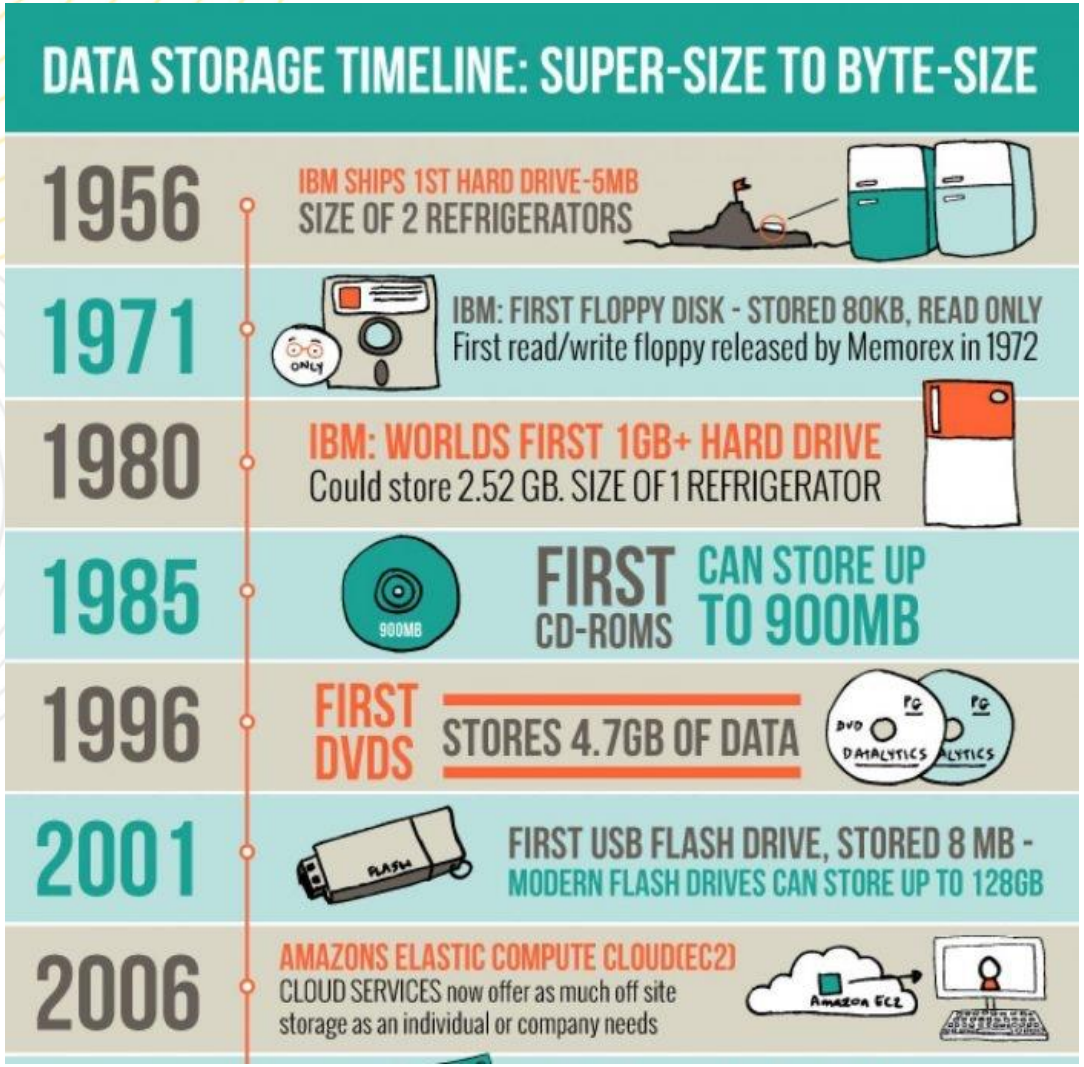
