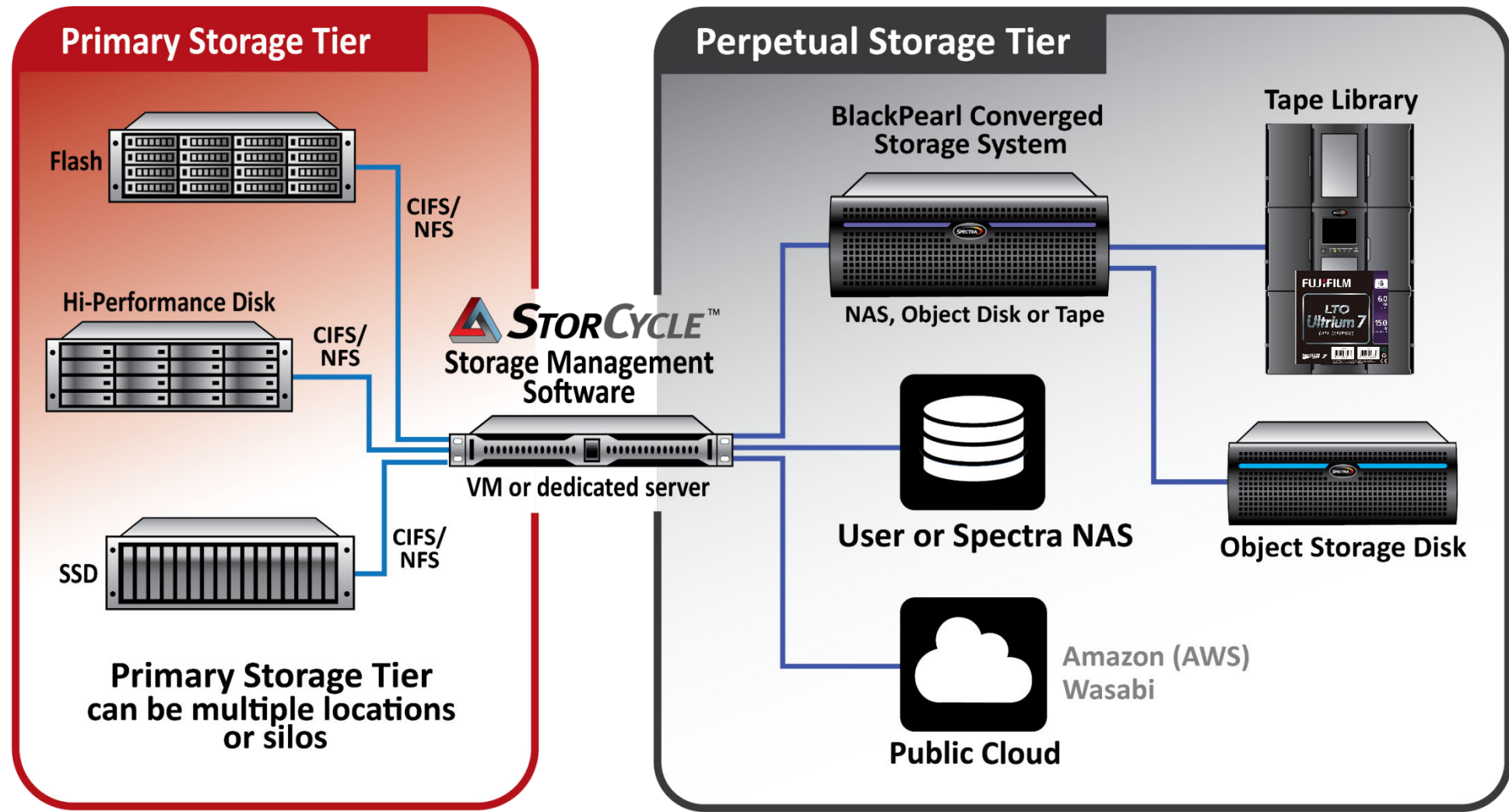
- Easily accessed
- Securely stored
- Faster Primary Tier
- Limitless storage capacity
- Easy collaboration and use of cloud workflows
- Maximum cost savings





