The Role of Tape Technology in Managing the Exponential Growth of Cold Data

Osamu Shimizu
Take Nambo
Hitoshi Noguchi

FUJIFILM Corporation
Exponential growth of data and storage

Demand of data storage has been increasing along with the exponential data growth.

*1) IDC’s “The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things” sponsored by EMC (April, 2014)
Most data is never accessed

- Most data is very rarely accessed (Cold)*, however, data must be retained for preservation, legal/compliance requirements and for finding new business opportunities.**

  **Storage for COLD data is HOT topic**

- But budget is limited.

  **Reliable and Inexpensive storage for cold data is required.**

*90% data in NAS is never accessed. (Source: University of California, Santa Cruz)

**Retention of 20 year or more is required by 70%. (Source: SNIA-100 year archive survey)
## Storage media comparison for cold data

<table>
<thead>
<tr>
<th>Current</th>
<th>Future</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blue characters show advantages</strong></td>
<td><strong>Tape</strong> (Latest formats data)**</td>
</tr>
<tr>
<td><strong>Capacity</strong> [TB/unit]</td>
<td>6 to 10</td>
</tr>
<tr>
<td><strong>Sustained transfer rate</strong> [MBps]</td>
<td>252 to 360</td>
</tr>
<tr>
<td><strong>Access time in libraries</strong> [s]</td>
<td>≈30(shorter tape)-80 (incl. loading)</td>
</tr>
<tr>
<td><strong>Media lifetime</strong> [year]</td>
<td>30</td>
</tr>
<tr>
<td><strong>Cost/GB</strong> [$/GB]</td>
<td>≈0.01 (LTO)</td>
</tr>
<tr>
<td><em><em>CO2</em> Relative value</em>*</td>
<td>1/10 to 1/30</td>
</tr>
<tr>
<td><strong>Hard error rate</strong></td>
<td>1E-19 to 1E-20</td>
</tr>
<tr>
<td><strong>Write after verify</strong></td>
<td>Yes (No transfer rate loss)</td>
</tr>
<tr>
<td><strong>Latest media tech</strong></td>
<td>BaFe</td>
</tr>
<tr>
<td><strong>Capacity</strong> [TB/unit]</td>
<td>220(Demonstrated in 2015) 48(LTO10)</td>
</tr>
<tr>
<td><strong>Transfer rate</strong> [MBps]</td>
<td>Multi Ch / Linear density 1,100(LTO10)</td>
</tr>
</tbody>
</table>

*Source: JEITA tape storage committee (2013)*

**Bits per rotation at an inner position are less than at an outer, so transfer rate is slower at an inner position. (up to -50%)**


---

Tape storage is suitable for archiving cold data!!

---

2016 Data Storage Innovation Conference. © Insert Your Company Name. All Rights Reserved.
Still Tape?

**Consumer / Business (-1990s)**

VHS for video

Cassette tapes for music

- **Recording signal**: Analog
- **Capacity (2k recording)**: 2 hours
- **Transfer rate**: 2MB/s (MPEG2 format)

**Business**

Tape storage for any data

- **Recording signal**: Digital
- **Capacity (2k recording)**: Up to ≈470 hours
- **Transfer rate**: Up to 360MB/s
Tape Advantages 1: Capacity (Density)

- Tape roadmap was extended in 2015 and now covers till 2025.
- Recoding density of tape has increased at 30+% yearly, which is now faster than HDD.
- CAGR of volumetric density of tape is 40% or more contributed by thinner (longer) tapes.

Tape is the highest capacity media now!!
Tape Advantages 1: Capacity (Cont’d)

Latest technical demonstration in 2015 (IBM and Fujifilm)

- 123Gbpsi was achieved, which is 40+% larger than 85.9Gbpsi IBM and Fuji hit in 2014.
- This density enables a single tape cartridge to have 220TB which is 37 times larger capacity than the latest LTO format. (6TB)
- This capacity would be able to cover the estimated capacity of coming 5 LTO generations.
# Tape Advantages 1: Capacity (Cont’d)

Tape’s Capacity has been growing faster than past!!