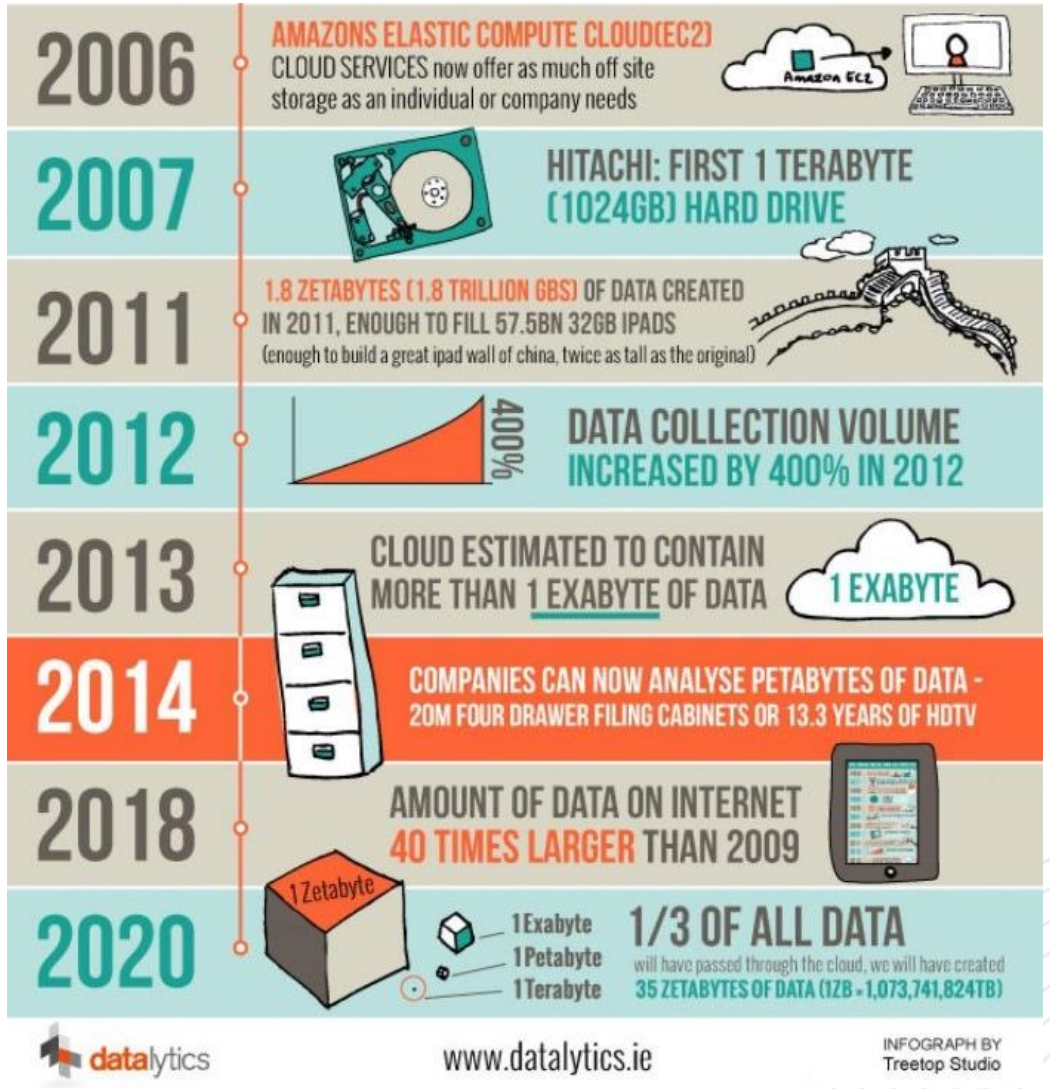
Archive Storage