New Storage Management Software from Spectra Logic

Spectra's Approach to the Two-Tier Model



Required Data Migration Types

Migrating Project Directories

- Projects, experiment results, etc.
- Want to move all files in directory
- One-time event
- Will move other projects in same way

Auto-Migrating Inactive Data

- Automated policy based on age & size
- General IT data
- Find and move infrequently accessed data
- Recurring, ongoing event

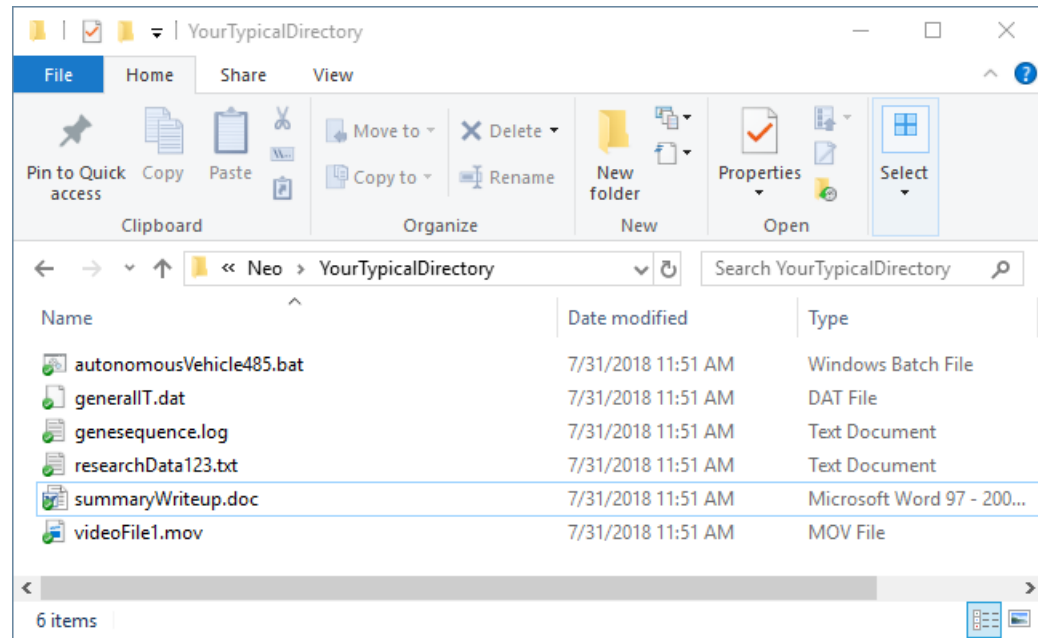
Access Methods for Different Types of Storage

How do users get to data that is moved to Perpetual Tier?

