SPECTRA
SINCE 1979

Answers for the Data Explosion Phenomenon

Changing The World of Storage

Molly Rector

- Spectra Logic (www.spectralogic.com)
  - VP of Product Management & Marketing
- Active Archive Alliance (www.activearchive.com)
  - Chairman of the Board
- INSIC (www.insic.org)
  - Tape roadmap committee member

Key Take Aways:
When looking at a 1PB of data

1. Metrics to use in assessing the reliability of your BigData / MassStorage environment
2. Technologies to consider in keeping your data secure
3. Considerations for disaster recovery
4. Scaling and accessing information…today and years into the future
5. Finding the right balance of speed, density and cost

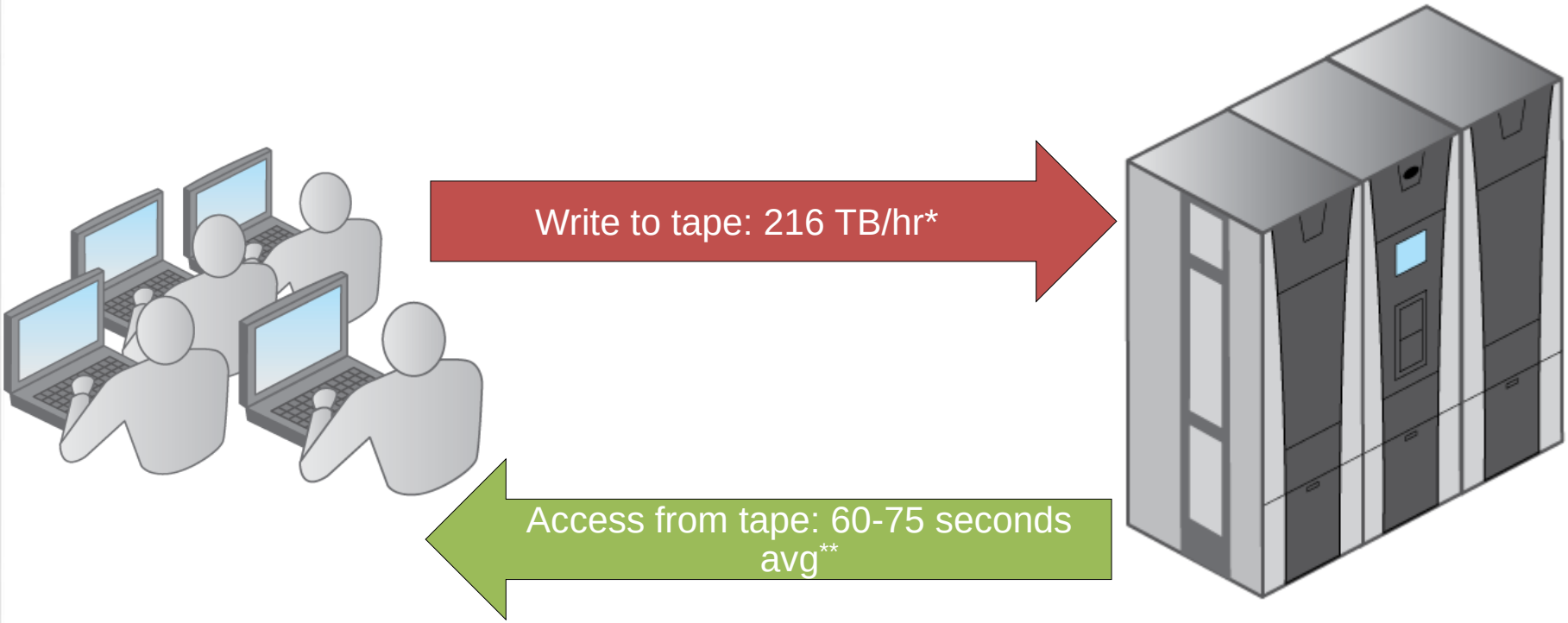

SPECTRA

"Big Data" and "Information Explosion": different definitions to different people

a. A mass of data so large it's difficult to store, move or manage
b. Unstructured, file-based data growth
c. Infrastructure and tools to drive value out of created and stored data

All Data Intensive markets MUST have the greatest Reliability, Density, Scalability and Energy Efficiency as possible