Matt Ninesling

Storage throughout time





Today's Data Storage

- 2023 - 200 layer plus NAND Flash
  - Micron's 232-layer NAND
- 22 TB conventional hard drives
  - 26TB SMR drives now available
- LTO 9 is 18 TB going to 36 TB, TS is 20 TB going to 40 TB or more
  - LTO-1 was 100GB per tape when released in 2000
- Holographic Storage, DNA, Glass, still products of the future

SPECTRA

Customers struggling to find cost-effective ways to manage growing data

68%

68% of companies report storage costs as their top concern with data storage.

42%

The average enterprise saw data increase 42% over the last two years.

79%

79% of IT teams struggle to manage the tiering of their data.

40%

Spectra predicts storage footprint will grow by 40% to 14 Zettabytes by 2031.

https://www.hitachivantara.com/en-us/pdf/ebook/long-road-ahead-to-ransomware-preparedness.pdf
https://spectralogic.com/white-papers/data-storage-outlook-2022/

total-enterprise-data-volume-ocation
https://www.lightedge.com/blog/the-data-explosion-and-hidden-data-storage-costs-in-the-cloud-could-object-storage-be-the-

Top 8 Reasons to Archive

1. Large dataset management
2. Ransomware resiliency
3. Digital preservation
4. Compliance
5. Disaster recovery
6. Future utilization of data
7. Storage cost control
8. Sustainability

Best Data to Archive

Medical records, video footage, broadcast media, aerial images, genomics data, geological study data, compliance data, completed projects, and more.

Archive Options

Project Archive
Preserve large datasets for long-term retention

Bulk Archive
Migrate, manage and protect on cost and carbon-footprint efficient storage

Active Archive
Preserve growing datasets while maintaining user accessibility

SPECTRA

Data Locality is Changing

Cloud data will continue to accelerate

95%

Cloud

Approximately 95% of **new digital workloads** will in the cloud by 2025[1].

85%

Cloud First

85% of organizations will have a **cloud first principal** by 2025[4] – in two years.

80%

Multiple Clouds

80% of IT organizations will be **using multiple clouds** with three years[3].

https://www.lightedge.com/blog/cloud- trategy/#:~:text=This%2C%20coupled%20with%20other%20advantages.up%20from%2030%25%20in%202021.

On-Premise vs. Public Cloud
Scale of Data Makes a Big Difference

The public cloud is cost effective when archiving small amounts of data.

On-premises storage is far more economical for archiving large datasets.

On-Premise Digital Archives relieve the financial burden of managing large datasets

- Reducing backup capacity and cost
- Reducing recovery time from backup – improved RTO / RPO
- Avoiding ongoing cloud bills
- Reducing WAN bandwidth costs
- Reducing / eliminating cloud egress fees and access charges
- Reducing primary storage costs
- Providing near-instant access to archived files and projects

SPECTRA

Amazon Cloud Glacier vs. On-Prem Glacier

Amazon S3 Glacier

100% cloud accessed and retrieved

Archival data
Medical records, broadcast media, aerial images, consumer photos, videos, and more.

Expect costs for bandwidth and egress fees

Amazon S3 Glacier Instant Retrieval storage class
Milliseconds retrieval of data in a low-cost S3 storage class

Amazon S3 Glacier Flexible Retrieval storage class
Minutes to 12 hours retrieval of data in a lower cost archive S3 storage class

Amazon S3 Glacier Deep Archive storage class
12 to 48 hours retrieval of data in the lowest cost archive S3 storage class

Optimizes your storage costs with low-cost storage options for long-term digital preservation for rarely accessed data.

Milliseconds vs. Milliseconds

Minutes to 12 Hours vs. Seconds

12 to 48 Hours vs. Minutes

On-Prem Glacier

On-premises digital archive with hybrid cloud option

Archival data
Medical records, broadcast media, aerial images, consumer photos, videos, and more.

