Spectra Logic

Matt Starr - CTO

About me

- CTO at Spectra Logic
- BSEE in VLSI design
- 25+ years of computer and storage experience
- 43 US patents, 63 applications under file
- 23rd year at Spectra Logic
- Architected and Designed disk and tape subsystems

Spectra Logic: Leading Tape Innovation Since 1979

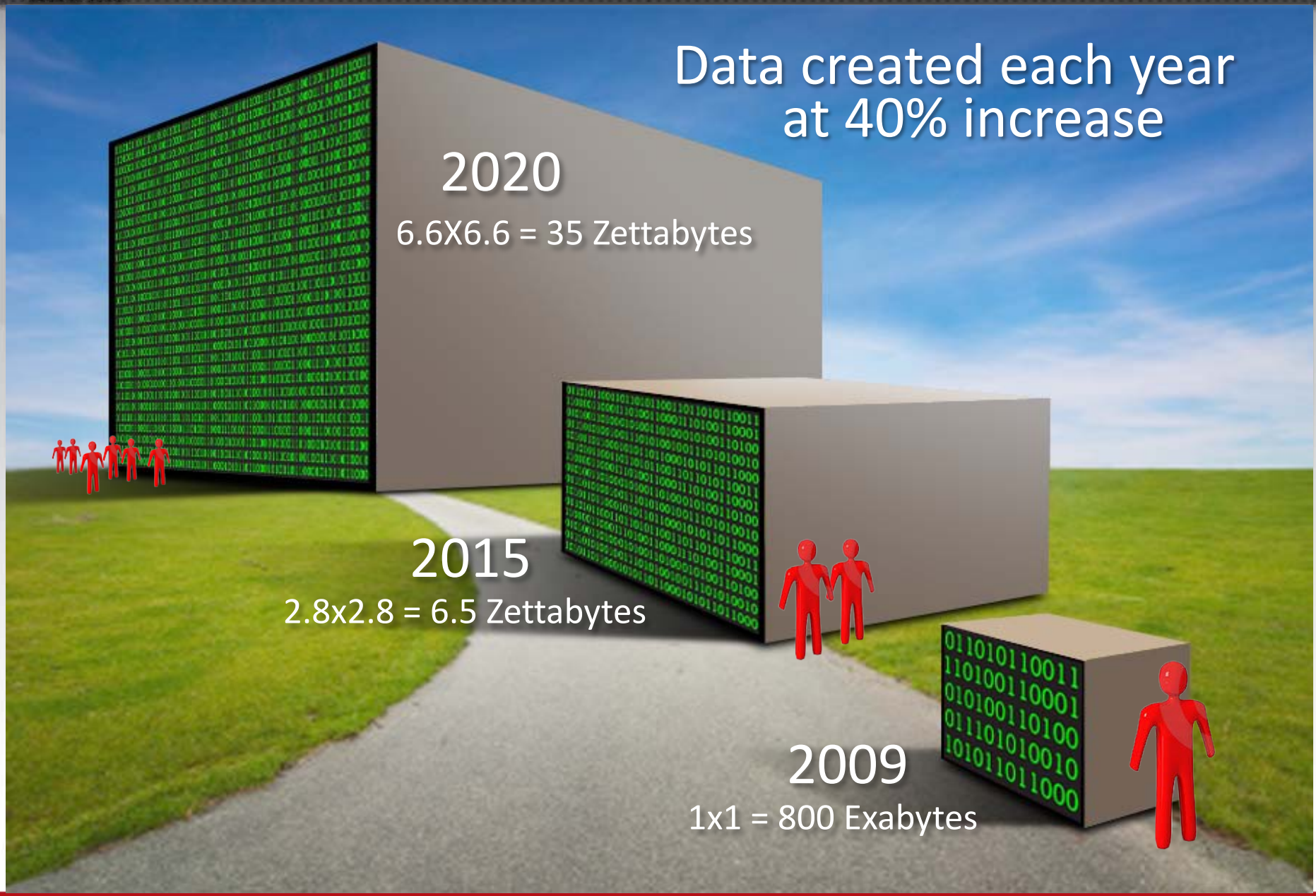
- Proven Innovator and History of Success
  - Intelligent integration of complete data protection solutions
  - Founded in 1979, self-funded, profitable, debt-free growth
  - Continuous innovation
  - High customer satisfaction & support ratings
- Long-term Market Traction and Growth
  - Years of Yr/Yr growth in enterprise and mid-range tape libraries;
    media and support services
  - Focus on Tape Automation and Archive Grade Storage
  - Leader in data intensive verticals: HPC, M&E, Federal