<table>
<thead>
<tr>
<th>Year</th>
<th>Tape Type</th>
<th>Tracks</th>
<th>GB Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1951</td>
<td>Open Reel Tape</td>
<td>9</td>
<td>200 MB</td>
</tr>
<tr>
<td>1985</td>
<td>MTC 3480 Tape</td>
<td>18</td>
<td>210 MB</td>
</tr>
<tr>
<td>1992</td>
<td>MTC 3490E</td>
<td>36</td>
<td>800 MB</td>
</tr>
<tr>
<td>1995</td>
<td>MTC Magstar 3590E</td>
<td>128</td>
<td>10 GB</td>
</tr>
<tr>
<td>1999</td>
<td>DLT4</td>
<td>168</td>
<td>40 GB</td>
</tr>
<tr>
<td>2000</td>
<td>LTO 1</td>
<td>384</td>
<td>100 GB</td>
</tr>
<tr>
<td>2003</td>
<td>LTO 2</td>
<td>512</td>
<td>200 GB</td>
</tr>
<tr>
<td>2005</td>
<td>LTO 3</td>
<td>704</td>
<td>400 GB</td>
</tr>
<tr>
<td>2007</td>
<td>LTO 4</td>
<td>896</td>
<td>800 GB</td>
</tr>
<tr>
<td>2007</td>
<td>3592 JA</td>
<td>896</td>
<td>500 GB</td>
</tr>
<tr>
<td>2010</td>
<td>LTO 5</td>
<td>1,280</td>
<td>1,500 GB (1.5 TB)</td>
</tr>
<tr>
<td>2011</td>
<td>T10K B / 3592 JB</td>
<td>1,152</td>
<td>1,000 GB (1 TB)</td>
</tr>
<tr>
<td>2012</td>
<td>LTO 6</td>
<td>5160</td>
<td>10,000 GB (10 TB)</td>
</tr>
<tr>
<td>2013</td>
<td>T10K C</td>
<td>3,584</td>
<td>5,000 GB (5 TB)</td>
</tr>
<tr>
<td>2014</td>
<td>LTO 7</td>
<td>3,584</td>
<td>60,000 GB (6 TB)</td>
</tr>
<tr>
<td>2015</td>
<td>3592 JD</td>
<td>5,160</td>
<td>10,000 GB (10 TB)</td>
</tr>
<tr>
<td>2015</td>
<td>T10K D</td>
<td>4,608</td>
<td>8,500 GB (8.5 TB)</td>
</tr>
<tr>
<td>2016</td>
<td>LTO 8</td>
<td>7,152</td>
<td>14,000 GB (14 TB)</td>
</tr>
<tr>
<td>2017</td>
<td>T10K E</td>
<td>9,216</td>
<td>18,000 GB (18 TB)</td>
</tr>
<tr>
<td>2018</td>
<td>LTO 9</td>
<td>11,280</td>
<td>21,000 GB (21 TB)</td>
</tr>
</tbody>
</table>

**CAGR**
- 1951 to 2007: 12%
- 2007 to 2015: 39%
Tape Advantages 1: Capacity (Cont’d)

1996
- 6,000 carts per Library = 12,000 cartridges!
- TimberLine 9490 – 0.8 GB (3490 cartridge)
- 357 sq ft x 2 = 714 sq ft
- 8200 lbs x 2 = 16,400 lbs

2015
- Just 1 cartridge
- Up to 10.0TB
- 0.15 sq ft
- 0.6 lbs

Data in 6,000 carts in 1996 can be recorded into one cartridge in 2015.
Migration to advanced cart does not necessarily mean negative!!
Tape Advantages 2: Transfer rate

\[ \text{Transfer rate} = \text{Linear density} \times \text{rotations} \]

\( (5.4\, \text{to} \, 15\, \text{rpm}, \, 7.2\, \text{k is current standard}) \)

- Linear density is hitting a current technology’s limit.
- Faster rotations cannot be expected.

\[ \text{Transfer rate} = \text{Linear density} \times \text{tape speed} \times \text{number of Channels} \]

Still have room for Growth!
Tape Advantages 2: Transfer rate (Cont’d)

• Tape is slower than HDD for smaller size data when a media needs to be loaded.
• But for bigger size data and/or sequential transfer, tape is faster or impact of loading time would be smaller.
Tape Advantages 3: Reliability

1. Verify-after-write

   - In tape system, data can be verified just after write operation without losing transfer rate.