Writing & Retrieving data from tape



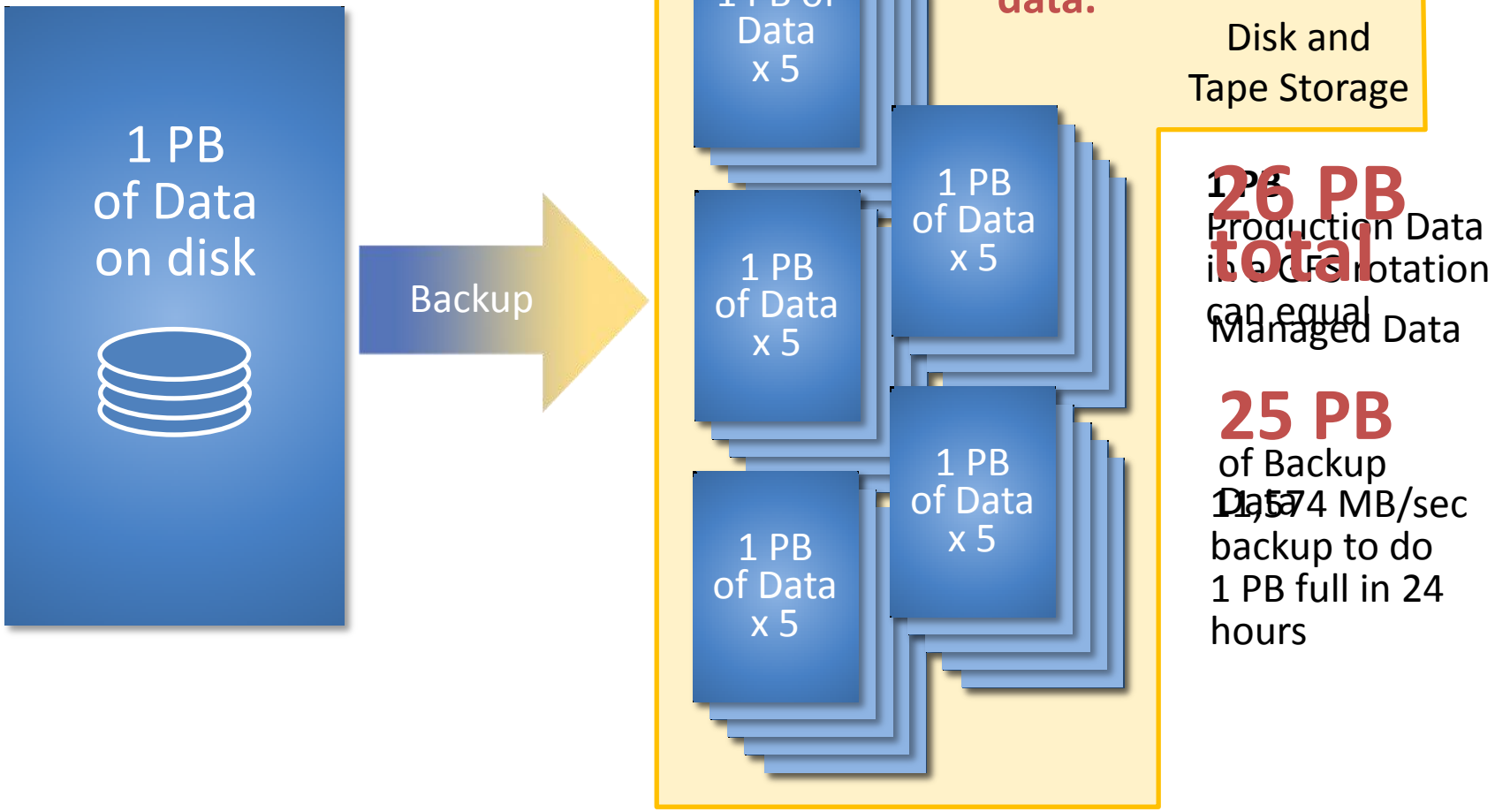
* Based on benchmarked data; assumes 2:1 data compression

** Times vary based on library and tape drive in use. Assumptions based on 120 IBM TS1140 technology tape drives in a Spectra T-Finity library

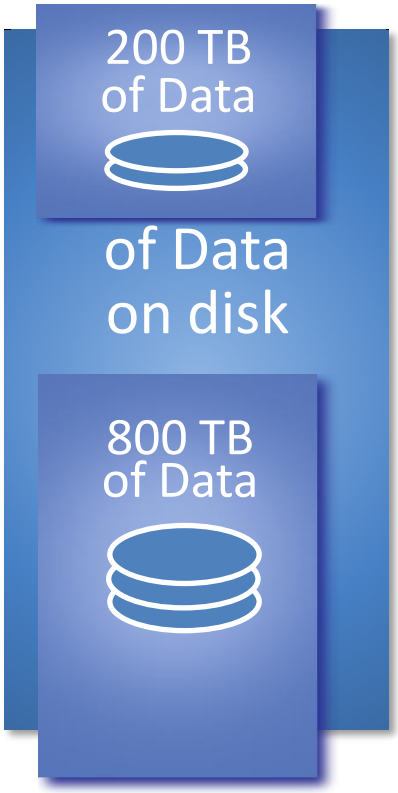
Portable – How do you move 1 PB of data?

- Bandwidth is limited buy the size of the FedEx plane.
  - New Boeing 747-800 Freighter can carry 1,335,900 TB of uncompressed data on LTO5 tapes or 2,315,560 TB of uncompressed data on TS1140 tapes

Before Archiving

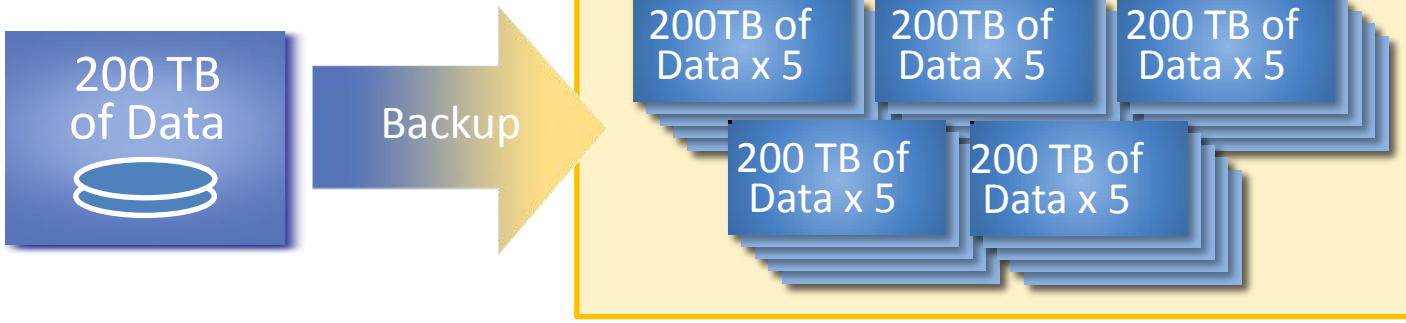


Archive Process



- In this example, we are assuming 80% of the data can be archived.
- By splitting the transactional data from the archive data, each type of data can be treated correctly

