SPECTRA

FUJIFILM

Fujifilm's 6th Annual Global IT Executive Summit

Deep Storage for Exponential Data

Nathan Thompson

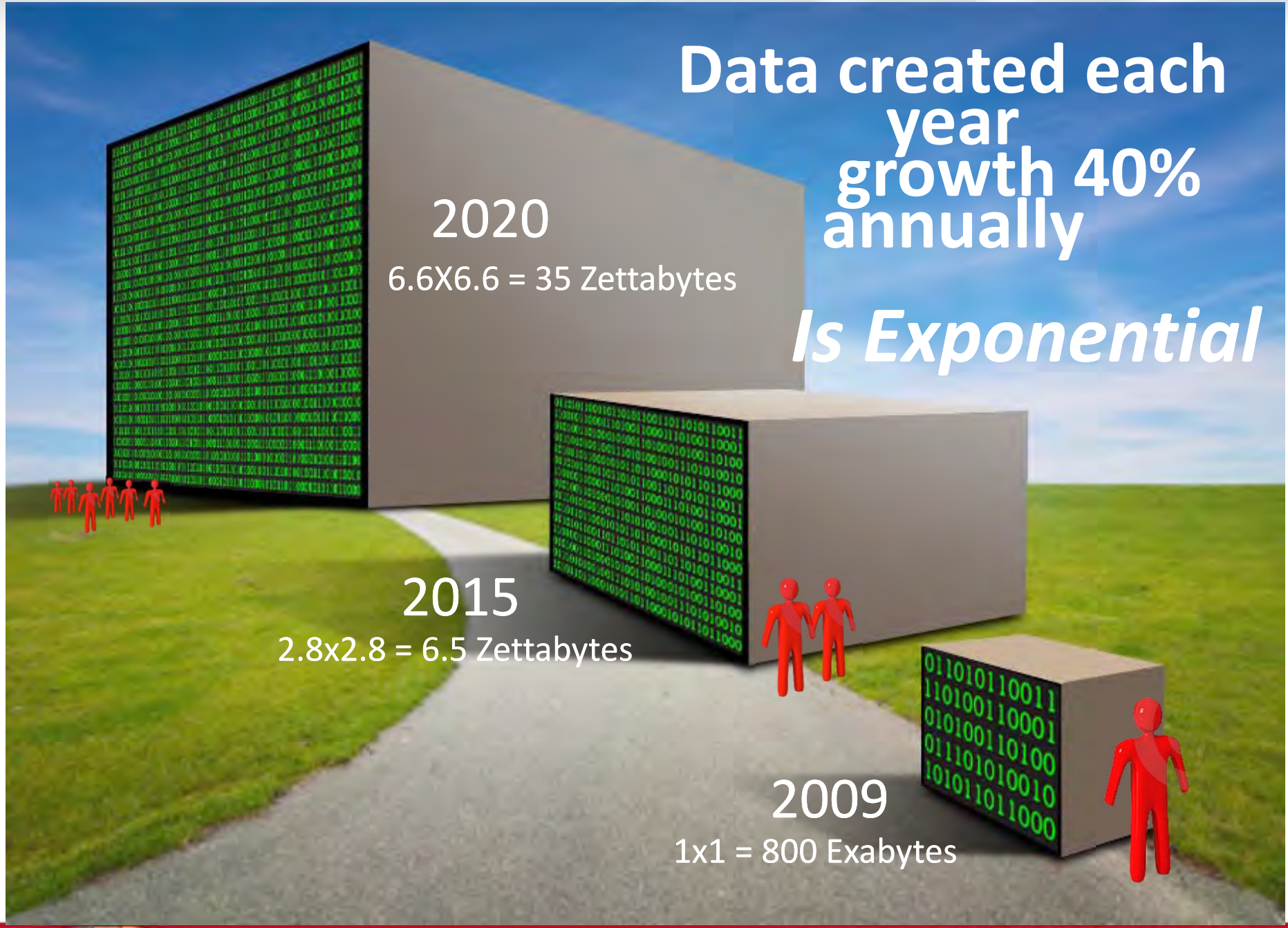
CEO, Spectra Logic

DEEP STORAGE EXPERTS

HISTORY

- Partnered with Fujifilm on a variety of projects
- HQ in Boulder, 35 years of business
- Customers in 54 countries

- Spectra builds tape libraries
  - LTO and TS Libraries
  - nTier Verde low cost high reliability disk arrays
  - BlackPearl DS3/S3 Appliance


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

What is "Exponential Data"?