   - In HDD, Write / Read elements are on the same head module, verification would impact on the write transfer rate.
     
     eg. HGST Ultrastar Archive Ha10;
     Read transfer rate : 157MBps, Write transfer rate : 68MBps

   - ODA is also slower for write operation because of its verification.
     http://www.sony.co.uk/pro/product/archiving-storage-oda-stand-alone-drive/ods-d77u/overview/
Tape Advantages 3: Reliability (Cont’d)

2. Low hard (unrecoverble) error rate

<table>
<thead>
<tr>
<th>Storage</th>
<th>Hard error rate</th>
<th>Hours to reach possible hard error rate (50 devices)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SATA Enterprise HDD</td>
<td>1E-15</td>
<td>3</td>
</tr>
<tr>
<td>SAS/FC Enterprise HDD</td>
<td>1E-16</td>
<td>30</td>
</tr>
<tr>
<td>LTO7</td>
<td>1E-19</td>
<td>23,000 (960 days)</td>
</tr>
<tr>
<td>Enterprise tape</td>
<td>1E-19 / 1E-20</td>
<td>&gt;=27,000 (1,140 days)</td>
</tr>
</tbody>
</table>

Tape has 10’s to 1M’s of magnitudes better hard error rate.

http://www.spectra.com/pdfs/lto_ulltrium.pdf,
Tape Advantages 3: Reliability (Cont’d)

Google’s Use Case

✓ In 2011, an outage in Gmail occurred due to software bugs which affected multiple copies of data stored on disk in multiple data centers.
✓ Google was able to restored Gmail from tape backups which were stored offline.

Source: http://gmailblog.blogspot.com/2011/02/gmail-back-soon-for-everyone.html,

• Tape drive & media are separable, so drive software bugs does not necessarily mean data loss, and tape media can be stored off-line.

• Off-line tape media can protect data assets from software bugs, viruses, hackers and drive failures.
Tape Advantages 4: Cost

Total Cost of Ownership Study:

SATA Disk vs. LTO Tape Library

• 9-year analysis of archiving data, from 1PB growing to over 52PB

• Considers all costs to grow and maintain storage (hardware, maintenance, real estate and energy)

The Results

• Disk storage costs 6x the average tape TCO ($2.4M)

Source: The Clipper Group, 2015
Tape’s new role for cold data

**Hot tier**
Access frequently (50-80%) / Very small capacity (<10%)

**Warm tier**
Access sometimes / small capacity (20%)

**Cold (Archive) tier**
Access rarely / Huge capacity (80%)

- **Very fast access time**: PO-HDD to Flash
- **Moderate access time**: Tape → CO*-HDD
- **Days access time**: Backup
- **Reliable/Inexpensive**: Archive

*PO: Performance Optimized  **CO: Capacity Optimized*
Tape use case -Salesforce.com-

- Salesforce is a leading company in SaaS industry. Tape backup is utilized as a last resort.
- They backup all data to tape at each data center, on a rotating schedule of incremental and full backups.
- The backups are cloned over secure links to a secure tape archive at different facilities.
- Even if a customer completely deleted their data by mistake, salesforce.com is able to recover customer’s data at a specific point in time.

https://developer.salesforce.com/page/Salesforce_Backup_and_Restore_Essentials_Part_1
Tape use case (Cont’d) -KEK-

KEK (English name: High Energy Accelerator Research Organization)

- Tape has been selected for tons of experimental data because of its many advantages including TCO (especially in electricity) and reliability for data loss.
- Thousands of tapes with tens of IBM TS1140 drives & TS3500 libraries now (8PB).
- Has increased 1 – 2PB /year and probably increase more in future.
- 70PB scale tape system will be available from the middle of this fiscal year, and will introduce 500PB scale tape system in future.
- KEK has deployed a file-based tiered storage for 20+ years, file allocation among disk-cache and tape is well-controlled automatically by utilizing HPSS, GPFS and own policy, so that access time to the file on the tape can be shortened.
Tape use case (Cont’d) – NASA (NAS) –

- The NASA Advanced Supercomputing (NAS) Division’s archival storage system allows to archive and retrieve important data quickly, reliably, and securely.
- Copies of users' data are written to two separate tape media in silos located in two different buildings.
- Tiered storage system automatically control data allocation, such as disk-to-tape and tape-to-disk.