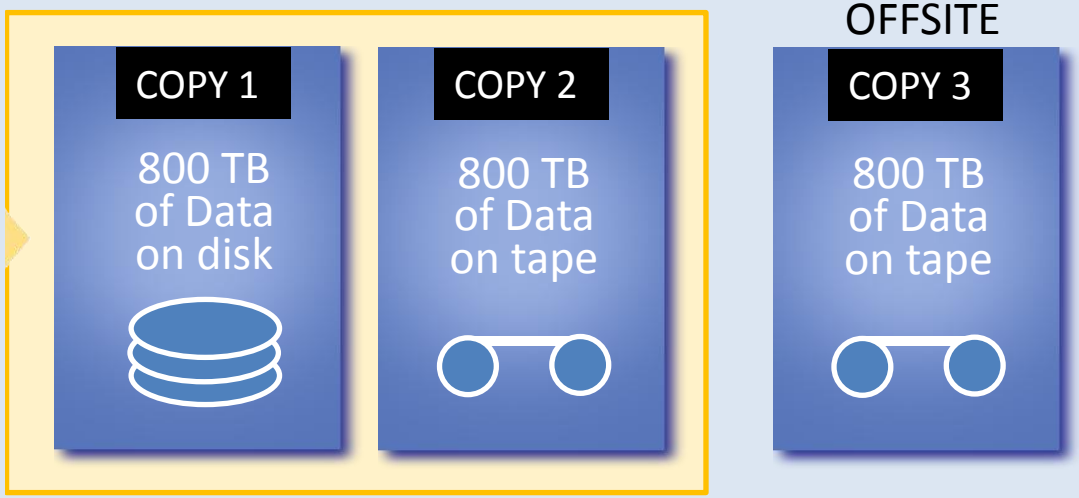
After Archive



200 TB Transactional Data

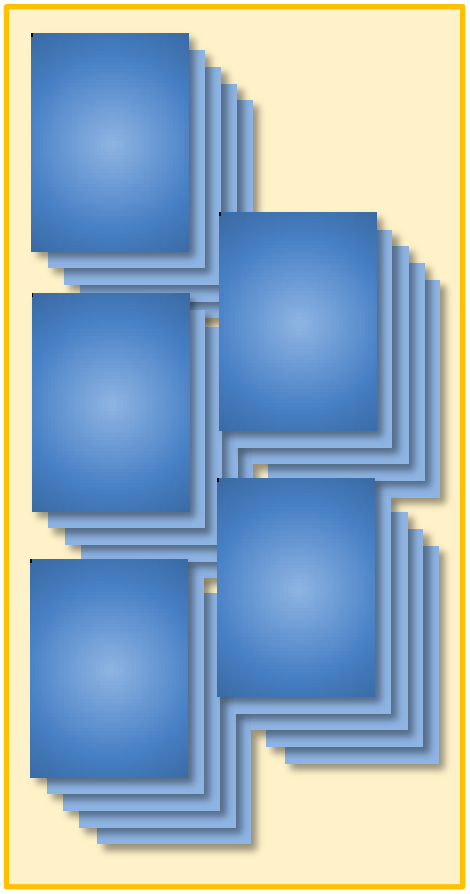
5 PB Backup Data

5.20 PB Total data 1,157 MB/Sec backup



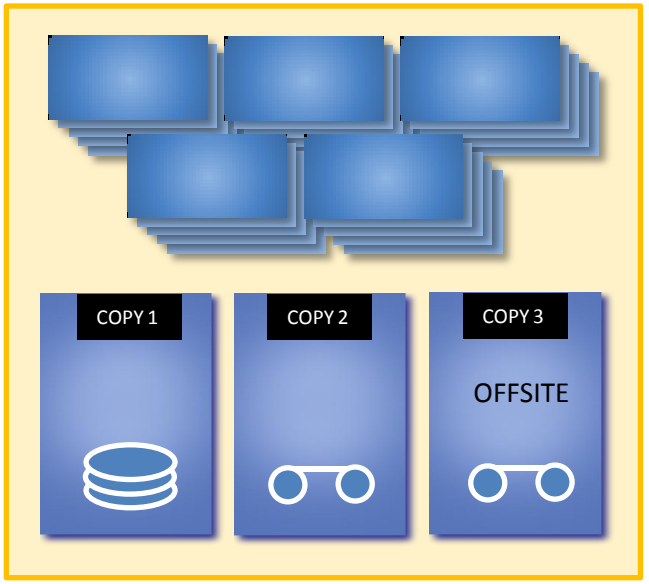
3 total copies of archived data for **2.40 PB** total in archive…Total data managed = **7.60 PB**

Before Archiving and After Archiving



26 PB total

Managed Data



Compared to

7.60 PB total

Managed Data

Active Archive

Active Archive provides an affordable, online solution to access and store all created data.

An archive that contains production data, no matter how old or infrequently accessed, that can still be retrieved online. It may exist on both disk and tape.