No additional fees

Instant Retrieval storage class
MILLISECONDS retrieval of data in a low-cost S3 storage class

Eco Object storage
Retrieval of data in **SECONDS** with spin-down disk for reduced carbon footprint

Archive S3 storage class
Retrieval of data in **MINUTES** the lowest cost & most environmentally positive solution

Almost infinitely scalable

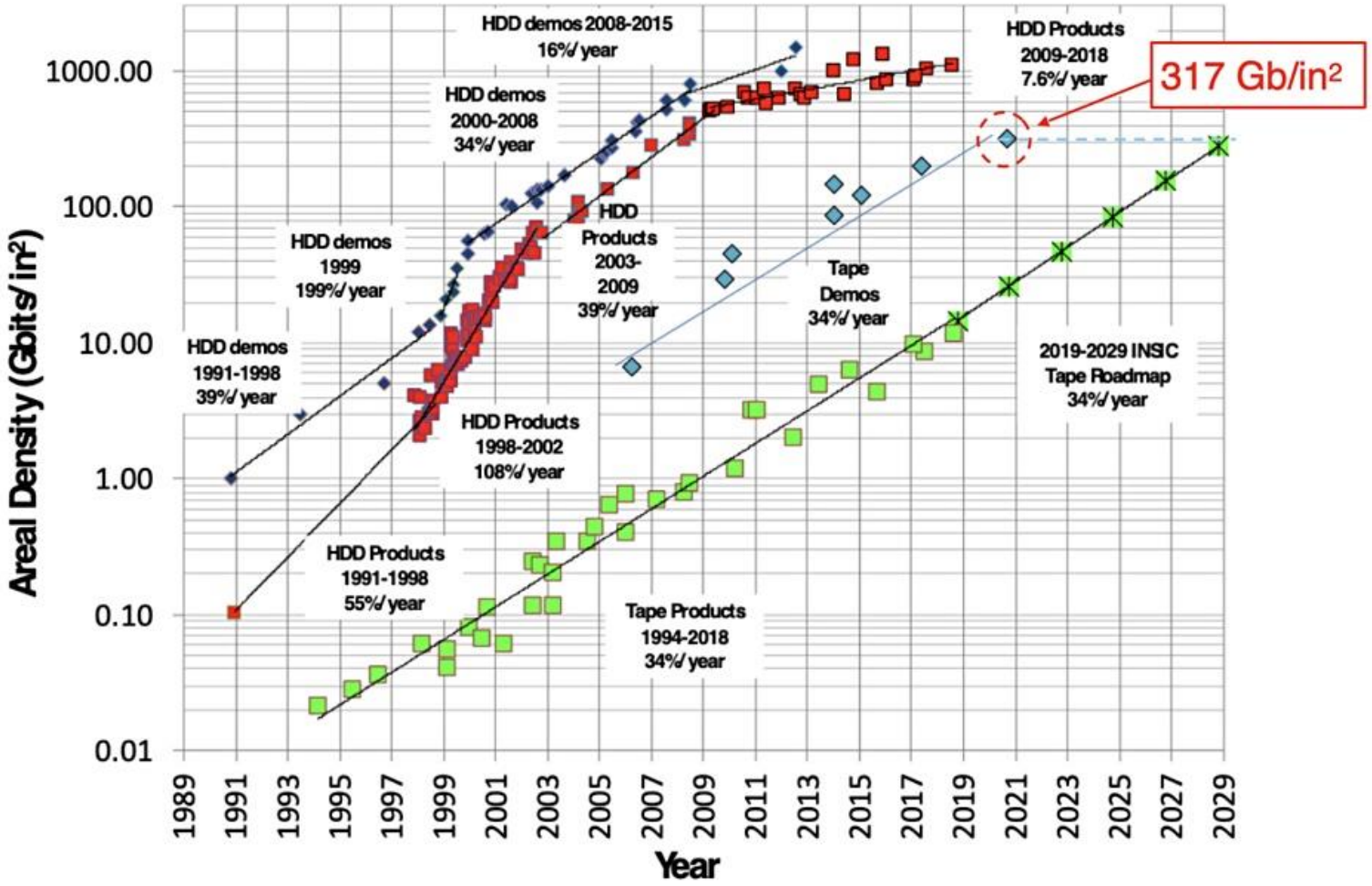
>50% lower costs and no surprises with hidden fees.