Disk tier: 2.9 PB locally attached SATA RAID
Tape tier: 380 PB maximum capacity tape
- 6 tape libraries, 64,820 slots total
- 84 tape drives (LTO-5&7)

http://www.nas.nasa.gov/hecc/resources/storage_systems.html
### Who else have used tape?

<table>
<thead>
<tr>
<th>Research Institute</th>
<th>CERN</th>
<th>NCSA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KEK</td>
<td>NERSC</td>
</tr>
<tr>
<td></td>
<td>NASA</td>
<td>NIH</td>
</tr>
<tr>
<td></td>
<td>QBI</td>
<td>ORNL</td>
</tr>
<tr>
<td>Media and Entertainment</td>
<td>Endemole</td>
<td>MLB</td>
</tr>
<tr>
<td></td>
<td>Fox Sports</td>
<td>Red Bull Media House</td>
</tr>
<tr>
<td></td>
<td>Los Angeles Lakers</td>
<td>NBC</td>
</tr>
<tr>
<td></td>
<td>NHK</td>
<td>Fuji TV</td>
</tr>
<tr>
<td>Meteorology</td>
<td>German Weather Service</td>
<td>Météo France</td>
</tr>
<tr>
<td></td>
<td>Deutscher Wetterdienst</td>
<td>ECMWF</td>
</tr>
<tr>
<td>Engineering &amp; Manufacturing</td>
<td>Bombardier</td>
<td>Qualcomm</td>
</tr>
<tr>
<td></td>
<td>Google</td>
<td>T3 Media</td>
</tr>
<tr>
<td></td>
<td>Salesforce.com</td>
<td>NetSuite</td>
</tr>
<tr>
<td></td>
<td>Oracle</td>
<td>IBM</td>
</tr>
</tbody>
</table>

Approximately 90% of Fortune 500 companies have tape implemented in their infrastructures.

http://www.lto.org/2015/10/lto-7-specifications-questions-answers/
Have new tape technologies come out?

Many products and technologies have come out!!

[Quick abstract in 2014 and 2015]

<table>
<thead>
<tr>
<th>Year</th>
<th>Moth</th>
<th>Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>Mar</td>
<td>Fujifilm announced Dternity a storage and archiving solution with leveraging tape technologies. (Cloud, Tape-NAS, Media vault)</td>
</tr>
<tr>
<td></td>
<td>May</td>
<td>Sony Sputtering Tape Technology Able to Store 185TB in LTO Cartridge</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IBM and Fujifilm answers to Sony, Native 154TB into LTO cartridge</td>
</tr>
<tr>
<td></td>
<td>Sep</td>
<td>Sept, 2014 IBM announced new LTFS(*) Library Edition (LE) version extends tape library support to third-party tape libraries.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>New LTO roadmap extends to generation 10 (48TB / 1,100MBps)</td>
</tr>
<tr>
<td></td>
<td>Oct</td>
<td>IBM TS1150 released (10TB / 360MBps)</td>
</tr>
<tr>
<td>2015</td>
<td>Mar</td>
<td>Oracle updated its HSM with LTFS support and providing an OpenStack swift interface</td>
</tr>
<tr>
<td></td>
<td>Apr</td>
<td>IBM and Fujifilm demonstrated 220TB tape cartridge</td>
</tr>
<tr>
<td></td>
<td>Dec</td>
<td>Crossroads, StrongBox 3.0 NAS for unstructured data, offers solution that provides both file and object storage with LTFS tape. (1st StrongBox was released in April, 2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LTO generation 7 released (6TB / 300MBps)</td>
</tr>
</tbody>
</table>