The Evolving Role of Tape

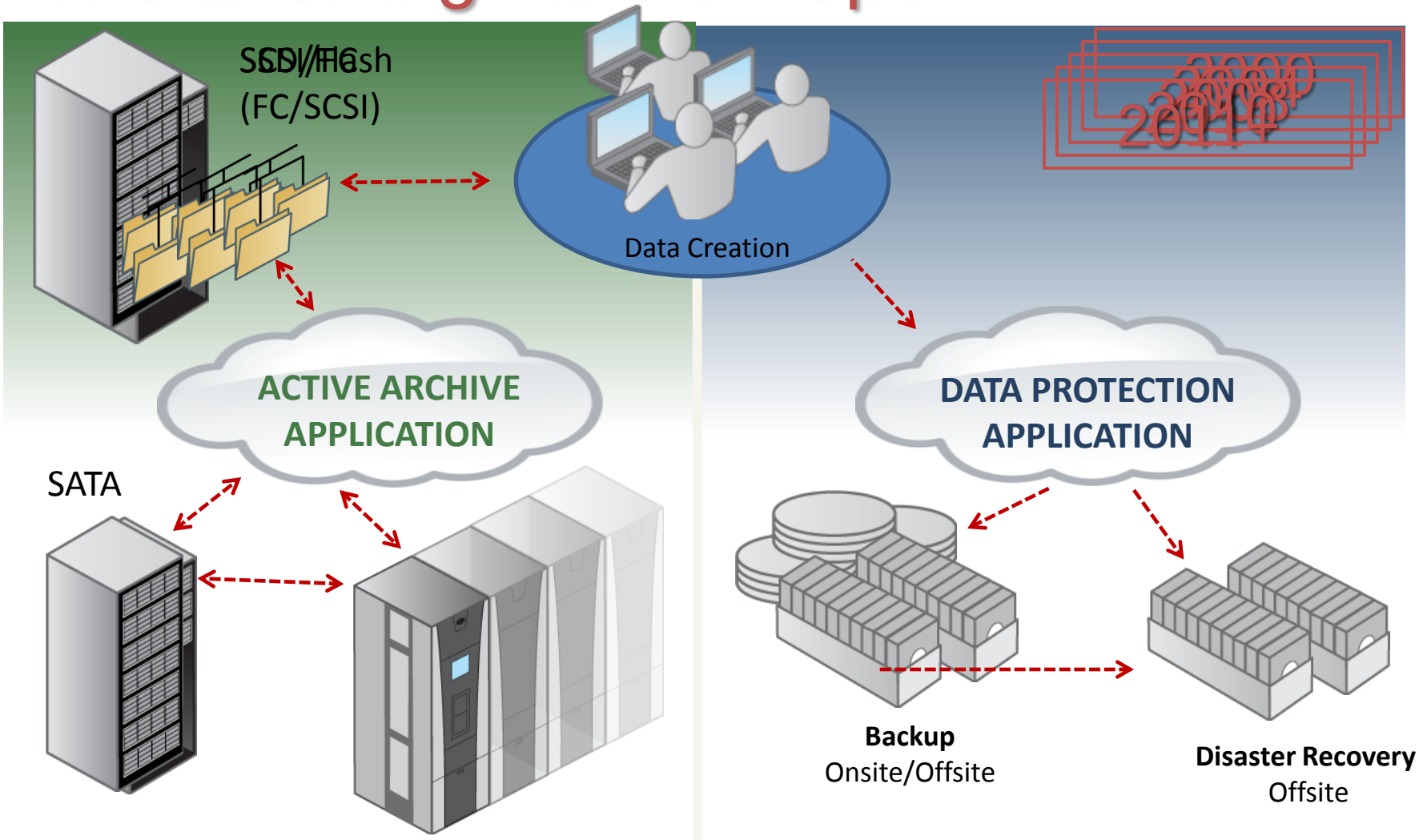
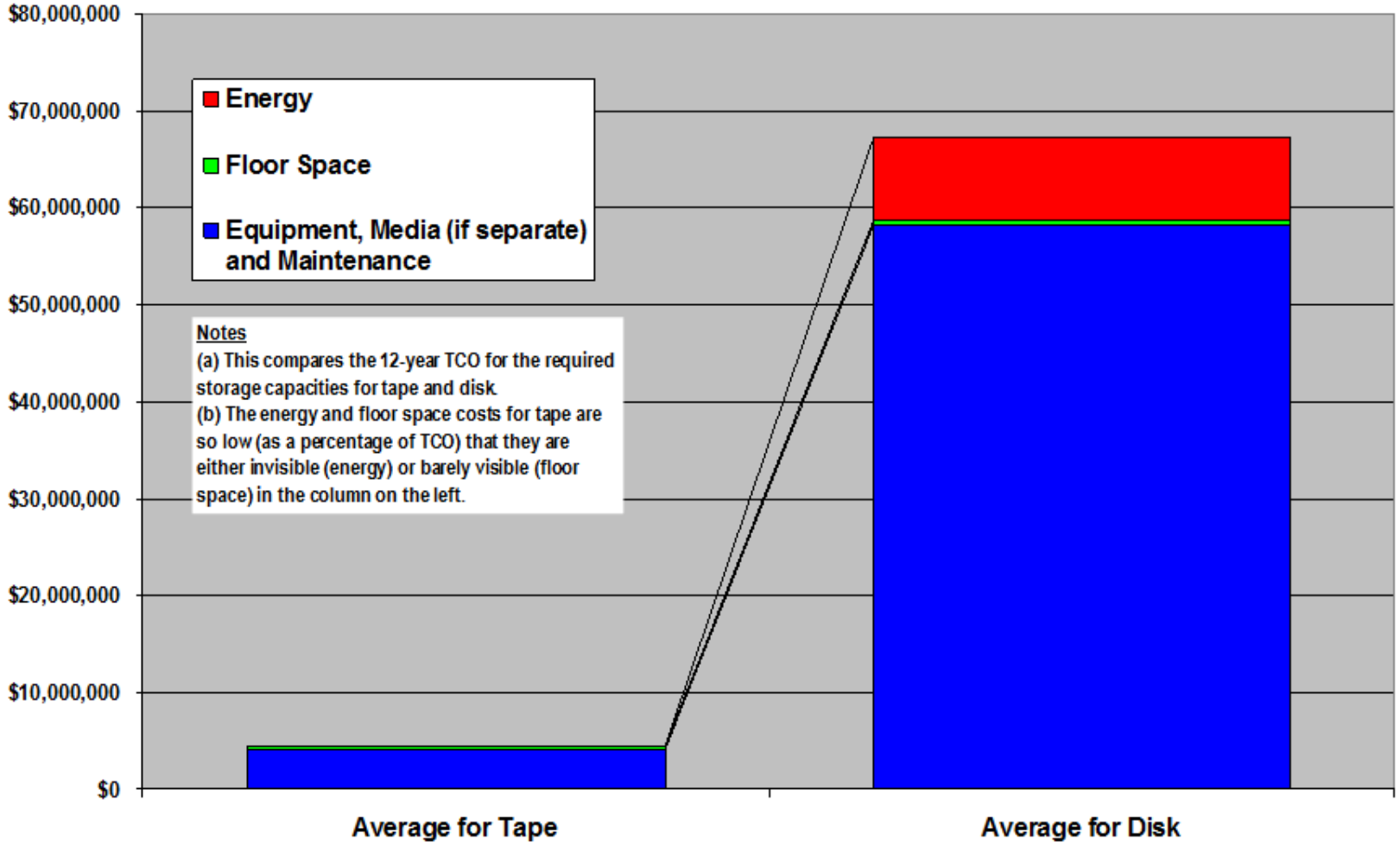