Data Growth

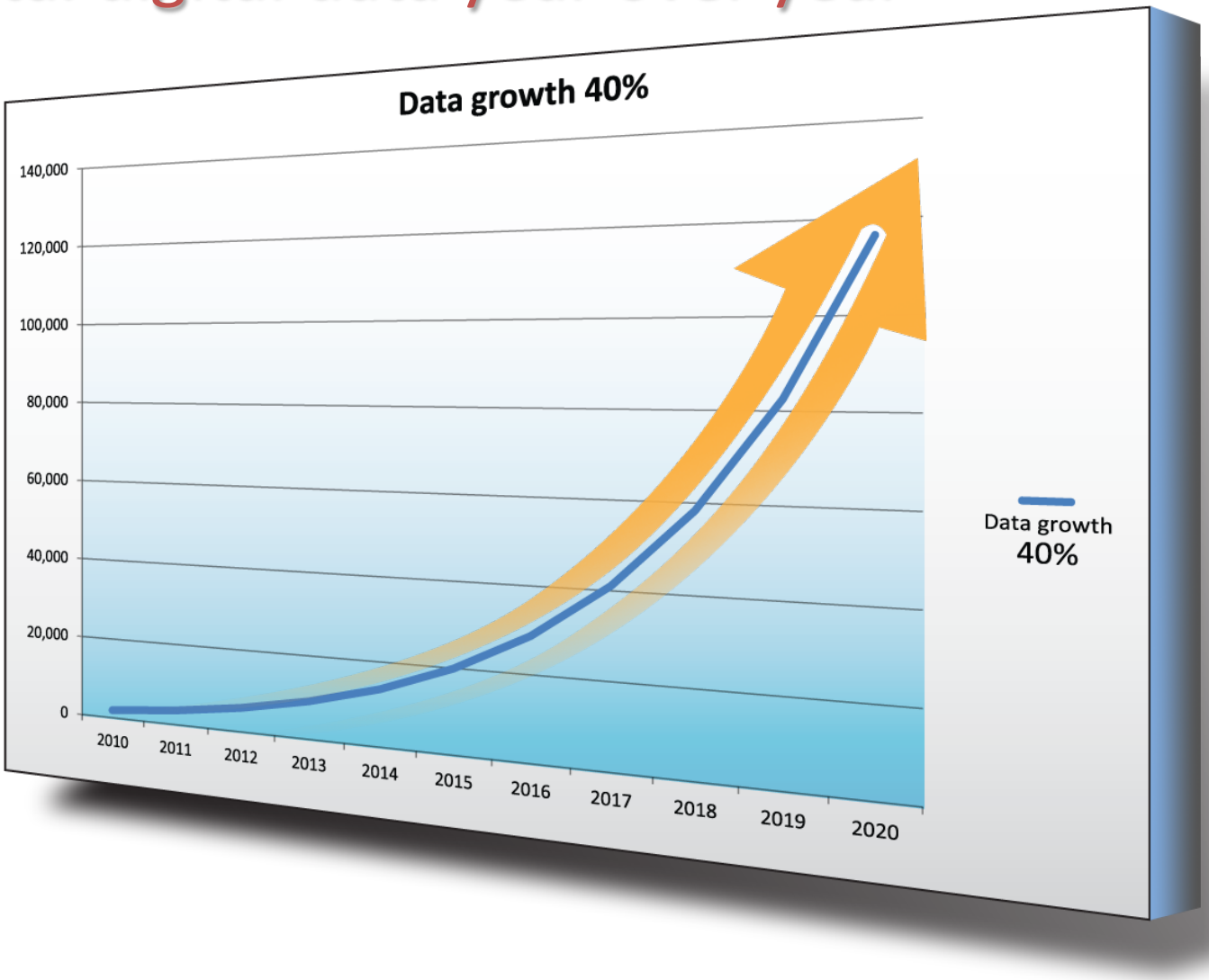
Data Growth over the next few years

- 281 Exabytes of data under management in 2008
- 161 Exabytes of data was created year 2008
- 800 Exabytes in 2009
- 1.2 Zettabytes were created in 2010
- Data rates are growing at 40-60% per year
- 1.8 Zettabytes expected by 2011
- We will create 6.5 Zettabytes in 2015
- Two areas
  - Social media and Digital content

Source: IDC 2007/2008 – IDC 2010


Data created each year
at 40% increase
2020
6.6X6.6 = 35 Zettabytes
2015
2.8x2.8 = 6.5 Zettabytes
2009
1x1 = 800 Exabytes

Total digital data year over year



What to do?

Archive your data!

Active Archive Defined:

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 indexed, searched and retrieved online.

By extending a file system across high performing disk, capacity disk and now tape, the need for IT intervention to retrieve and an archived file is minimized, if not eliminated.