- By 2020 the Digital Universe will hold 40-85 Zettabytes*
  - 1 Zettabyte = 1000 Exabytes or 10^{21} bytes
  - 1 Exabyte = 1000 Petabytes or 10^{18} bytes
  - 1 Petabyte = 1000 Terabytes or 10^{15} bytes
  - 1 Terabytes = 1 Million * 1 Million bytes

Who here today is working with a Petabyte of data?

- Clayton Christensen wrote about Disruptive Technology in his book, "*The Innovator's Dilemma*"

- *Spectra sees a "disruptive" usage of tape for storing much of the "Digital Universe"*

*Source: IDC, The Digital Universe in 2020

Tape Storage Usage

Backup

Archive

ACTIVE Archive

Deep Storage:

A Public/Private cloud interface to Tape Storage

State-Of-Art Storage Media*

400GB Flash Drive (Seagate)**

100GB Blu-Ray rewritable disk (Verbatim)

6TB Enterprise grade disk drive (Seagate)

10TB Enterprise Tape Cartridge (Fuji/IBM)

*Highest capacity shipping now or within 1 month. **2TB flash drives now available

We Have Worked With Thousands of Users

Our Perspective Of Storage Has Changed

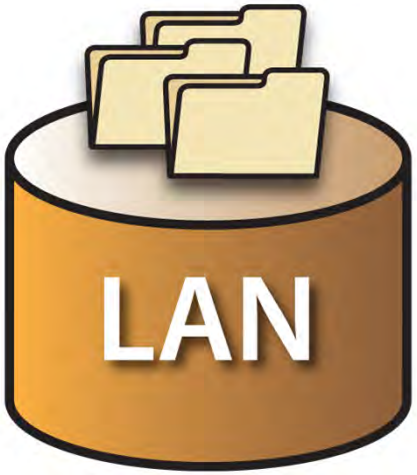
We Now Think About Deep Storage

Six Attributes of Deep Storage

- DS3/REST Interface for Object Storage
- Persistent
- Cost Effective
- Efficient
- Secure
- Easy to integrate/deploy