Exhibit 1 —
Comparing 12-Year TCO for Tape to Disk for Long-Term Archived Data

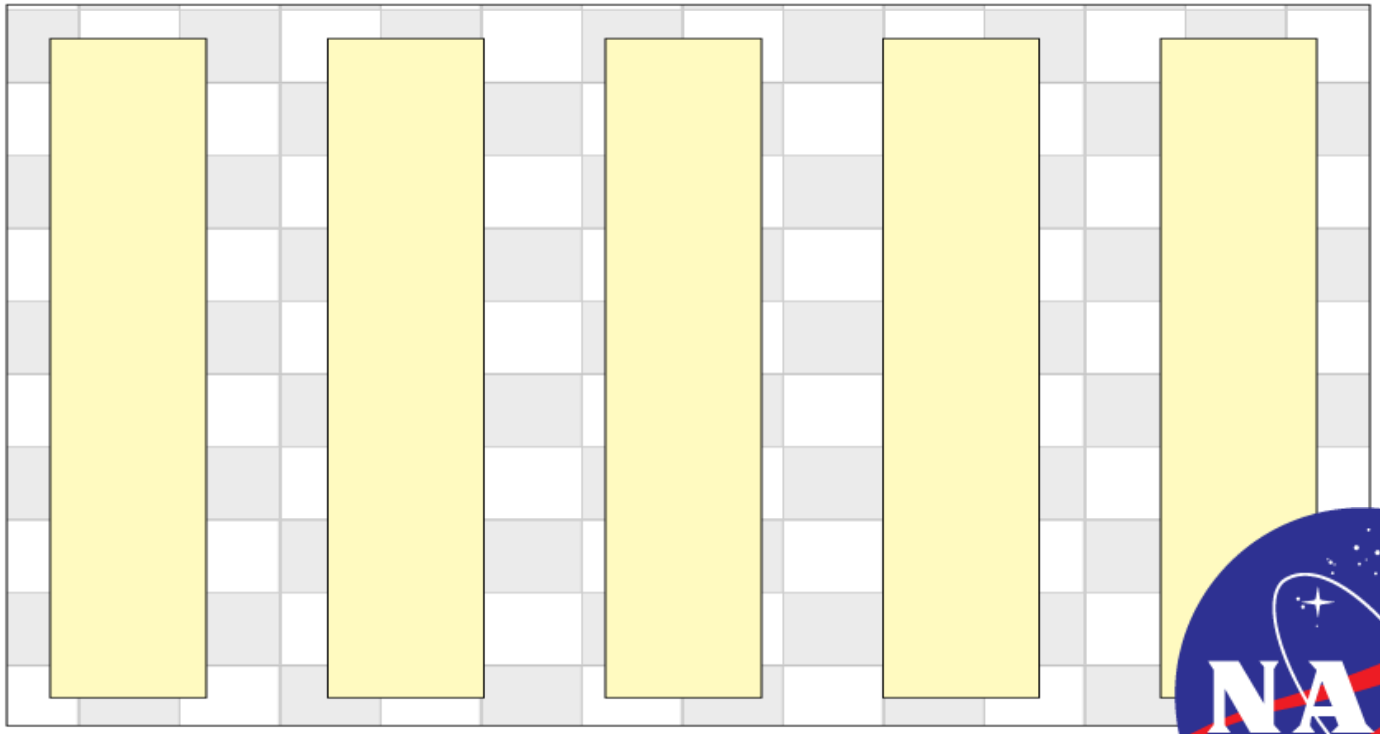
TCO for Disk is Approximately 15 Times Tape Using Clipper's Case Study Model