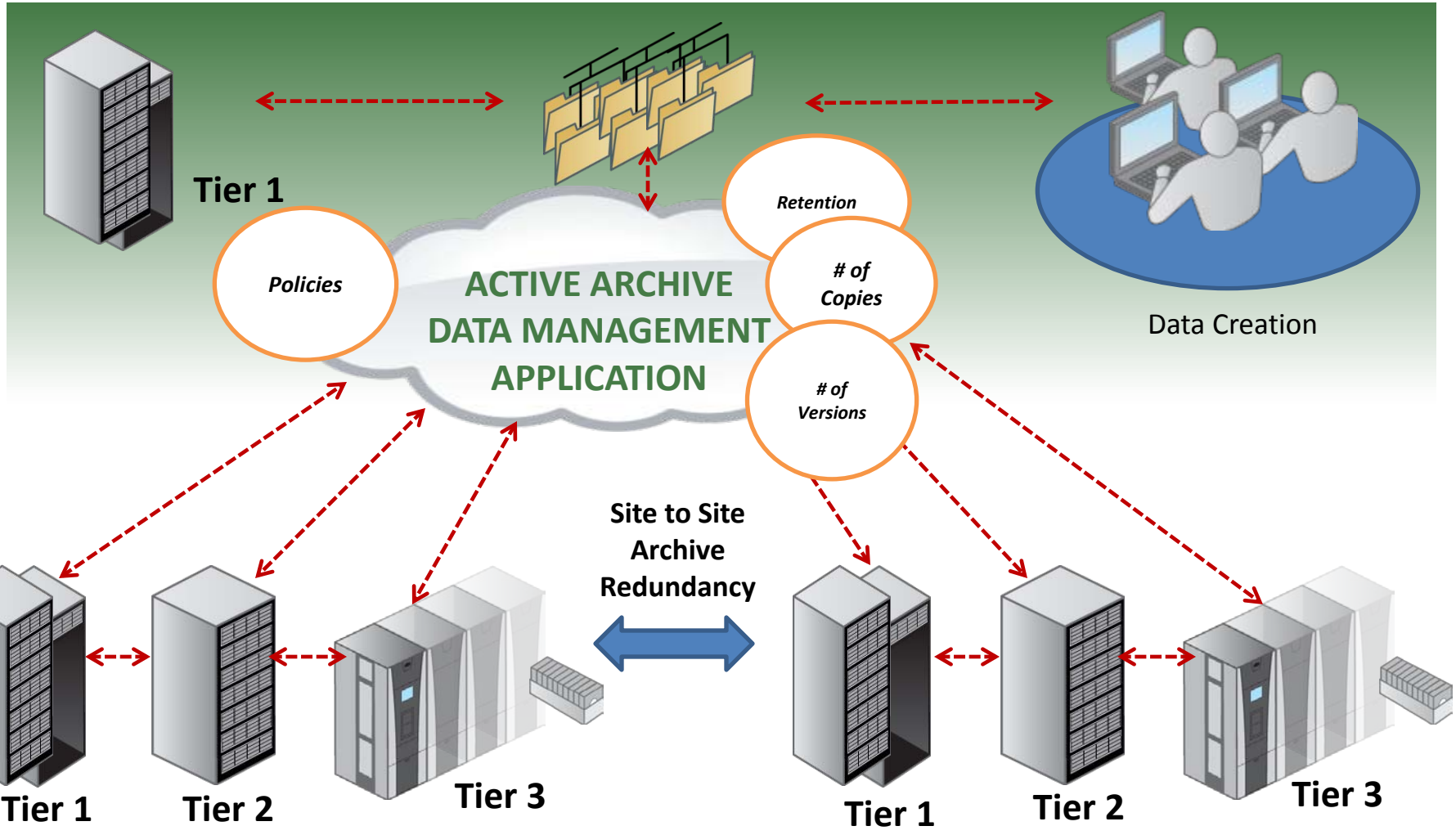
Introducing the Active Archive Concept

- An Active Archive contains **native file format data transparently accessible to end users through a file system interface (CIFS, NFS).**

- Active Archiving is not a single product. It's a **collaborative solution** offered by software and multiple hardware vendors, and can also take advantage of existing equipment.

- **Vendor Agnostic-** Consisting of data management software, disk, and tape options.

The Concept of an Active Archive



Active Archive Primary Instance

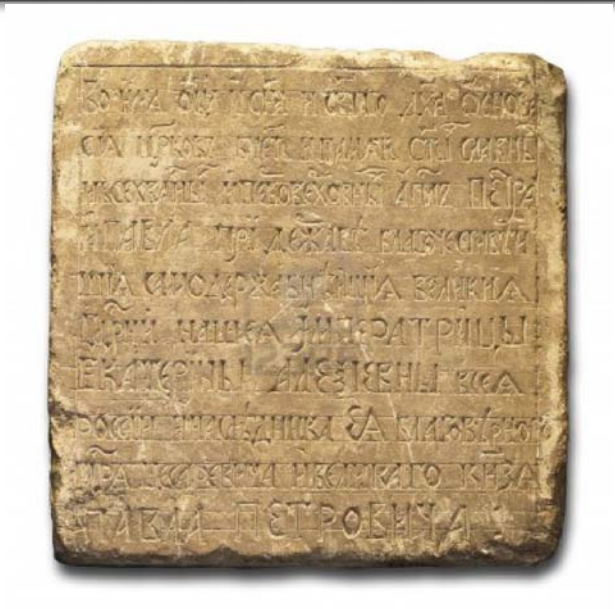
Active Archive Secondary Instance

Best practices

Where you can lose data

- Humans
  - The number one cause of data destruction
- Hardware / Firmware / Software failure
  - Firmware is usually not listed in ECC
- File Format
  - Application or software to use data
- Data format
  - Operating system, tape or disk format