SPECTRA

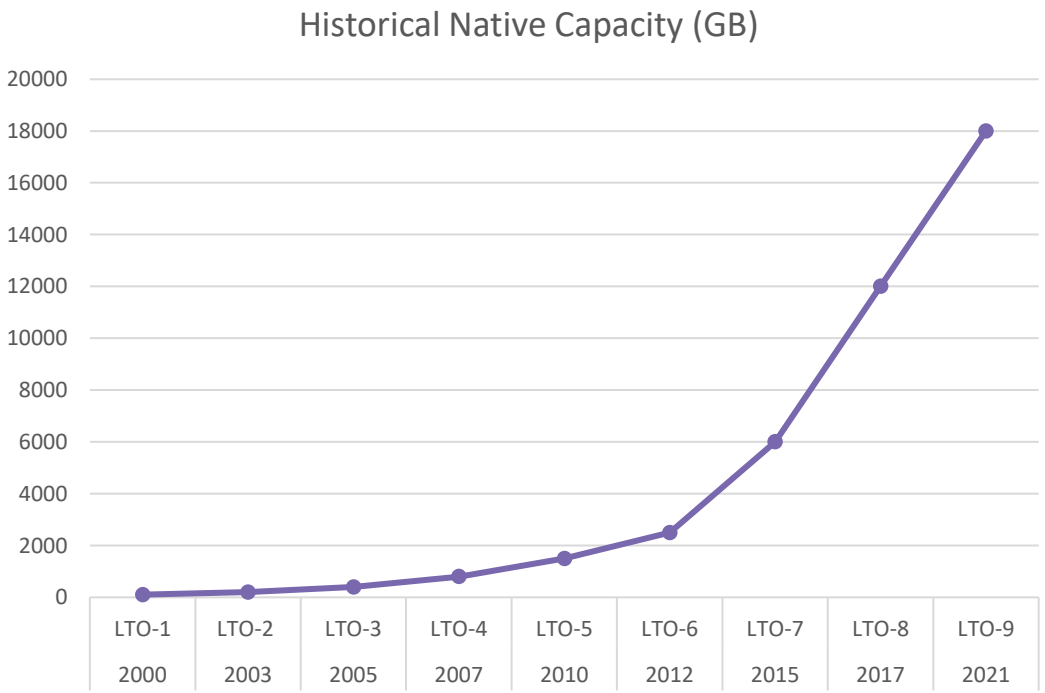
Modern Tape Archive Storage

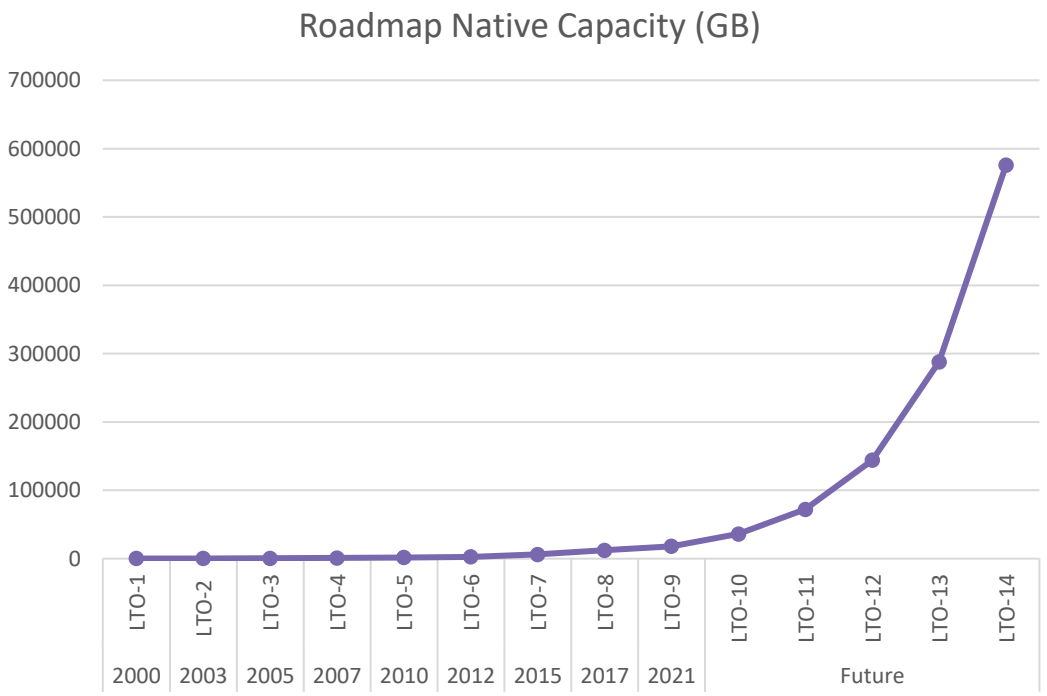
INSIC Storage Roadmap

317 Gb/in2 demonstrates the sustainability of the INSIC Tape Roadmap
34% CAGR in Areal Density for the next decade



LTO roadmap


Historical Native Capacity (GB)
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
LTO-1
2000
LTO-2
2003
LTO-3
2005
LTO-4
2007
LTO-5
2010
LTO-6
2012
LTO-7
2015
LTO-8
2017
LTO-9
2021


Roadmap Native Capacity (GB)
700000
600000
500000
400000
300000
200000
100000
0
LTO-1
2000
LTO-2
2003
LTO-3
2005
LTO-4
2007
LTO-5
2010
LTO-6
2012
LTO-7
2015
LTO-8
2017
LTO-9
2021
LTO-10
LTO-11
LTO-12
LTO-13
LTO-14
Future

Best Practices for Tape

- Dual Copy
  - Put the second copy offsite
  - Either in a vault or in another tape library
  - Use Cloud for second copy
- Store tapes in a cold archive vertically
  - This keeps them from sliding down on the hub and creating possible edge damage
  - Mount the tape and retention it periodically
- Use automated data integrity verification/fixity checking if data on tape is rarely or never accessed