Source: The Clipper Group

Achieving Efficiency with Active Archive

NASA Ames Case Study



Achieving Efficiency with Active Archive

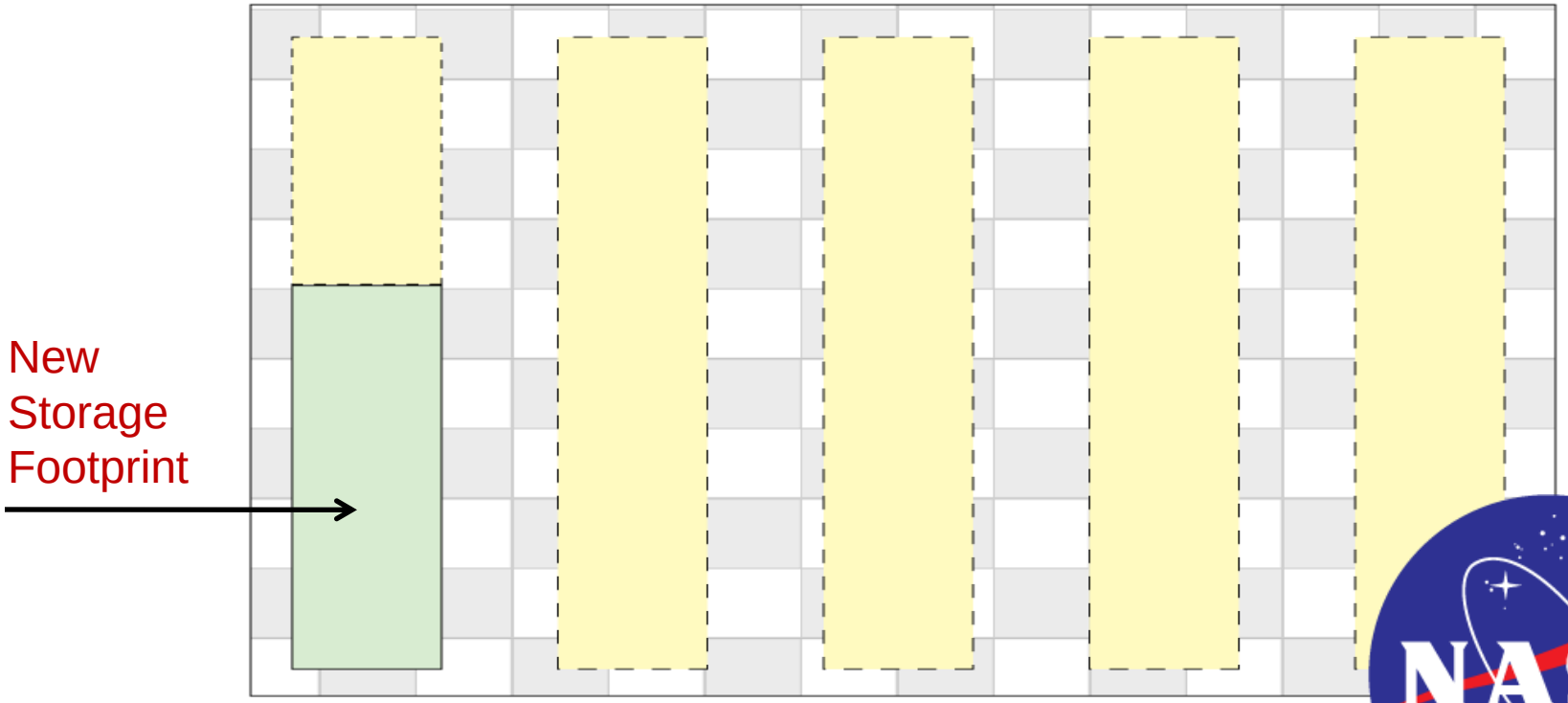
Results—NASA benefits greatly migrating to an Active Archive:

SGI's DMF and Spectra's T950 (8 Frame)

- Extended file system capacity on tape
- Reclaimed 1400 sq. ft. of data center space
- Increased online archive capacity 12 PB to 32 PB
- Increased data storage reliability