Avoid extinction



- Power
  - Water wheel, Steam…
- Language
  - Mohican, Tillamook…
- Humans
  - Homo Erectus, Homo Habilis, Homo Neanderthal
- Storage
  - Punch Cards, IBM 3340 Winchester, Cipher M990, FW SCSI
- Software
  - Dbase, CP/M, Quattro

Best practices

- Finger print data at time of acquisition
  - Create a CRC, SHA-x…of data
- Index and Categorize on ingest
  - Create meta data at creation, you'll never go back and do it
- Two or more copies
  - Apollo 11, D-Day landings…
- Two Genome System
  - Avoid firmware issues destroying data
- Migrate forward
  - Move off older technology, check data with CRC
- Tectonic separation
  - Move data across country/world

Lifecycle Management Reduces Your Risk

- Improve system and data availability
- Proactively avoid problems before they occur
  - Media Lifecycle Management (MLM)
    - Prevents tape-related failures
    - Alerts you when tapes should be replaced and data re-mastered
  - Drive Lifecycle Management (DLM)
    - Always on drive reporting and trouble-shooting
    - Reduces drive error risks during critical back-ups
  - Library Lifecycle Management (LLM)
    - Pro-active notification of upcoming service events
    - Reporting on hardware utilization and issues

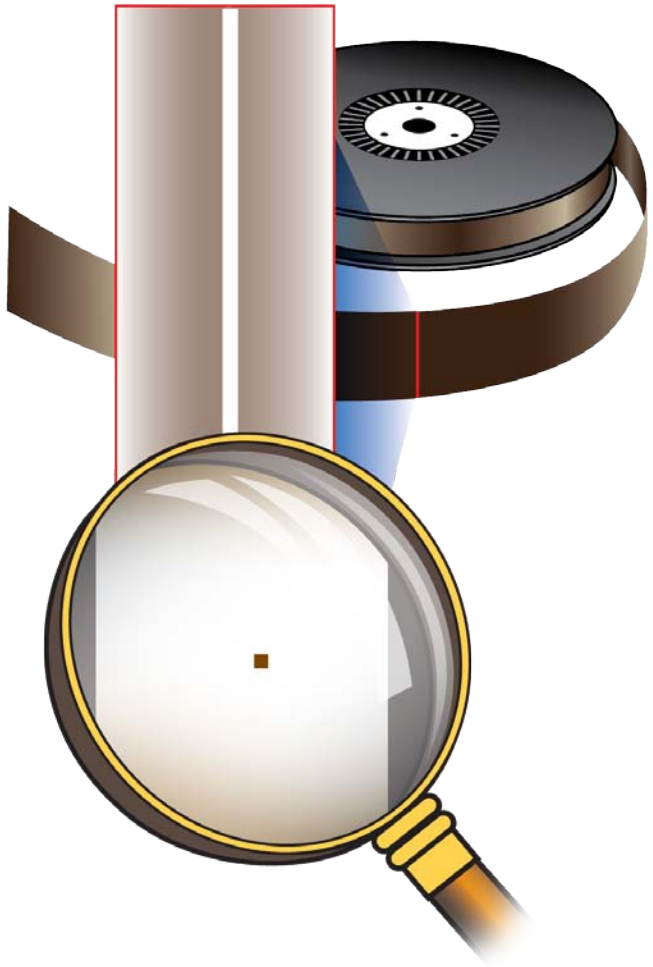
POWERED BY BlueScale INTELLIGENCE

Why nearly 80% of the world's data is on tape?

Fred Moore

Bit density, now and future (growth trend)

- HDD (slowing)
  - 81 nm x 15 nm
  - 530 Gbit/in^2
- Patterned Media 2015/16
  - 25 nm x 25 nm
  - 1000 Gbit/in^2
- NAND FLASH (Steady)
  - 45 nm x 45 nm
  - 310 Gbit/in^2
- Current Tape (Growing)
  - 8000 nm x 67 nm Current Tape
  - Next Gen 446 nm x 49 nm - 35 TB per tape