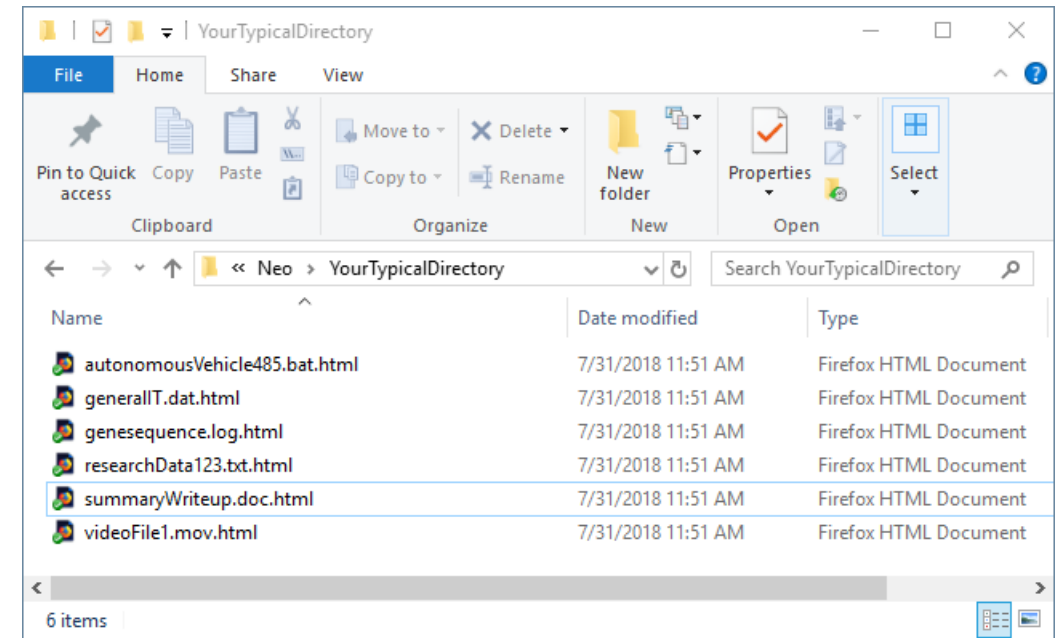
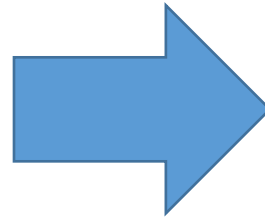
- Moved to immediately accessible storage (i.e. disk)
 - Transparent access can be provided via Symbolic Links or other methods
 - Users and applications unaware data has been moved
 - Most data management solutions are good at this
- Moved to storage with longer restore time (e.g. tape, AWS Glacier)
 - Can't fool user into thinking data hasn't been moved
 - HTML Links allow users to restore data themselves
 - Most data management solutions aren't good at this
- Good data management solutions will also include search restore tool

The Right Way to Address High-Latency Targets

Addressing User Access to Data on Targets like Amazon Glacier and Tape




Before Move



After Move

Spectra StorCycle

File | slmktg01.sldomain.com/Dept/Marketing/JeffBr/neotesting/ResearchProject-06/data7/fi...



Spectra StorCycle


This File has been Archived

Files have been moved from this device to archival storage, both for long-term preservation and more efficient use of primary media. The file has been replaced by this html file, but it can be recovered using the directions below.


File Information

Name: data7\file21.txt
Archived date: Tuesday, 25-Jun-19 22:51:44 UTC
Last accessed: Tuesday, 25-Jun-19 21:45:09 UTC
Size: 1.0 MiB (1,048,576 bytes)
Archived to: BP Five
Task: project 6
Task Description:

Recovery Options

 **Restore Now**

Restore *data7\file21.txt* to original location

 **Restore Project Files**

Open Spectra StorCycle to Recover files in *project 6-1*

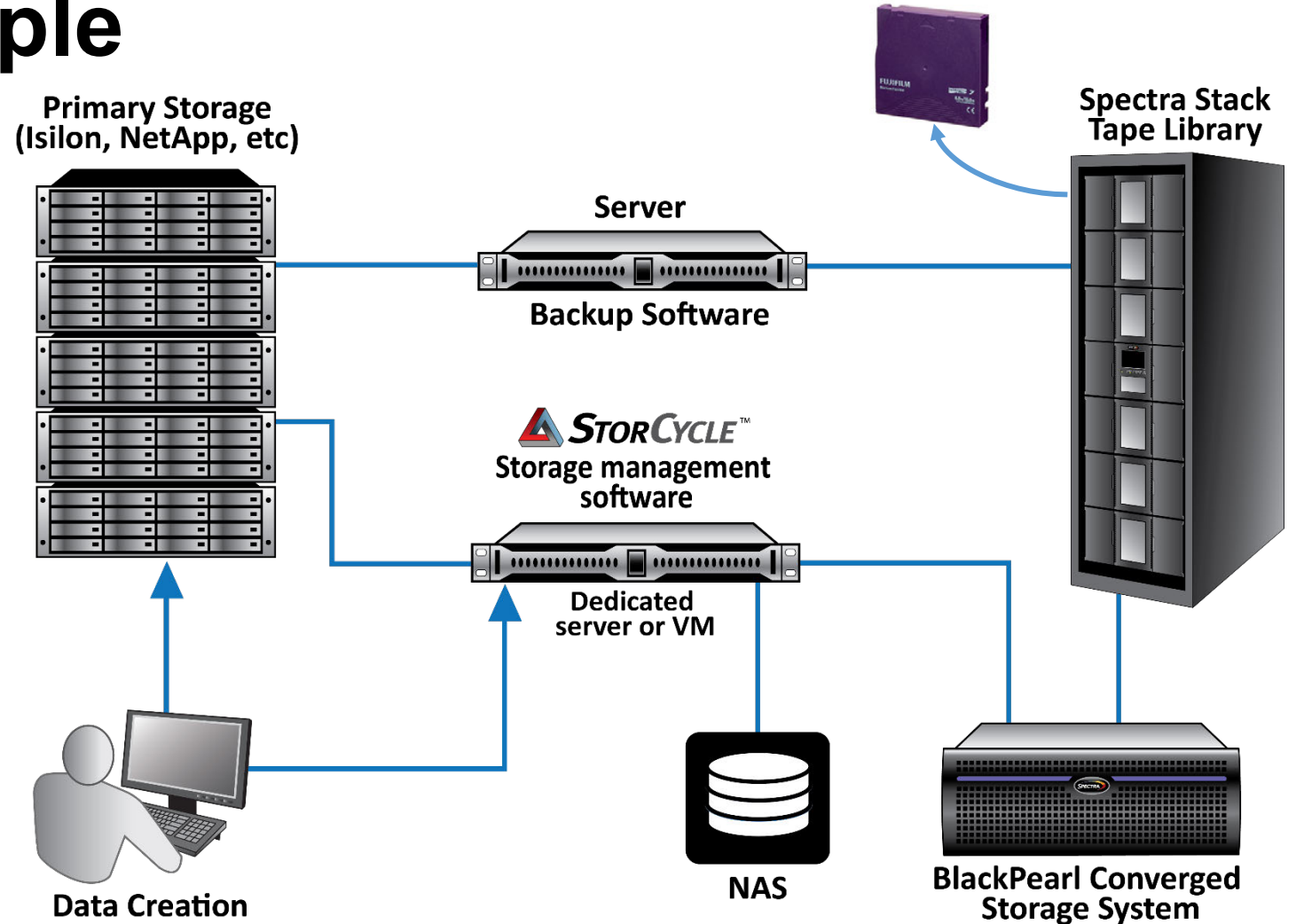
Manual Recovery

This file was written to the following device(s):