Achieving Efficiency with Active Archive

Results

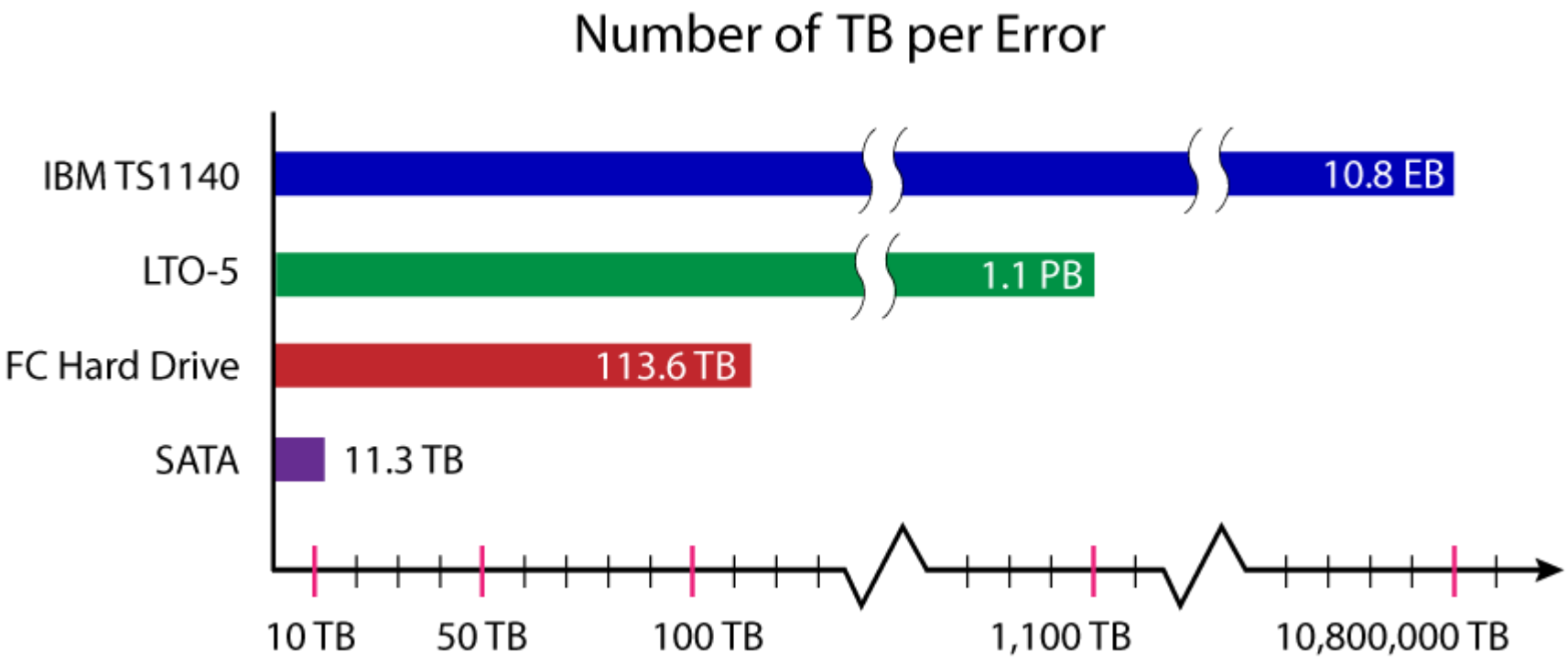


Hard Error Rates

Instrumental
HIGH PERFORMANCE INNOVATION

Device	Hard error rate in bits	Equivalent in bytes	PB equivalent
SATA consumer	1.00E+15	1.25E+14	0.89
SATA Enterprise	1.00E+16	1.25E+15	8.88
Enterprise SAS/FC	1.00E+17	1.25E+16	88.82
LTO	1.00E+18	1.25E+17	888.18
Enterprise Tape	1.00E+20	1.25E+19	88817.84

Device Type	1	10	100	1000	10000
	Device	Devices	Devices	Devices	Devices
	Hours before hard error is reached				
Consumer SATA	4083.9	408.4	40.8	4.1	0.4
Enterprise SATA	31530.5	3153.1	315.3	31.5	3.2
Enterprise SAS/FC	267591.9	26759.2	2675.9	267.6	26.8
LTO-4	2207583.4	220758.3	22075.8	2207.6	220.8
T10000B	220757938.7	22075793.9	2207579.4	220757.9	22075.8


Number of TB per Error
IBM TS1140
10.8 EB
LTO-5
1.1 PB
FC Hard Drive
113.6 TB
SATA
11.3 TB
10 TB
50 TB
100 TB
1,100 TB
10,800,000 TB

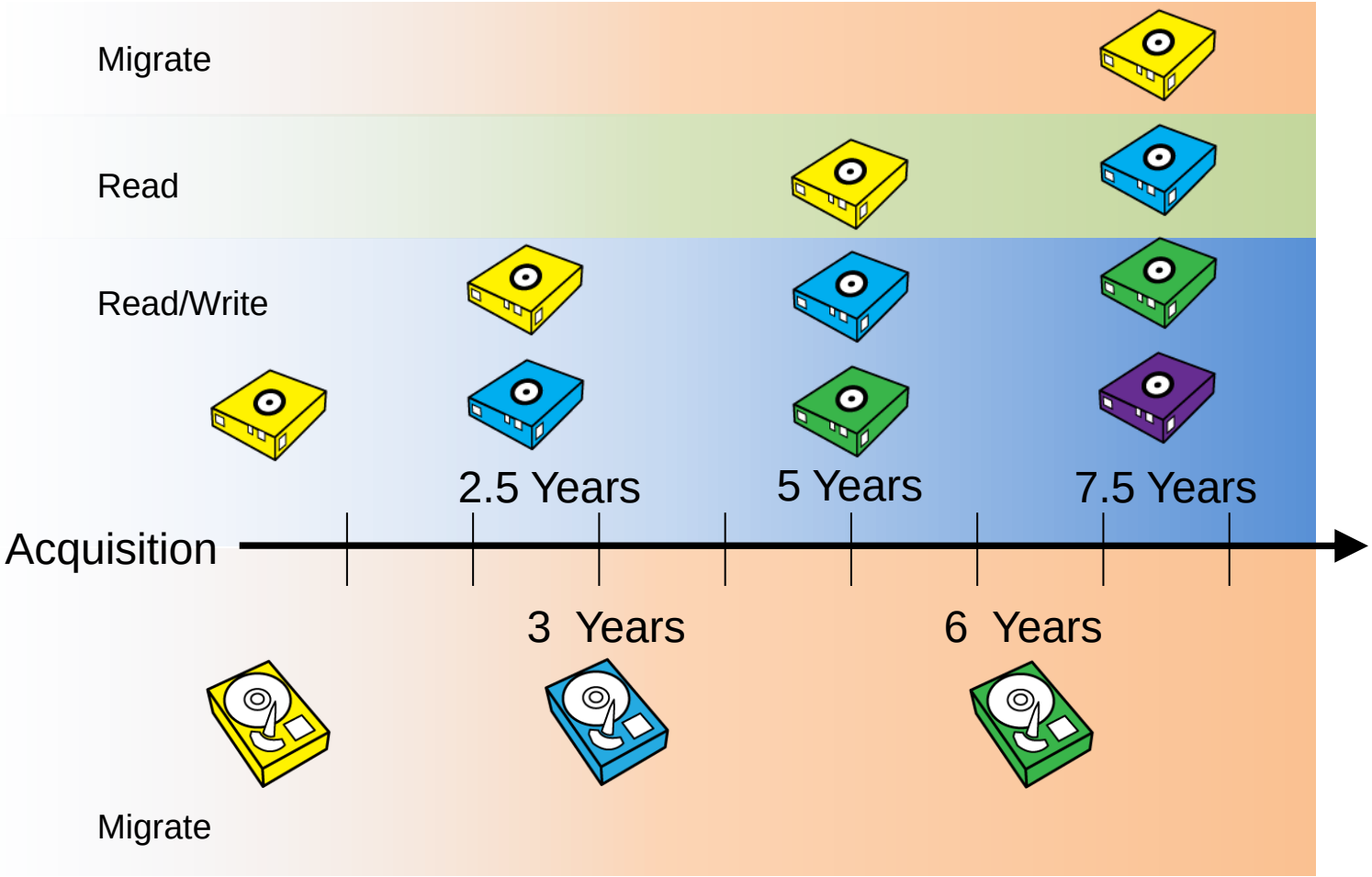
Securing Data – from theft and corruption