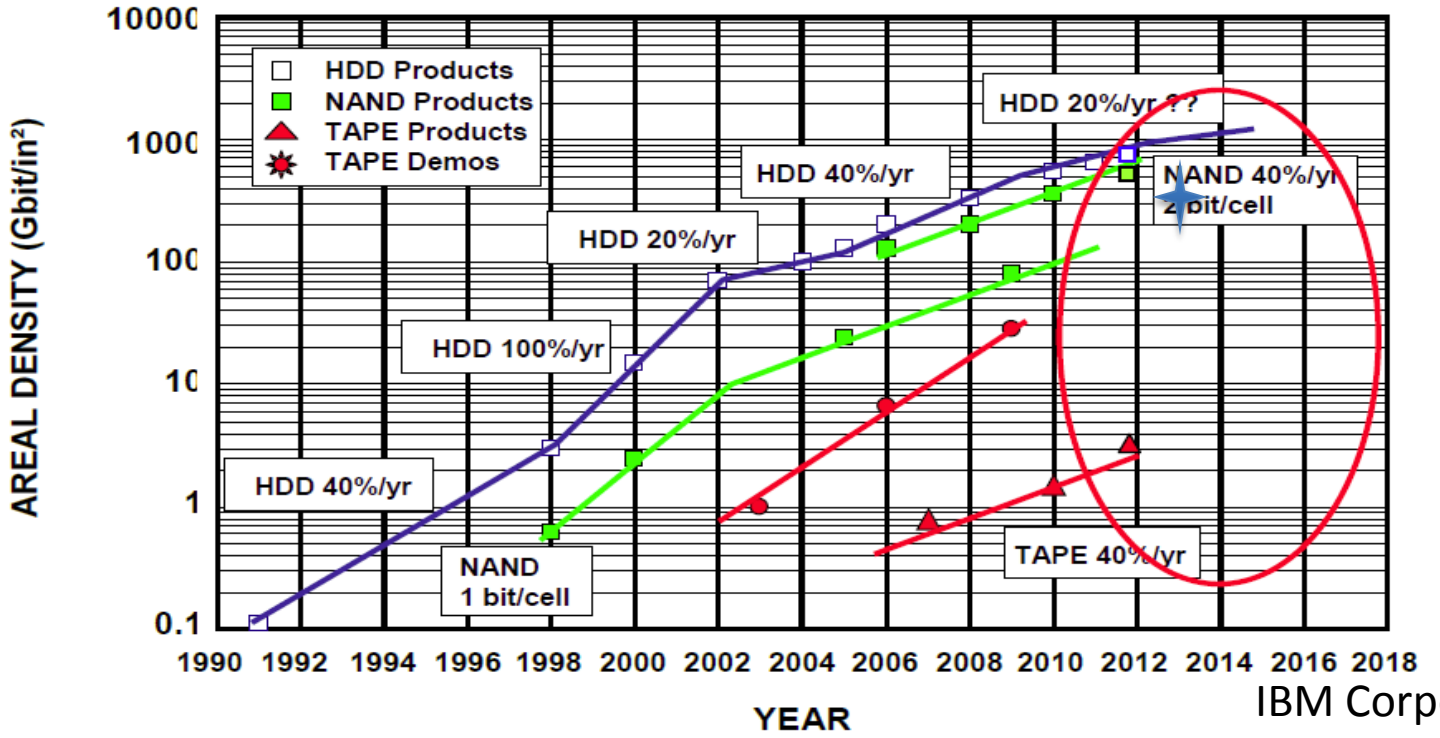


Future bit density

IBM

Storage Device Density Landscape – A Summary

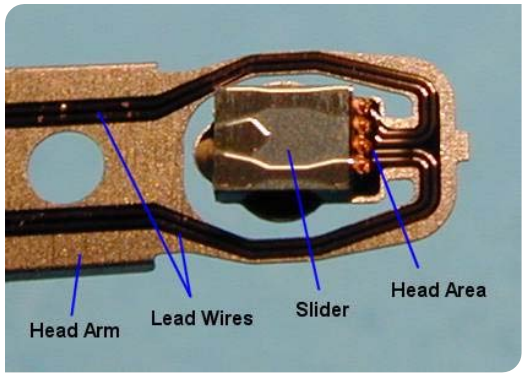
- **HDD** → 20% to 25% annual density increases
- **NAND** → 25% to 30% annual density increases
- **TAPE** → 40% annual areal density increases; likely greater (80%??)

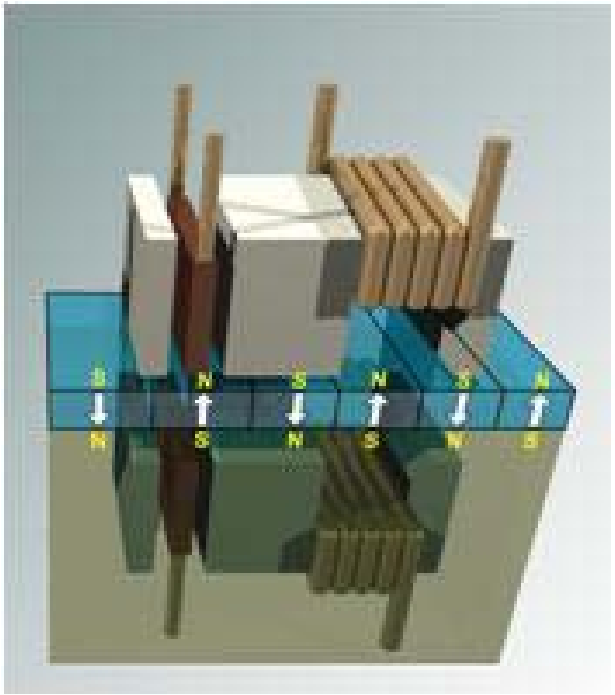


IBM Corporation

Tape Leverages Disk Technology

- Disk Drive Head Technology
  - GMR disk drives were release in 1993
  - CMR, HAMR, SMR, Pattern – new technologies
- Tape Drive Head
  - LTO GMR in 2008
  - Tape is 2 – 4 head generations behind disk head technology
- Disk Capacity current hurdles
  - Adding platters back in to keep capacity curve, 7 to 8 may be needed.
  - Helium filled drives
  - Disk throughput not keeping up, shingling will slow this down.
- Recording area
  - HD < 100 sqin, ~12.5 per platter face
  - Tape >14,000 sqin





Bit Error Rate

- Bit Error Rate (BER)
  - SATA disk 1:10^14
  - SAS/FC disk is 1:10^15
  - LTO-5 tape 1:10^17.
  - IBM TS1140 tape 1:10^20.

