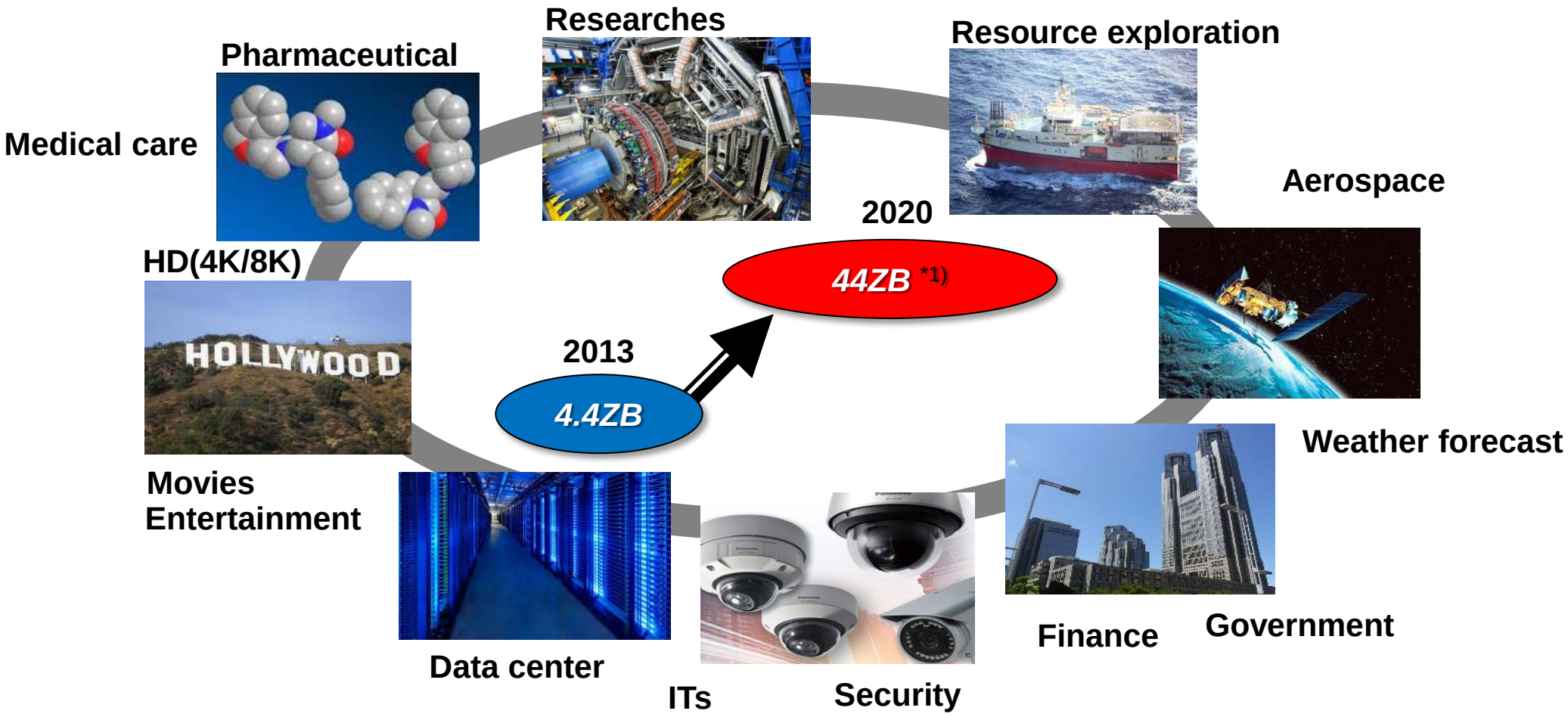
SNIA DSI '16 DATA STORAGE INNOVATION CONFERENCE
Innovation in Storage Products, Services, and Solutions
June 13-15, 2016 | Marriott San Mateo | San Mateo, CA