*LTFS (Linear Tape File system) enables tape to be non-proprietary data storage and to drag-and-drop for data transfer like other storage devices. LTO has this function from LTO5, Oracle T10000X and IBM TS11XX also have similar function. SNIA(Storage Network Industry Association) TWGs have also studied a tape related item such as “Object storage on LTFS”, “LTFS as a cloud backing store”, “LTFS bulk transfer to/from/between clouds for standardization. (Ref: [http://www.snia.org/ltfs](http://www.snia.org/ltfs), [http://www.snia.org/sites/default/files/LTFS_PPT_FINAL.pdf](http://www.snia.org/sites/default/files/LTFS_PPT_FINAL.pdf))
LTFS (Linear Tape File System)

- Conventional tape system was inconvenient for searching and accessing data because of its proprietary backup software.
- LTFS, has been used from LTO5, enables tape system to be file-based storage to have drag & drop function like USB memory and HDD.
- Since then, Tape NAS products have been released from many companies including Crossroads, IBM, HP, Quantum, Spectra Logic and Fujifilm.
- LTFS have been used in a wide variety of industries including film, media, R&D and medical (http://www.lto.org/resources/case-studies/)

LTFS got Emmy award in 2011 as a contribution for “improving the ability of media companies to capture, manage and exploit content in digital form, fundamentally changing the way that audio and video content is managed and stored. “ (http://www.research.ibm.com/articles/linear-tape-file-system.shtml)

- LTFS became LTFS standard in Apr, 2016. The data content (not the physical media) of the LTFS format shall be interchangeable among all data storage systems claiming conformance to this format. http://www.research.ibm.com/articles/linear-tape-file-system.shtml
LTFS use case

Los Angeles Kings NHL Hockey Team

**CHALLENGE for LA Kings**

- Increase of video contents: 32 – 35TB per season (only 100GB in 2007/2008 season)
- Need to archive from legacy data (1967) to future data
  
  Format must be non-proprietary and solution must be reliable & scalable
- Minimal management resources
  
  Solution must be easy to use
- CAPEX reduction

LA Kings has selected Tape NAS for active-archive and deep archive.

Demand for tape

- Demand for tape has continued to increase!

Total Capacity by Year (PB Compressed)

Tape Technologies
Fujifilm media development history

Fujifilm has developed own media which has enabled tape system have more capacity and reliable.

- **Single layer-MP**: 0.25μm (Particle Size)
- **ATOMM-MP**: 0.1μm
- **NANOCUBIC-MP**: 35 to 45nm
- **NANOCUBIC-BaFe**: 10 to 20nm
- **NANOCUBIC-BaFe**: 50 to 100nm
- **1μm**: 50 to 100nm
- **0.1μm**: 35 to 45nm
- **0.25μm**: 2 to 3μm
- **0.2 to 0.3μm**: 10 to 20nm

Advanced Super Thin Layer & High Output Metal Media Technology

- **Fine particle**
- **Thin / smooth coating**
- **Fine dispersion**

- **Pro videos**
- **Pro videos**
- **Pro videos**
- **Pro videos**

- **ATOMM**
- **NANOCUBIC**
- **Thin / smooth coating**
- **Fine dispersion**

- **LTO1 to 3**
- **LTO4 to 5**
- **LTO7 and future**
- **TS1140&50 and future**
- **T10000C&D and future**
- **T10000A&B**
- **JA1 to TS1120**
Cartridge Capacity trends

http://www-03.ibm.com/support/techdocs/atsmastr.nsf/5cb5ed706d254a8186256c71006d2e0a/82f67152325e844985257960005866fa/$FILE/IBM%20TS1140%20Technology%20White%20Paper%2011%20October%202011%20Final%20v3.pdf

http://www.oracle.co.jp/events/jpm120809/materials/20120809-10_StorageSumit_A-2.pdf

http://www.lto.org/technology/index.html  (The launch of a new generation of LTO is assumed to be every 2 and half years.)


2016 Data Storage Innovation Conference. © Insert Your Company Name. All Rights Reserved.
### BaFe magnetic particle

- All the latest tape systems use BaFe particle technology.
- BaFe magnetic particles have been investigated for 30+ years.
- BaFe has supported recent new tape roadmap and exponential capacity growth.
Long term reliability of Tape system

• High reliability of tape system is guaranteed by multi level error free mechanism including
  • Read-after-write verification
  • Strong C1/C2 Reed Solomon error correction code
  • Redundancy array, if needed

• Long term reliability of tape systems is also supported by
  • Chemical stability
  • Physical stability
  • Magnetical stability

of materials used in the tape media.
Read-after-write verification

- Reeder element array and writer element array are independently and simultaneously operated.