Deep Storage Uses DS3/REST for Objects

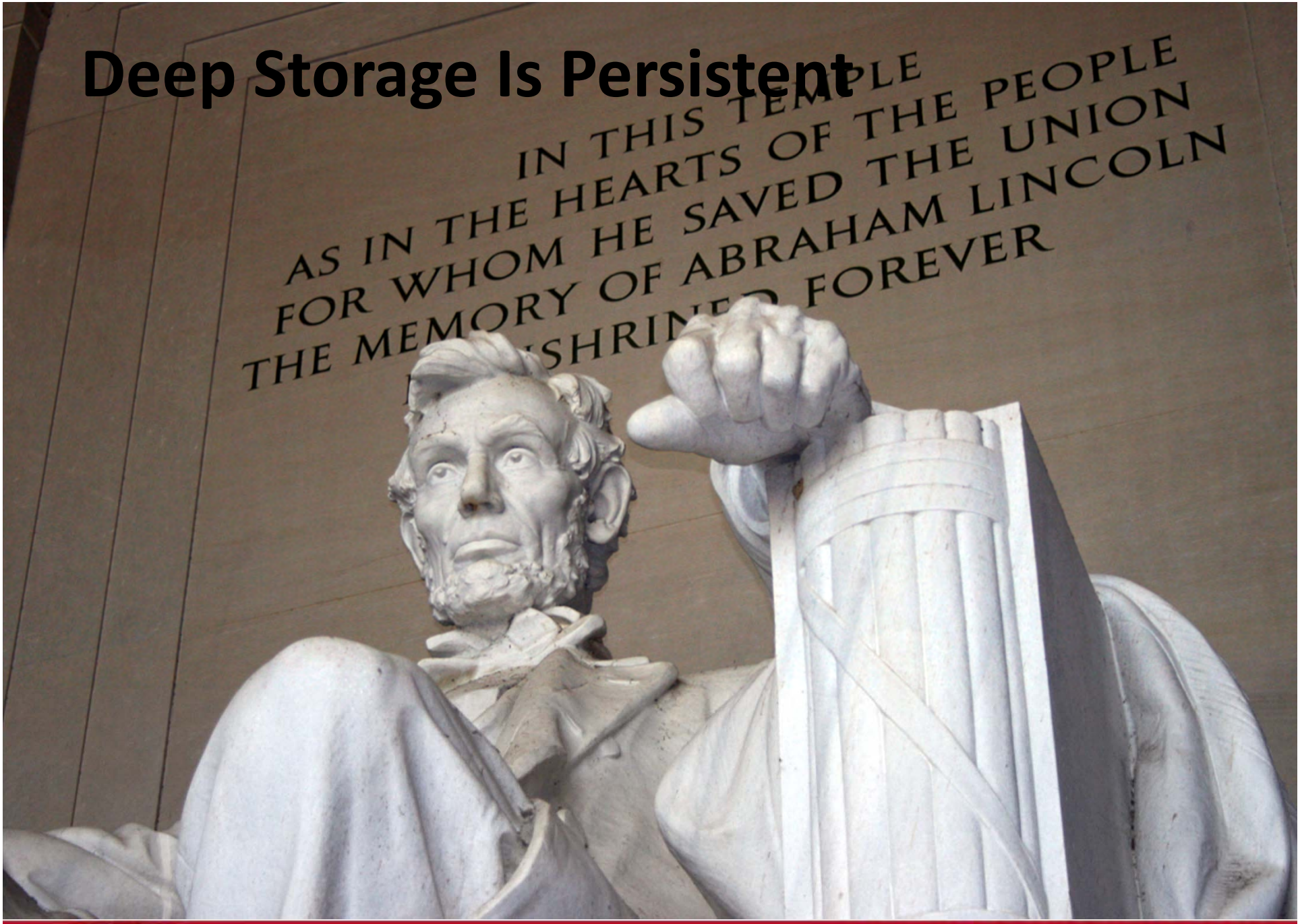
SCSI
Fibre Channel



NFS/CIFS
Ethernet

Public/Private Cloud
Web Infrastructure

Deep Storage Is Persistent



JS2

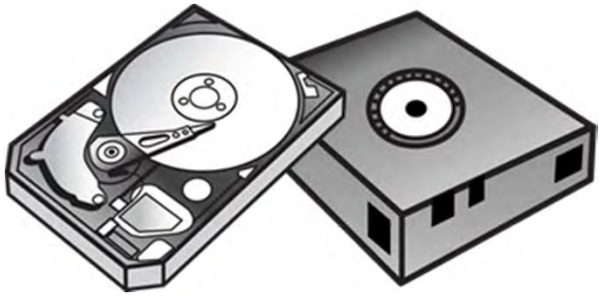
Deep Storage Is Cost Effective

1/5th to 1/50th the cost per PB of traditional disk storage

Deep Storage Is Efficient

Deep Storage Is Secure

- Connection is via https protocols
- Data is optionally encrypted at:



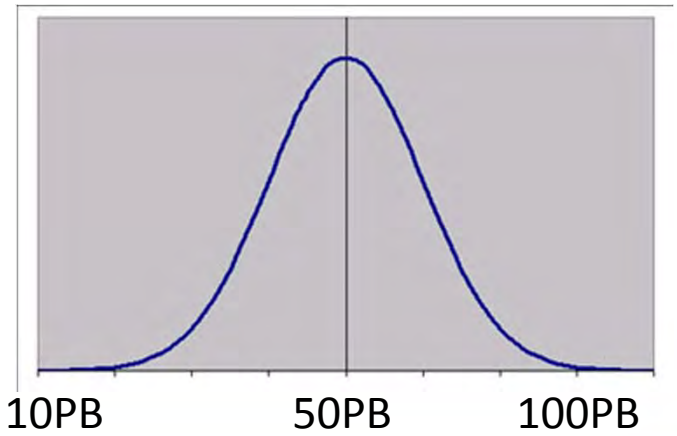
Deep Storage Is Easy To Deploy

- DS3 protocols are very easy to use, allowing many new applications. Some 500,000 applications have been written for S3
- Simple web/http methods vs. years of development for block storage programming
- Massive storage applications can be prototypes in weeks and deployed in months
- This will dramatically reduce the barriers to entry for new storage intensive applications