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

Osamu Shimizu
Take Nambo
Hitoshi Noguchi

FUJIFILM Corporation

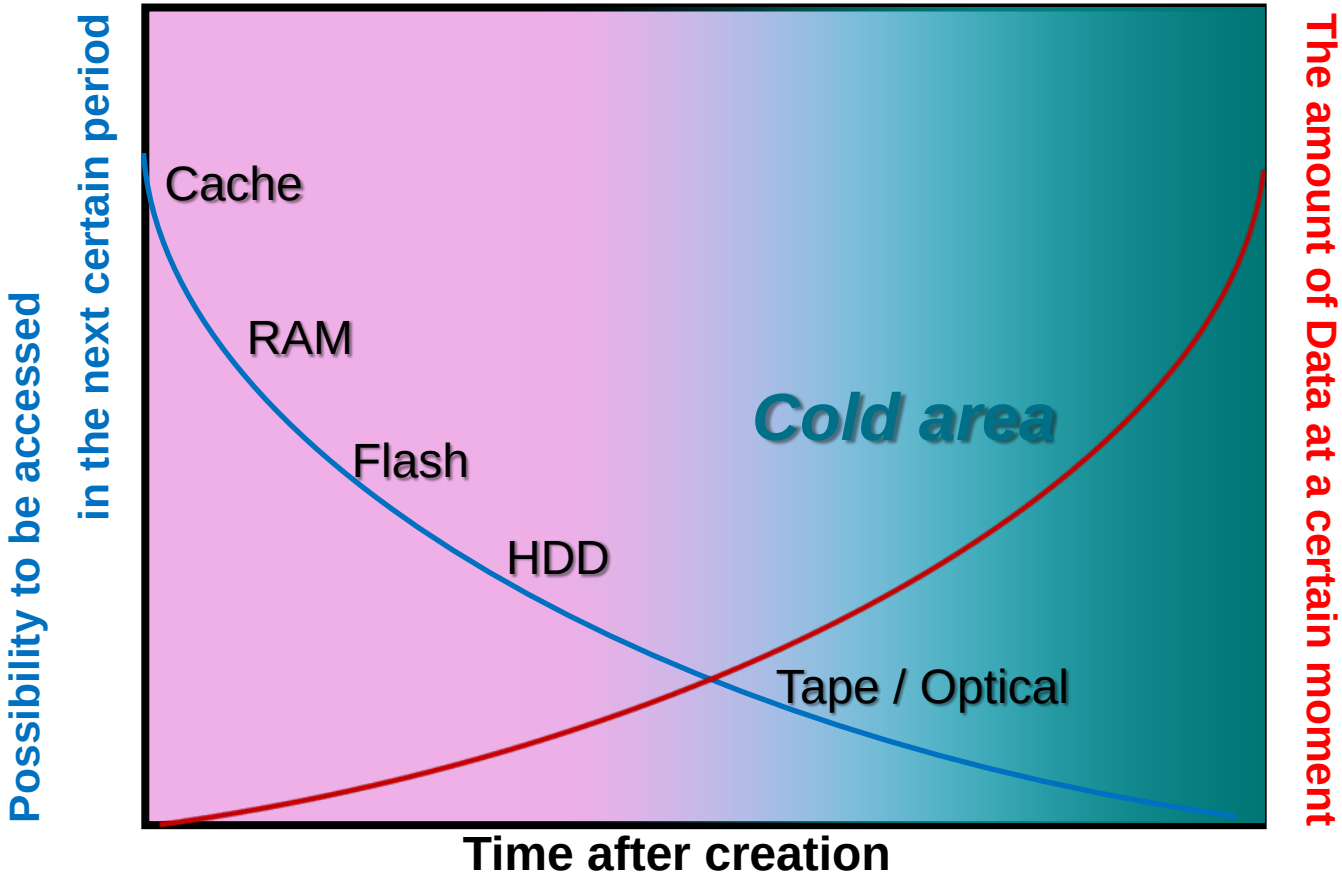
Exponential growth of data and storage



*1)IDC's "The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things " sponsored by EMC（April, 2014）