- Tape is more reliable than disk…
  - Tape has 1,000 to 1,000,000 times better in error correction!

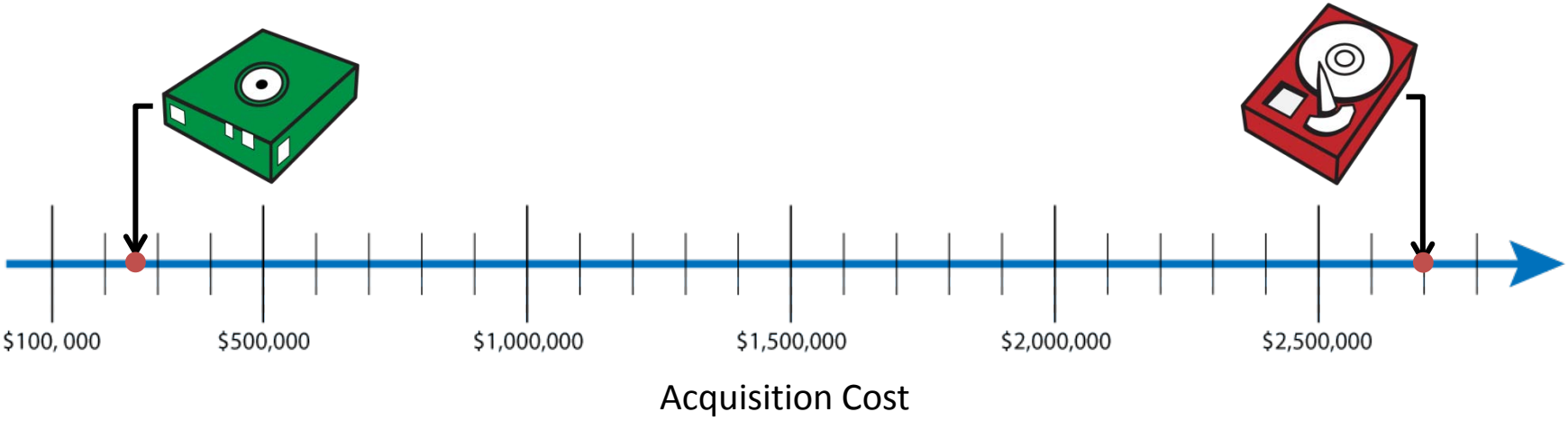
- While 10^15 may look really close to 10^17, it's not!
  - The difference in bits 113 TB and 11.1 PB of data! It means you are 100 times more likely to have bad data on disk than you are on an LTO-5 tape drive, and 10,000 times more likely than if the data is stored on TS1140 drive!

- I have RAID…
  - A two disk fault will happen every 300 TB on SATA and 3 PB on SAS!

Curtis Preston 3/12

Tape Has Lowest Acquisition Costs

- 10 to 15 times less expensive than disk
  - 2.7PB tape system has street price of $0.07 to $0.10 per GB. Larger systems are even lower.
  - IT grade has street price of ~ $1.00 GB; enterprise class disk is more expensive.

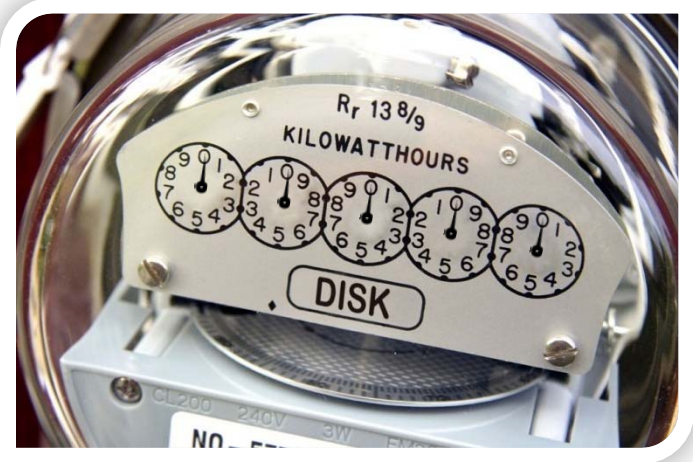


Tape Has Lowest Operating Costs

Operating costs are low and predictable

- Lower purchase means lower ongoing support costs
- Tape's power consumption is roughly 4% disks