All data are verified just after write process and will be re-written if necessary.
Strong C1/C2 Reed Solomon error correction

User data → RS encoder → ECC → No Errors

RS decoder → Error correction → Byte Error Rate ≈ 10^{-4} (before correction)

Decoded data → PRML block → Viterbi Decoder → Equalizer

User data 234x54 → C1=6
C2=10 byte

Examples of Redundancy array

$P = D_1 \oplus D_2 \oplus D_3 \oplus D_4$

If $D_1$ tape is damaged, it will be simply reproduced by

$D_1 = P \oplus D_2 \oplus D_3 \oplus D_4$

Redundancy array of inexpensive device (RAID)

https://www.youtube.com/watch?v=eNIoM9NtCM

or Redundancy Array of Independent Tape (RAIT)

http://www.activearchive.com/common/pdf/AA-Case-Study-NCSA.pdf
Chemical stability of tape media

Outline plot: stored in the office environment
Fill plot: stored in 60 °C dry (x-axis is expanded 10 times)

- LTO-1
- Barium ferrite

12 years
Physical stability of tape media

Outline plot: stored in the office environment
Fill plot: stored in 60 °C dry (x-axis is expanded 10 times)

Surface Roughness (nm)

duration (day)

LTO-1
Barium ferrite

12 years
Magnetical stability of tape media

Hc (kA/m)

LTO-1
Barium ferrite

Fill plot: stored in 60 °C dry (x-axis is expanded 10 times)

Outline plot: stored in the office environment

duration (day)

12 years
Future Tape prospect
# Tape used in 123 Gbpsi demonstration

## Properties of Magnetic Particle and Experimental Tapes

<table>
<thead>
<tr>
<th>Media type</th>
<th>New BaFe tape</th>
<th>TS1150 JD media</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume (nm³)</td>
<td>1600</td>
<td>1950</td>
</tr>
<tr>
<td>Coercivity (kA/m)</td>
<td>223</td>
<td>190</td>
</tr>
<tr>
<td>σs (A·m²/kg)</td>
<td>45</td>
<td>50</td>
</tr>
</tbody>
</table>

### Magnetic Properties

- **Longitudinal direction**
  - Coercivity (kA/m) | 146 | 182 |
  - SQ | 0.24 | 0.39 |

- **Perpendicular direction**
  - Coercivity (kA/m) | 263 | 214 |
  - SQ | 0.87 | 0.66 |

### Surface Roughness

- **Optical interferometry**
  - Ra (nm) | 0.9 | 1.6 |
- **AFM**
  - Ra (nm) | 1.8 | 2.0 |
  - Rz (nm) | 27 | 34 |

- **Coefficient of friction** | 0.21 | 0.44 |

1) With demagnetization compensation, 2) Measured with the HD2000 instruments from WYKO and a measurement area of 170 μm × 236 μm, 3) Measurement area of 40 μm square, 4) Against AlTiC cylindrical bar, at speed of 14mm/s, and back tension of 0.98 N

Tape used in 123 Gbpsi demonstration

Fig. 1. Surface profile images measured by optical interferometry. (a) New BaFe tape. (b) TS1150 JD media.

\[ R_a = 0.9 \text{ nm} \quad R_a = 1.6 \text{ nm} \]

Fig. 2. Surface profile images measured by AFM. (a) New BaFe tape. (b) TS1150 JD media.

\[ R_a = 1.8 \text{ nm}, \ R_z = 27 \text{ nm} \quad R_a = 2.0 \text{ nm}, \ R_z = 34 \text{ nm} \]
ECC used in 123 Gbps demonstration

Presented at TMRC2014 F2, pp.79-80, Aug. 2014
Details will be published in IEEE Transaction on Magnetics Vol. 51, No. 1 in January 2015.
Summary

*Tape system is the best solution for managing exponential growth of cold data*

*Because of its advantages including higher reliability, archivability, lower TCO and future prospect*