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

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

Blue characters show advantages			Tape (Latest formats data)	Capacity Optimized-HDD	Optical disc
Current	Capacity	[TB/unit]	6 to 10	4 to 10	0.128(Blu-ray) 1.5TB/cart(12 discs)
	Sustained transfer rate	[MBps]	252 to 360	≈160 to 249 (Slower at inner positions**)	Up tp 138 (Read) Up tp 55 (Write) (Slower at inner positions**)
	Access time in libraries	[s]	≈30(shorter tape)-80 (incl. loading)	mili	≈60-90 (inlc. loading)***
	Media lifetime	[year]	30	3 to 5	50
	Cost/GB	[$/GB]	≈0.01（LTO)	≈0.05	≈0.10 (Archival disc)
	CO2*	Relative value	1/10 to 1/30	1	Similar to tape
	Hard error rate		1E-19 to 1E-20	1E-15 to 1E-16	-
	Write after verify		Yes (No transfer rate loss)	Optional (Transfer rate may drop in write operation)	Optional (Transfer rate may drop in write operation)
	Latest media tech		BaFe	SMR/He-Shield	Multi layer
Future	Capacity	[TB/unit]	220(Demonstrated in 2015) 48(LTO10)	20 with HAMR / TDMR 100 by 2025?	≈0.46/disc
	Transfer rate	[MBps]	Multi Ch / Linear denisity 1,100（LTO10）	Up to ≈250? (Constrain of rpm)	≈250 (Read) ≈125 (Write)

*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%)

*** http://www.everspan.com/specs, http://hlds.co.kr/v2/HL200_eng.pdf, http://hlds.co.kr/v2/HL200_eng.pdf

Tape storage is suitable for archiving cold data!!

Still Tape?

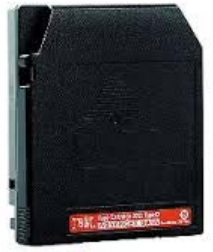
Consumer / Business (-1990s)