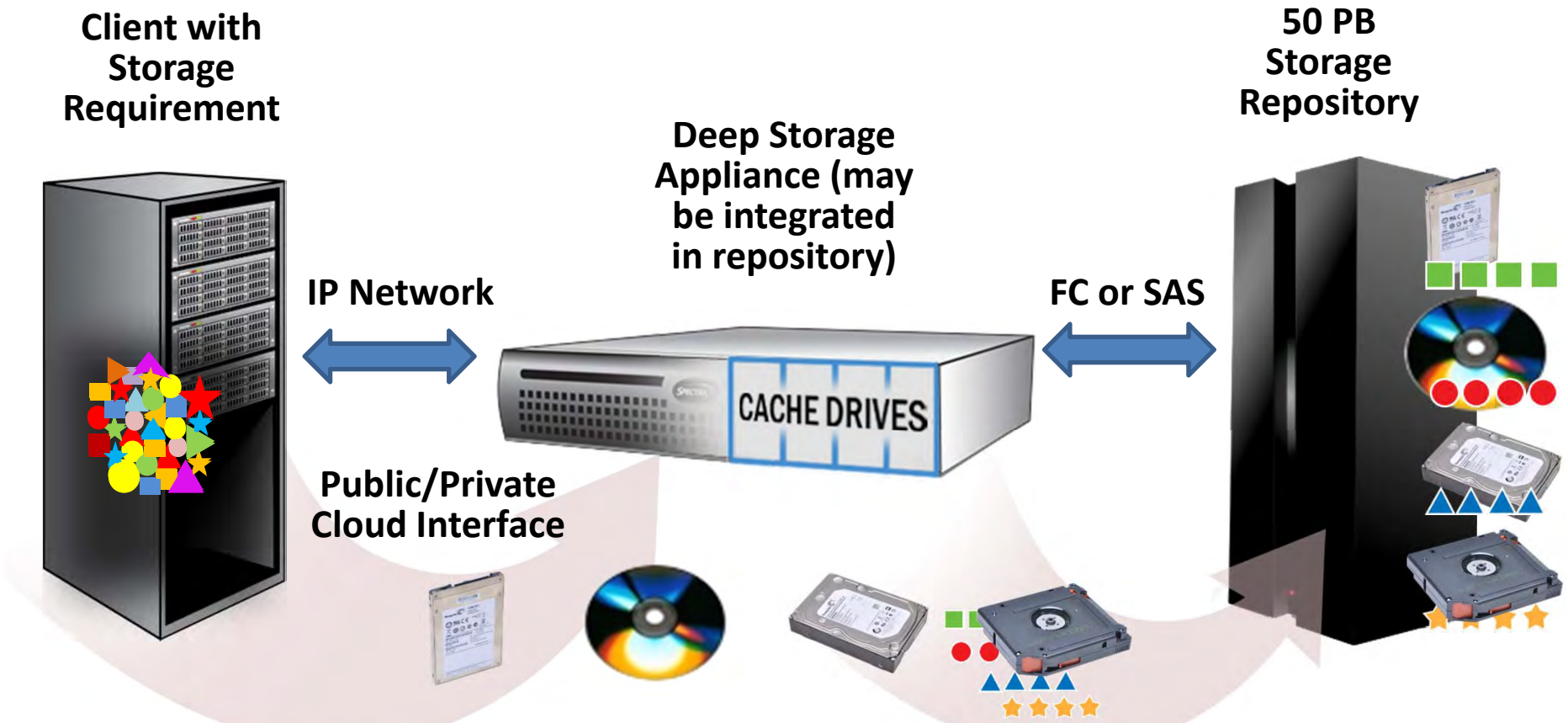
Deep Storage Applications

- Media & Entertainment
- Video Surveillance
- Seismic / Geophysical
- Health & Genomics
- Internet / Public Cloud
- Copy Management
- Historical preservation
- HPC

What does it take to build a 50 PB Deep Storage repository?



How Does Deep Storage Work?