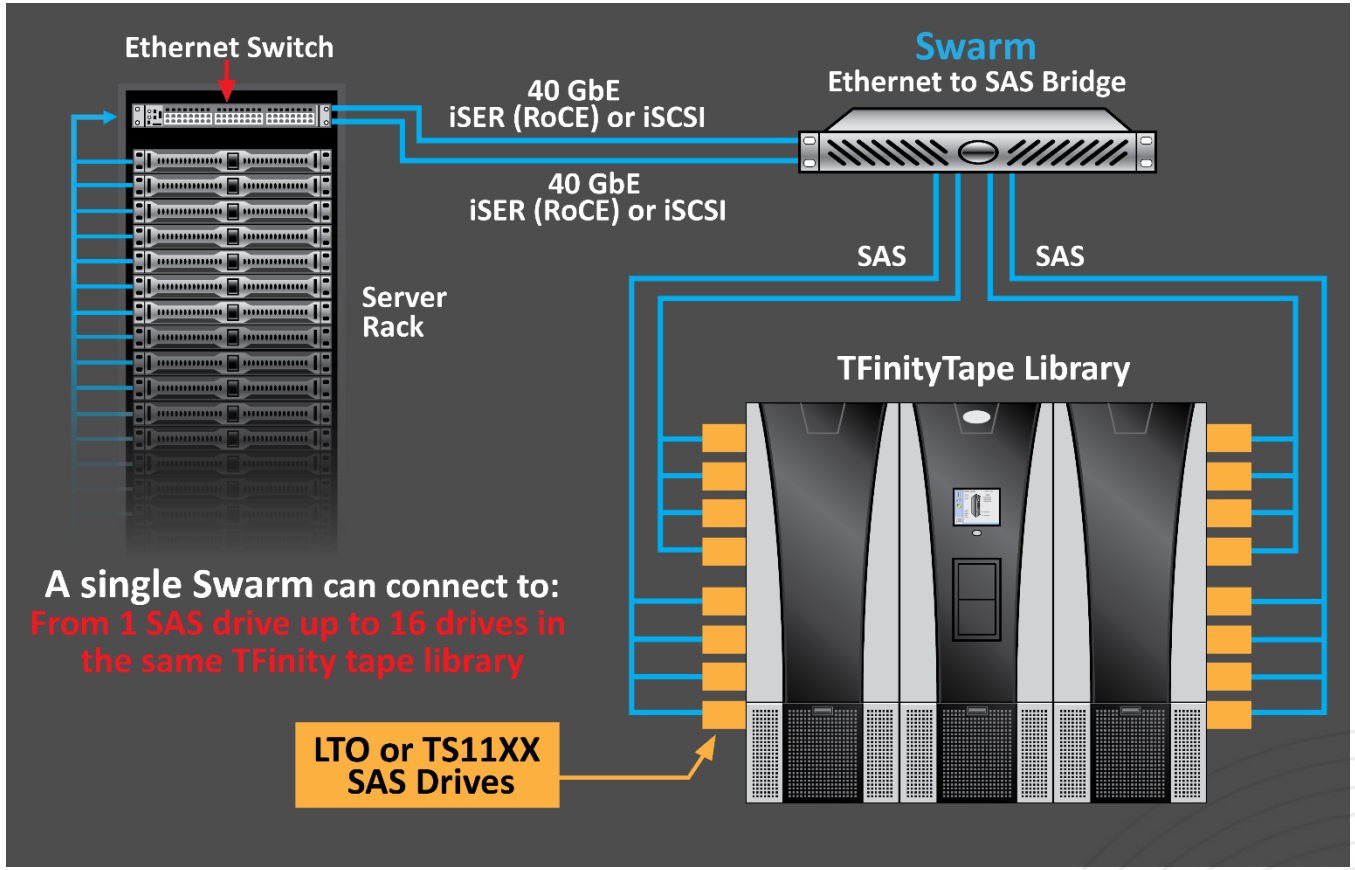
SPECTRA

Using New Interfaces to Tape

- Archiving in new ways like On-Premise Glacier require using S3
- When doing this use a non-proprietary format
  - LTFS - Don’t get caught with vendor lock-in
- This has all been driven by HyperScale

Using Ethernet with Tape

- Ethernet to SAS bridge mounts to top of library
- 2 x 40GbE Ethernet connections in
- 16 SAS connections to tape
- Full support for RoCE v2 and iSCSI
- Support for all SAS tape drives
  - FH LTO-9 and TS1160 SAS drives
- Can save thousands of dollars by moving away from fibre switches, HBAs, and dedicated personnel