Source: Clipper Group Study, "Tape Remains King", based on an initial 30 terabytes of data



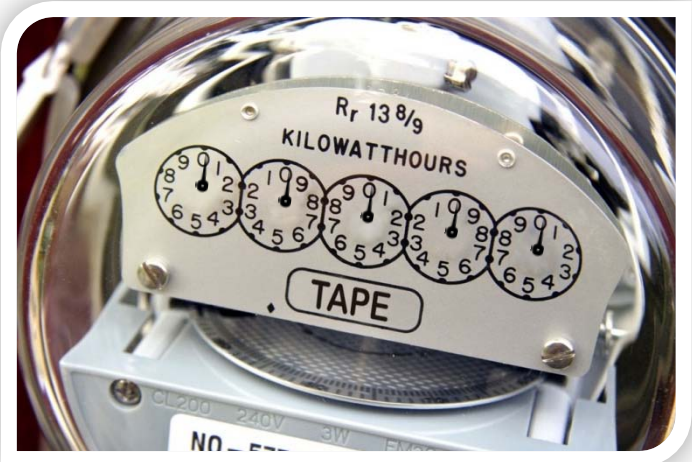
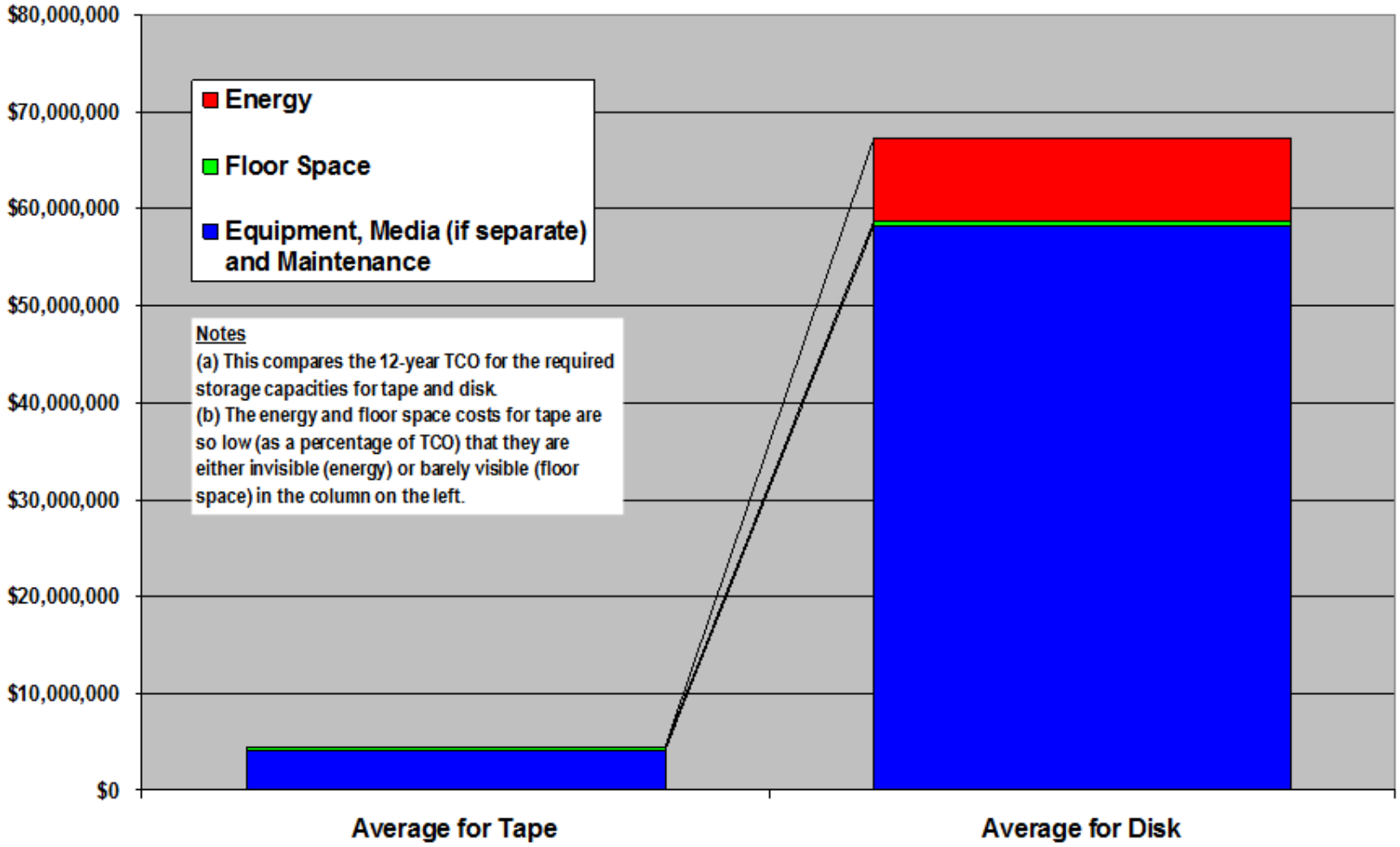


Exhibit 1 —
Comparing 12-Year TCO for Tape to Disk for Long-Term Archived Data

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



Source: The Clipper Group

Tape Offers the Best Power Efficiency

Spectra Logic T950 Enterprise Tape Archive

T950: The right size at the right price

- 1 to 8 Frames
- 2 to 120 Drives
- 100 to 10,050 slots (over 30 PB* of storage)
- Hi-speed robot accesses every drive and tape
- Buy ONLY what you need: Capacity on Demand in 10 slot increments
- Fast, easy expansion

Compact, rack-row design

- Improves hot / cold aisle air flow
- Makes service easy
- Saves time and money.