- Storage Endpoint: BP Five

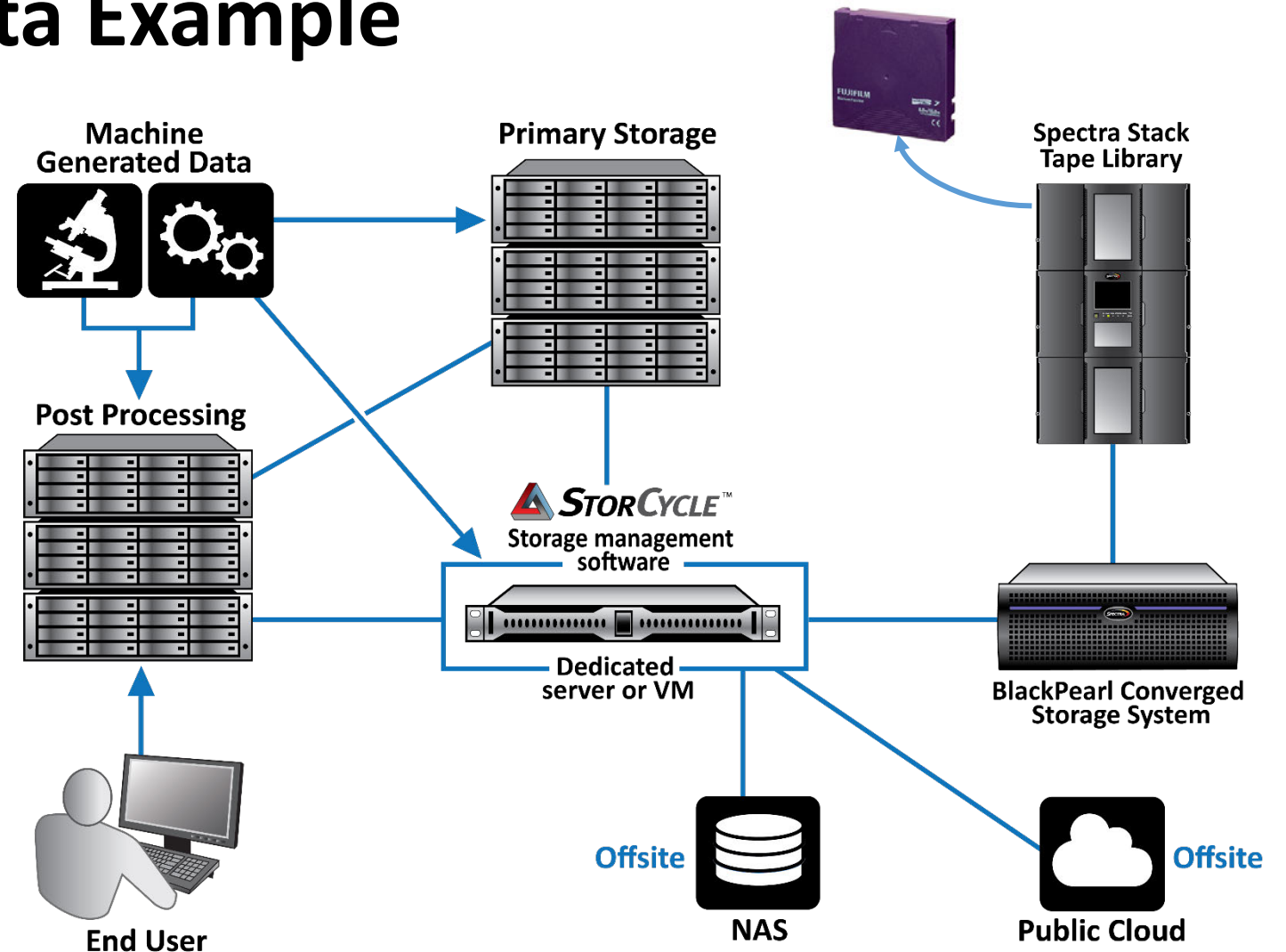
Corporate IT Example

- Goal 1: Provide unlimited storage to all employees
- Goal 2: Reduce the backup window and backup costs.
- Scan, identify, move, and provide a breakdown of storage contents helps the IT department needed to solve data growth problems.