Flash/SSD Array Storage: 50 PB

Verde	Capacity	Quantity	Price / Unit	Extended
1 Master, 9 expansions per rack, all with 96 2 TB SSD	1.920PB	31	$6,663,865	$206,579,883
TOTAL				$206,579,883

Floor Space
- 31 Racks
- 217 ft^2

Power
- 0.1 watts / drive
- 29,760 Drives
- 2,976 watts total

Chassis Power
- 15,500 watts

Total System Power
- 18,476 watts

Rack Performance
- 2,000 MB/s per system
- 209 TB/h solution

Reliability
- Undetected Bit Error Rate:
- 1 bit in 10^{17}

50PB Flash/SSD Storage System

- **Persistent**
  - ✓ 100 years storage life
  - ⛔ Second copy must be replicated
  - ✓ Low undetected error rates
- **Cost Effective, not at *($206M)***
  - ⛔ Second copy requires replication and doubles cost
  - ✓ Good SSD roadmap forward
    - But will never reach the cost per PB of disk, tape or optical
- **Efficient**
  - ✓ Lowest power consumption/performance ratio
  - ~ Software data compression
  - ✓ Proven growth path
- **Secure**
  - ~ Software driven data encryption
- **Easy to integrate/deploy**
  - ✓ Standard S3 commands
  - ✓ Microseconds to access

Facebook's Blu-Ray: 50 PB (100 GB Disks)

Blu-Ray	Capacity	Quantity	Price / Unit	Extended
Disk	100 GB	>500,000	$13.60	$6,800,000
Drives		2400	$60	$144,000
Robotics System (est.)	10,000 disks 48 drives	50	$80,000	$4,000,000
TOTAL				$10,944,000

Floor Space
- 50 Racks
- 330 ft^2

Power
- 8.8 watts / drive
- 400 watts/robotic system
- 30,560 watts total

Drive Performance*
- 18 MB/s per drive
- 103 TB/h solution

Reliability
- Undetected Bit Error Rate:
- 1 bit in 10^6 to 10^9

50PB Blu-Ray Storage System

- **Persistent**
  - ✓ 50 years storage life
  - ✓ Second copy can be transported off-site
  - − Poor error rate, additional ECC system (Forward Erasure Codes) would be needed
- **Cost Effective ($11M)**
  - − The current capacity of disks is only 100GB—so it takes a lot of them
  - − Second copy requires media only but is expensive ($6.8M)
  - − Uncertain cost roadmap forward, cost depends almost entirely on consumer adoption
- **Efficient**
  - ✓ Low power consumption
  - ✓ Growth path—300TB write once, 1PB in 5 years
- **Secure**
  - ~ Files would need to be encrypted separately
- **Easy to integrate/deploy**
  - ~ Requires DS3 priming commands with large cache of disk media, regular S3
- **Cons**
  - − Consumer grade storage. Failing drives are scrapped.
    - − Will there be a consumer market?
  - − Drives have short life, and new designs will not necessarily be compatible

Disk Arrays: 50 PB Usable

Verde	Capacity	Quantity	Price / Unit	Extended
Master with 9 Expansion Nodes	2.5 PB raw 2.2 PB available after ECC	23	$883,500	$20,320,500
TOTAL				$20,320,500

Floor Space
- 23 Racks
- 161 ft^2

Power
- 11.72 watts / drive
- 9913 disk drives
- 106,267 watts total

Chassis Power
- 11,500 watts

Total System Power
- 117,767 watts

Performance
- 2,000 MB/s system
- 158 TB/h - Solution

Reliability
- Undetected Bit Error Rate: 1 sector in 10^{15} bits read

50PB Enterprise Disk Storage System

- **Persistent**
  - ~ 5-7 years storage life
  - − Replication to make second copy
  - ✓ Low undetected error rate when implemented and corrected with ZFS RAID
- **Cost Effective ($20.3M)**
  - − High purchase cost
  - − Second copy requires replication of system (or write to other media)
  - − Roadmap challenge going forward with HAMR, SMR & patterned media
- **Efficient**
  - − Costly to acquire, maintain and operate
  - ✓ AFR predicts that only 67 drives will fail per year
  - ~ Compression can be accomplished by storage appliance
- **Secure**
  - ~ Requires external encryption unless more expensive encrypting drives are required
- **Easy to integrate/deploy**
  - ✓ Very easy to connect to S3
  - ✓ Ease of Integration makes Enterprise Disk an ideal caching system for other media

LTO-6 & Library List Price: 50 PB Configuration

Tape Storage	Capacity	Quantity	Price / Unit	Extended
Tapes	2.5 TB each	20,000	$60	$1,440,000
Drives		48	$18,950	$909,600
Tape Library	20,000 tapes 48 drives Installation	1	$2,154,922	$2,154,922
TOTAL				$4,504,522

18 Library Frames

Floor Space

- 157 Sq. Ft.