*LTO5 with 2:1 compression

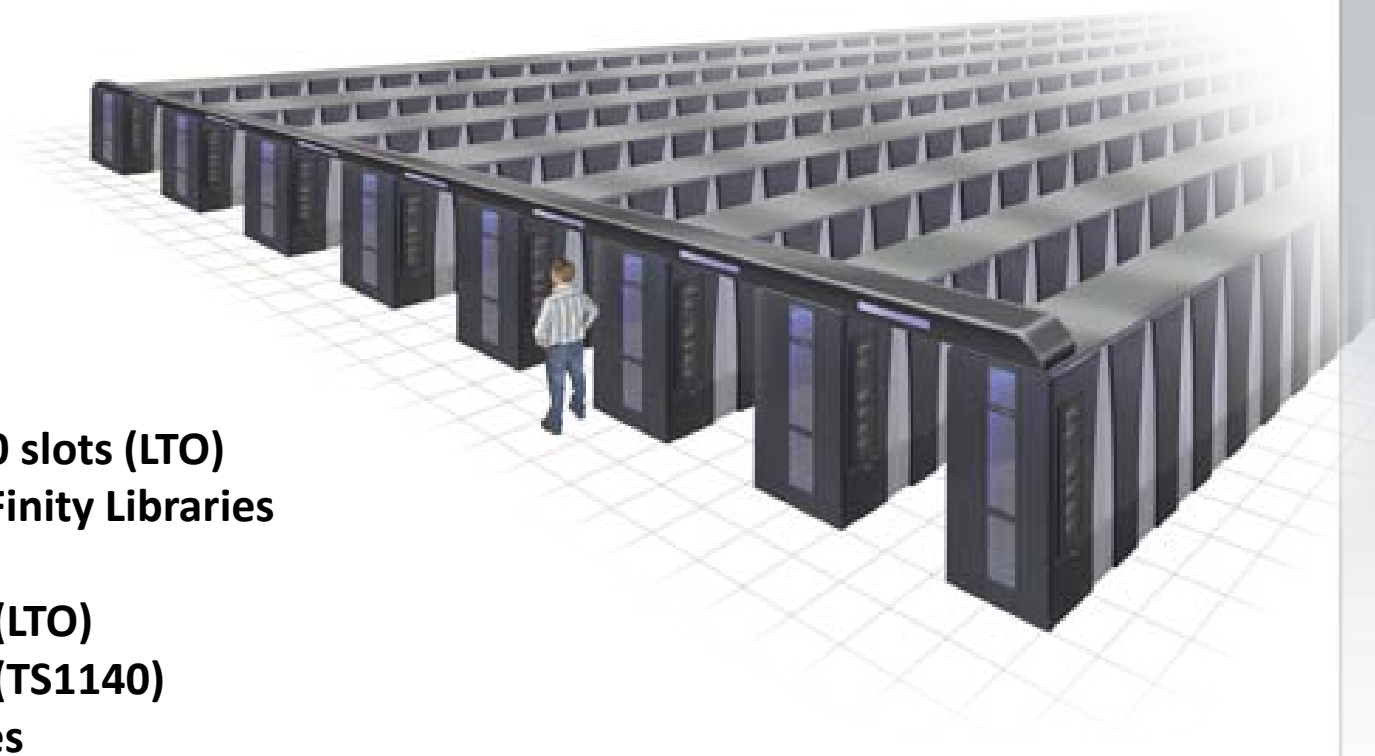
T-Series Scalability - Lowers Your Cost

The right size at the right price

- 1 Drive to 960 Drives
- 10 Slots to 400,800 slots (over 3 EB* of storage)
- Capacity on Demand – buy ONLY what you need
- TranScale – Investment protection

* With LTO-6 or TS1140 Technology (available now).

T-Finity Skyway Future Library Complex



- **Single Library 50,000 slots (LTO)**
- **8 Interconnected T-Finity Libraries**
- **320 Library frames**
- **Up to 400,800 slots (LTO)**
- **Up to 304,920 slots (TS1140)**
- **Up to 960 tape drives**
- **Up to 16 robots**
- **Up to 16 Bulk TAP service frames**

Tape archives

- Lowest power cost
- Lowest acquisition per TB
- Highly reliable
- Portable
- Air Gap (off site)
- Tape is the ideal storage media for long term archive.
- ~80% of the world's data is on tape.

Questions