1. AES-256 bit encryption build into every enterprise and LTO tape drive
2. Offline data copy
3. Integrated data integrity verification

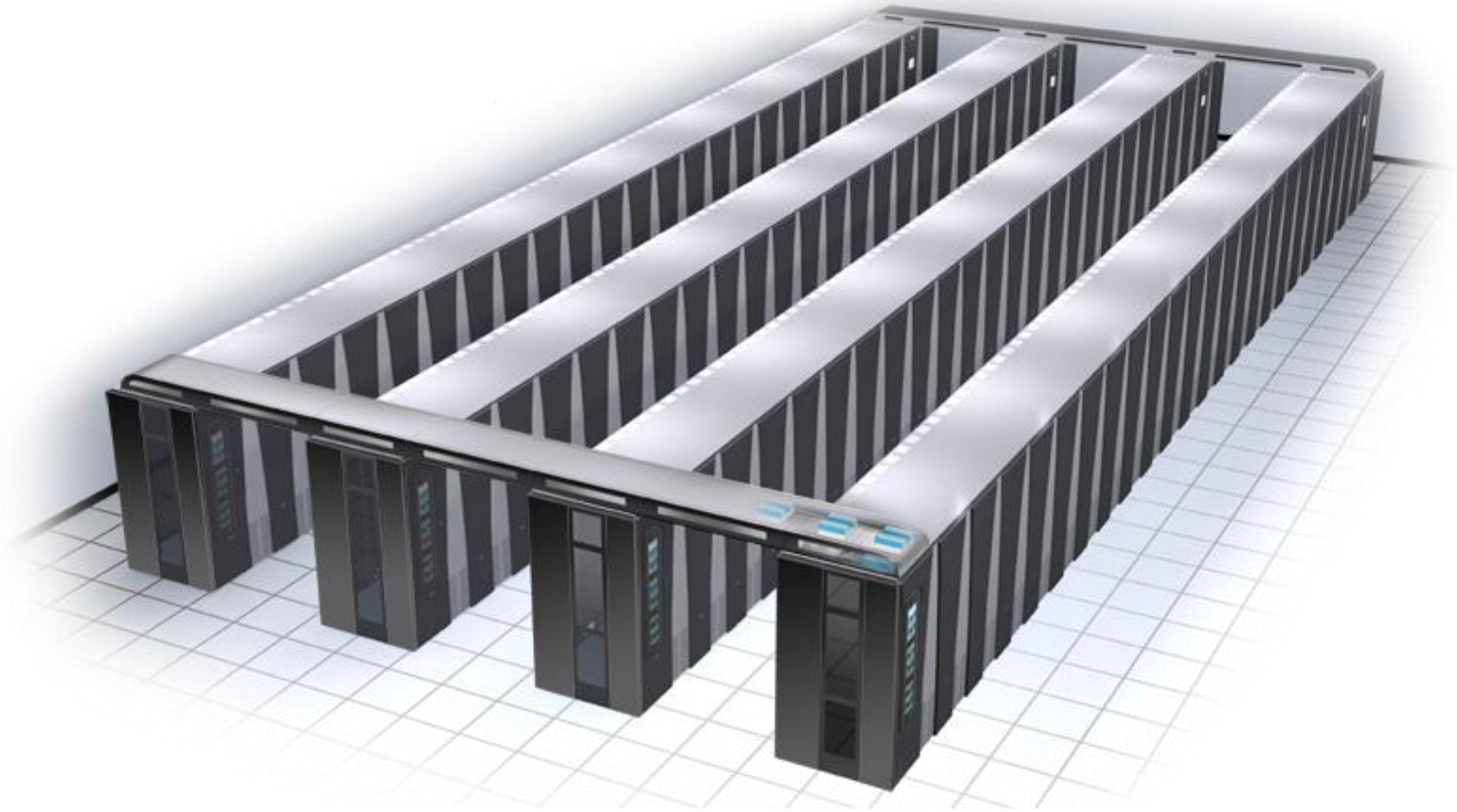
Tape has a proven lifespan

- Data has weight – large data sets present a significant migration challenge
- A modern tape library will be on the Data Center floor over a decade
- Look at LTO tape
  - Write back 1 Generation
  - Read back 2 Generations
  - If you want to stay current, you can migrate every 7 ½ years to new media in the same library, vs moving data to a new disk array every 3 – 4 years.

Tape has a great lifespan



…up to 740 PB in a single unit



Data Centric Organizations Choose Tape

NCSA

Discovery COMMUNICATIONS™

NATIONAL GEOGRAPHIC

NASCAR MEDIA GROUP

NASA

FMI

UNIVERSITY OF MINNESOTA

National Energy Research Scientific Computing Center… A fit for tape

At NERSC :

- There are 13PB of data on tape
- 30 – 40% of its tape's activity are reads
- It has a measured and proven reliability of 99.945%
- The tape cost is around 5% of that of its disk storage

This is *Primary* data with no secondary Copy!

Technical Brief – NERSC: Proving Tape as Cost-Effective and Reliable Primary Data Storage – ESG, Mark Peters, Senior Analyst December 2010

Growing and Profitable

- 34% revenue growth year over year
- 55% increase in mid-level tape shipments
- 100+% increase in enterprise tape shipments
- 10% revenue invested in R&D
- Profitable 5 years running &

  100% debt-free

Questions?

mollyr@spectralogic.com

303-449-6444 ext. 1358

VP of Product Management