Power

- 37 watts / drive
- 1,776 watts (48 drives)
- 1,557 watts library
- 3,333 watts system

Drive Performance

- 160 MB/s per drive
- 27.6 TB/h solution

Reliability

- Uncorrected error rate: 1 bit per 10^{17} bits read
- Undetected error rate: 1 bit per 1.6×10^{33} bits read

Newest Enterprise & Library: 50 PB Configuration

Tape Storage	Capacity	Quantity	Price / Unit	Extended
Tapes	10 TB each	5,004	$250	$1,251,000
Drives		48	$25,000	$1,200,000
Tape Library	5,000 tapes 48 drives Installation	1	$1,171,244	$1,171,244
TOTAL				$3,622,244

7 Library Frames

Floor Space

- 70.5 Sq. Ft.

Power

- 46 watts / drive
- 2,208 watts (48 drives)
- 1,323 watts library
- 3,531 watts system

Drive Performance

- 360 MB/s per drive
- 62.2 TB/h solution

Reliability

- Uncorrected error rate: 1 bit per 10^{20} bits read
- Undetected error rate: 1 bit per $1.6x10^{33}$ bits read

50PB Tape Storage System

- Persistent
  - ✓ 30 years storage life
  - ✓ LTFS Format for future read-ability
  - ✓ Second copy can be transported off-site
  - ✓ Lowest undetected error rate
- Cost Effective ($3.6-4.5M)
  - ✓ Lowest Cost Storage
  - ✓ Second copy requires media only ($1.2M)
  - ✓ Good roadmap forward
- Efficient
  - ✓ Lowest power consumption
  - ✓ Hardware data compression
  - ✓ Proven growth path
- Secure
  - ✓ Built in encryption with key standards
- Easy to integrate/deploy
  - ~ Requires DS3 priming commands for awareness of tape media, otherwise S3
  - ~ 60+ seconds to data <u>if no caching</u> is used

LTO Roadmap

	LTO-3	LTO-4	LTO-5	LTO-6	LTO-7	LTO-8	LTO-9	LTO-10
Shipment Year	2005	2007	2010	2013	2015	TBD	TBD	TBD
Native Capacity	400GB	800GB	1.5TB	2.5TB	Up to 6.4TB	Up to 12.8TB	Up to 25TB	Up to 50TB
Compressed Capacity	800GB	1.6TB	3.0TB	6.25TB	Up to 16TB	Up to 32TB	Up to 62.5TB	Up to 120TB
Native Transfer Rate	80 Mb/s	120 Mb/s	140 Mb/s	160 Mb/s	Up to 315 Mb/s	Up to 472 Mb/s	Up to 708 Mb/s	Up to 1100 Mb/s
Compressed Transfer Rate	160 Mb/s	240 Mb/s	280 Mb/s	400 Mb/s	Up to 788 Mb/s	Up to 1180 Mb/s	Up to 1770 Mb/s	Up to 2750 Mb/s

Undetectable Bit Error Rate With Tape

To put in perspective how reliable tape is with its undetected Error Rate of a single bit for every 1.6×10^{33} bits read:

On average for a million tape or disk drives running continuously, at 300 and 200 Mbytes respectively, you get one undetectable tape bit error **once every five times the age of the earth** in comparison to **1,577 undetected bad sectors EVERY YEAR with disk**.

How Can You Prevent Detected (Hard) Loss?

- Disk
  - Seagate states that .68% Annual Failure Rate / drive
  - RAID 6 and longitudinal ECC can reduce bit error rate to ~equal tape. But make sure your RAID system has this.
  - To provide geographic separation one needs to replicate

- Tape
  - If 1 in 10,000 tapes are destroyed per year
  - Make two copies. Then data loss is 1 tape in 100M.
  - If that is not enough, make three copies, then data loss is 1 in 1 Trillion
  - AND you get geographic separation

Areal Density

- Disk Seagate 6TB enterprise
  - 6 platters * 2 sides
  - 8.1 square inches/side useable
  - 97.2 square inches
  - 61 Gigabytes/square inch

- Next TS Enterprise tape
  - Single sided
  - 1132 useable meters
  - 22,283 square inches
  - .45 Gigabytes/square inch