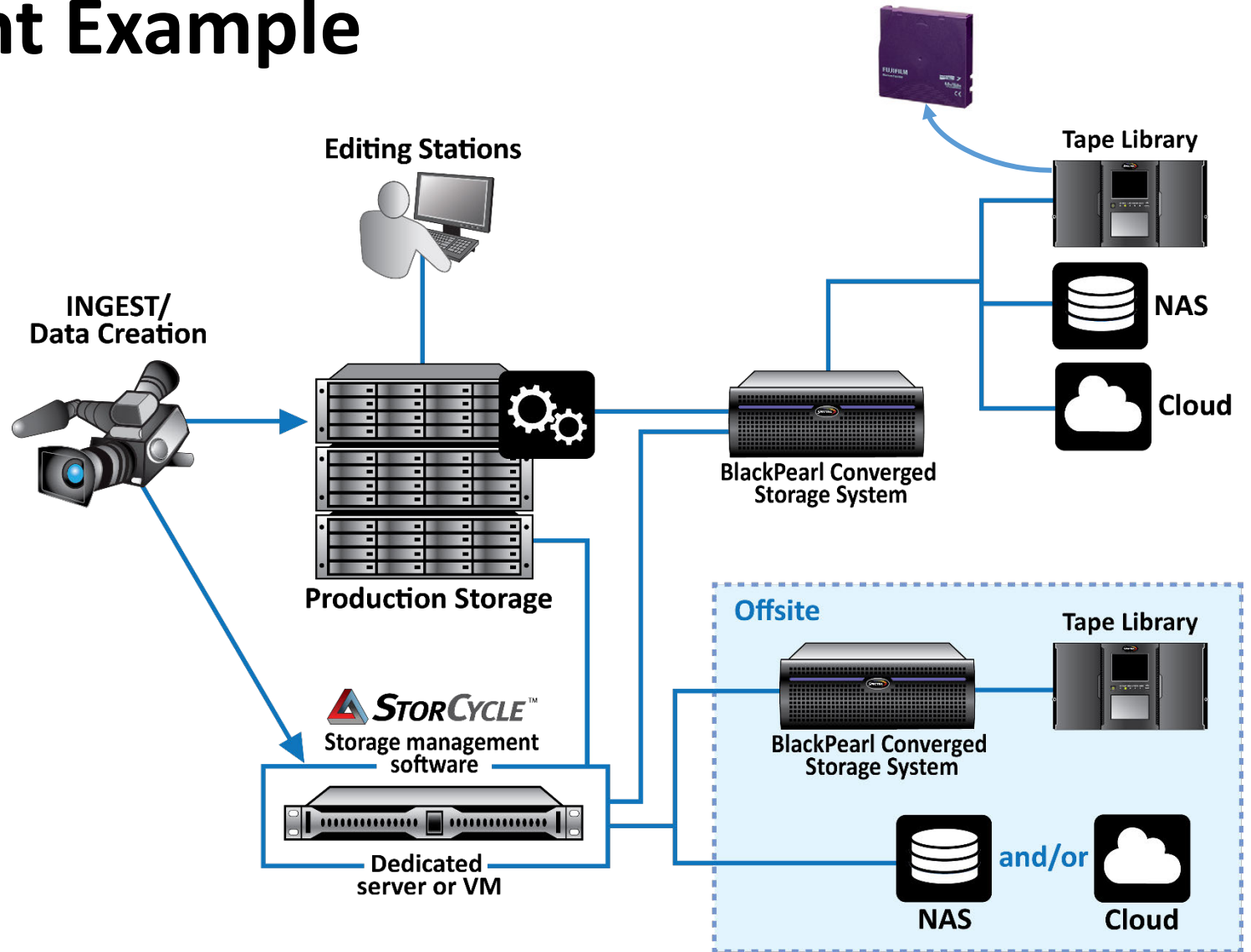
Machine-Generated Data Example

- Goal 1: Replace existing HSM system that does not work as advertised
- Goal 2: Archive full projects upon completion, freeing up primary storage space
- Instrument data to be written directly to the post processing storage tier, as well as a copy being written directly to the archive based on predetermined storage policies



Media & Entertainment Example

- Goal 1: Raw material ingest – store masters for redundancy
- Goal 2: Archive final projects/products
- Goal 3: Disaster recovery copy
- Migrate video and image assets to Perpetual Storage Tier while providing easy/transparent access to users





Thank You!

Nathan Thompson
nathant@spectrallogic.com

Jeff Braunstein
jeffbr@spectrallogic.com