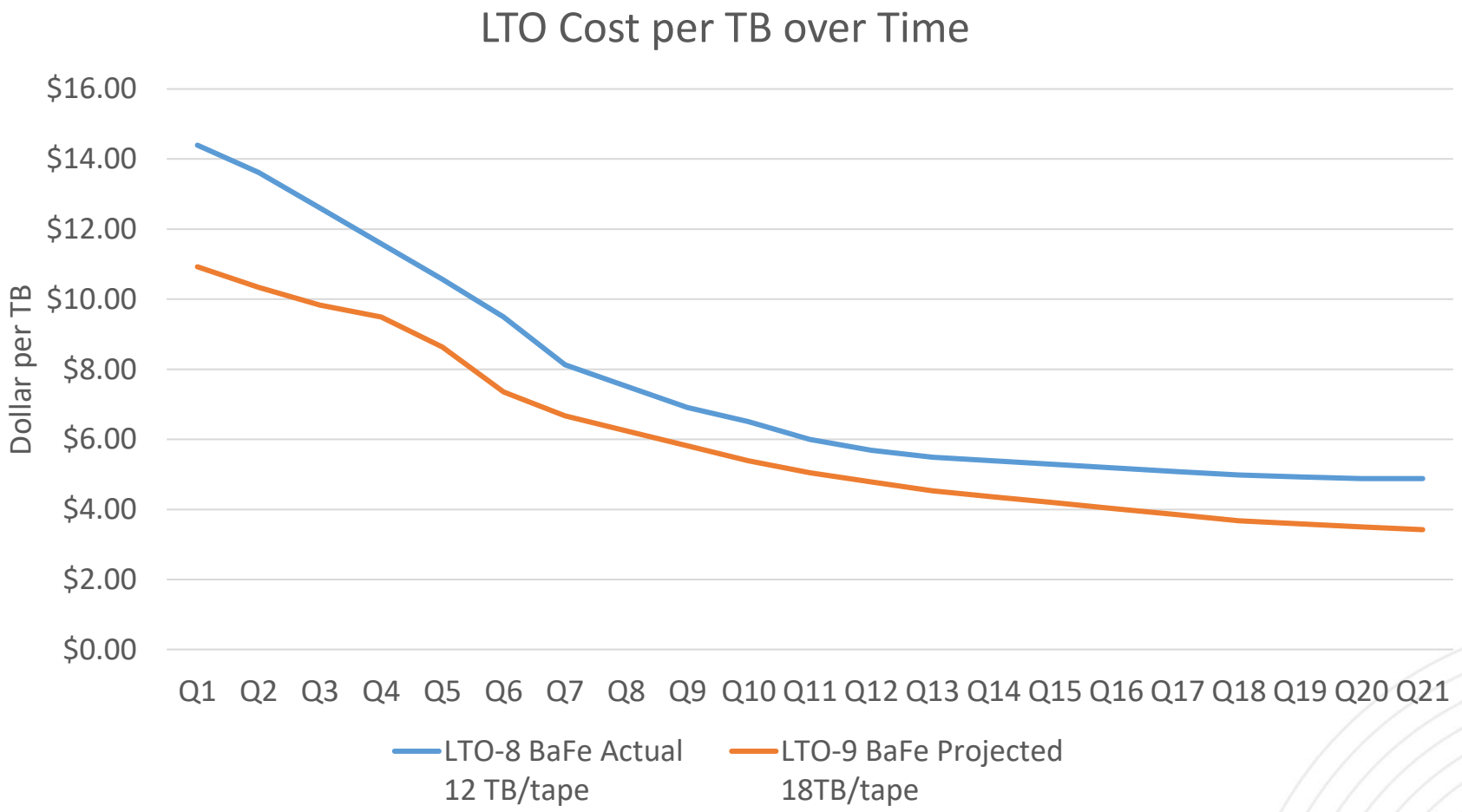


Real World Example

- Research University in Australia
- 2 TFinity libraries
  - One is 50km from Host
  - The other is 40km from Host
- Currently getting ~380 MB/s per drive and close to 500 MB/s with compression

Tape Media Price Trends – By Far the Lowest Cost Archive

- LTO-8 is currently under $5 per TB
- LTO-9 is currently under $7 per TB
- Likely almost equal on $/TB by end of this year



Benefits of Tape for Archive Storage

- Lowest Cost Storage Possible
- Longest lasting storage
  - Data can easily be kept on tape for 10+ years
  - Libraries run for 15 to 20+ years
- Safest way to archive is to have an Airgap
- Lowest Power and most environmentally friendly way to store your data
- Modern use cases with Archives and On-Premise Glacier solutions

SPECTRA

Thank You