Tape surface area / Disk surface area = **230x**

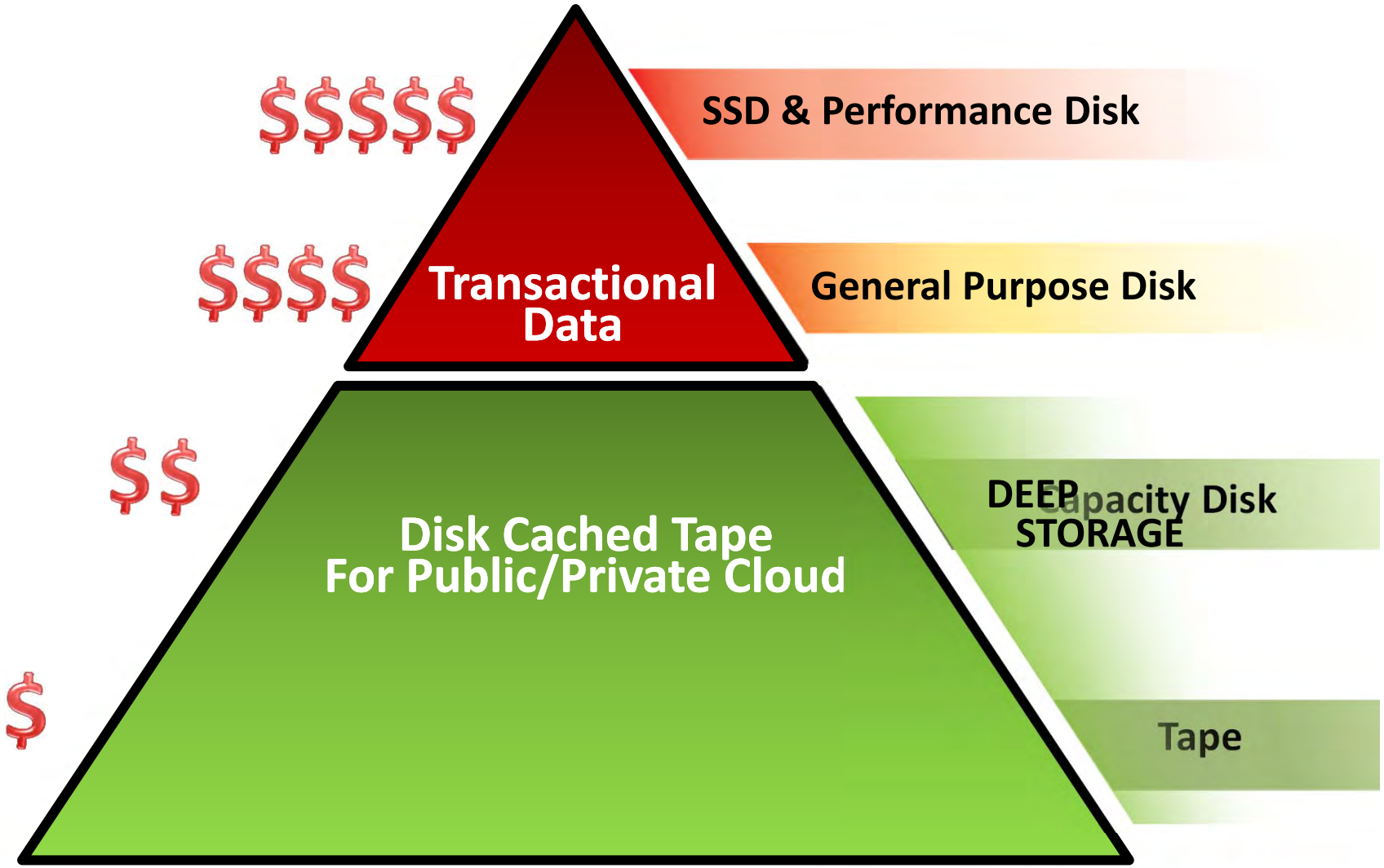
Disk areal density / Tape areal density = **127x**

If we could record at disk rates on tape we could achieve ~1.4PB on a single tape!

Transformation Of Cloud Storage

- At this point in time, cloud storage has been solely the domain of disk technology
- Today, I have shown how Tape can and will be used in private and public clouds
- And this will deliver
  - ✓ Greatest reliability
  - ✓ Greatest density
  - ✓ Greatest efficiency
  - ✓ Greatest persistence
  - ✓ Best cost model

“Disrupting” The Storage Pyramid



SPECTRA

Questions or Comments?

Thank You!

DEEP STORAGE EXPERTS