VHS for video

Cassette tapes for music

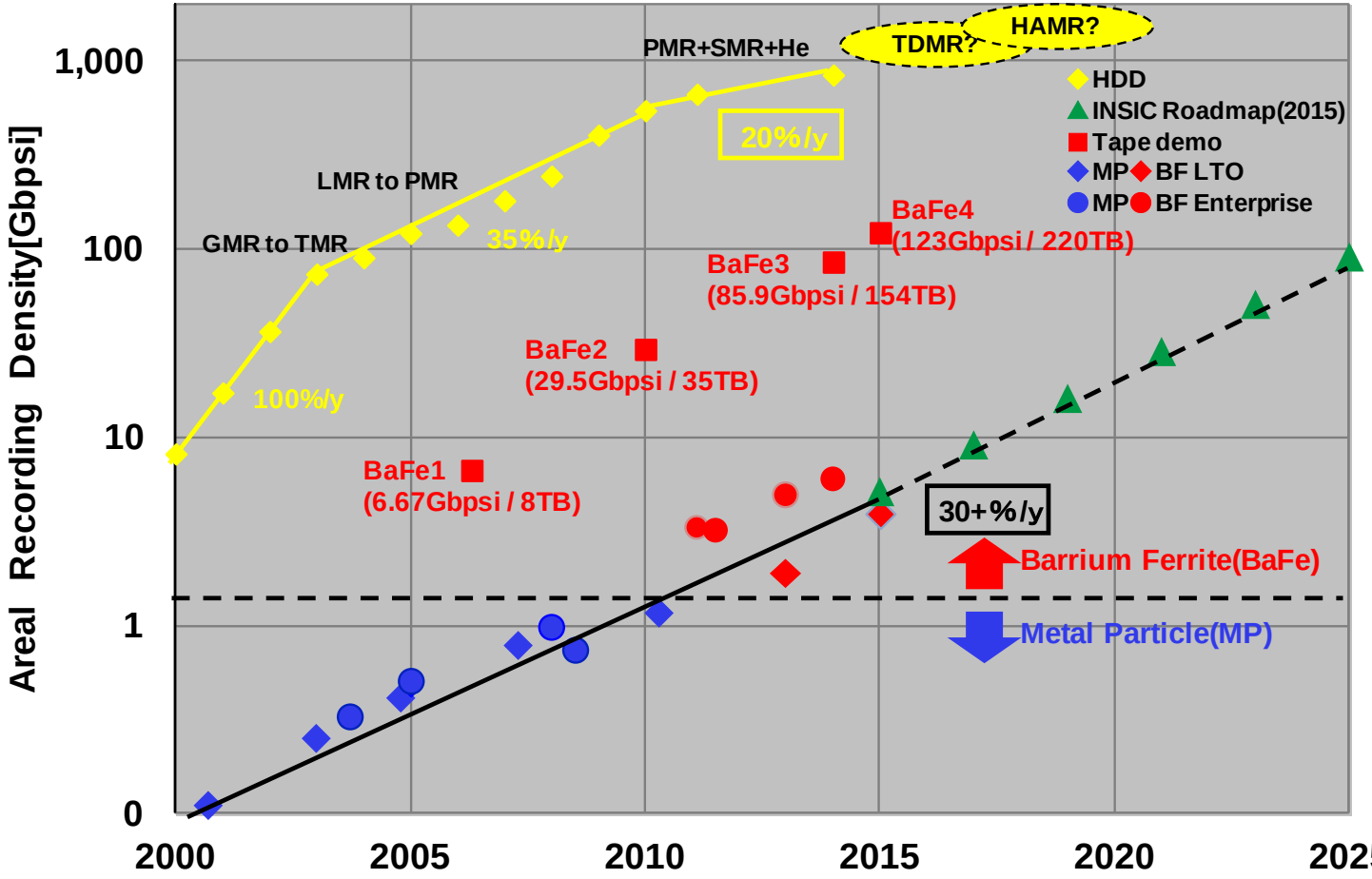
Business

Tape storage for any data



Recording signal	**Analog**		**Digital**
Capacity(2k recording)	**2 hours**	**200x**	**Up to ≈470 hours**
Transfer rate	**2MB/s (MPEG2 format)**	**125x**	**Up to 360MB/s**

Tape Advantages 1 : Capacity (Density)



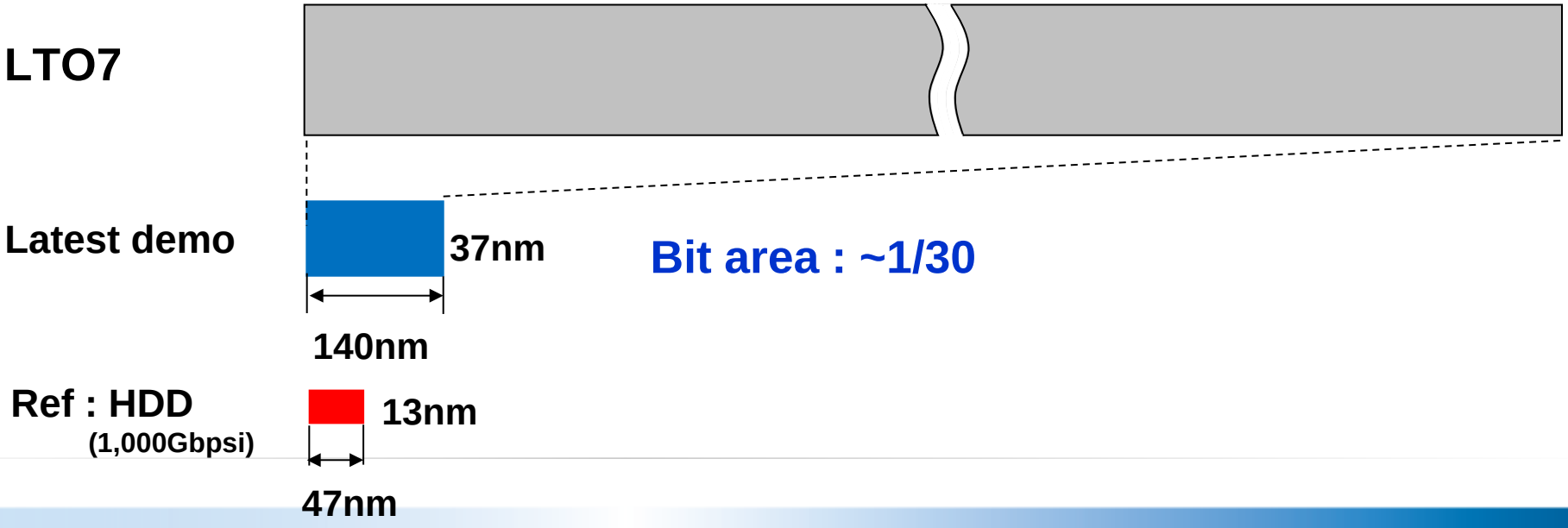
- Tape roadmap was extended in 2015 and now covers till 2025.
- Recoding density of tape has increased at 30+% yearly, which is now faster than HDD
- CAGR of volumetric density of tape is 40 % or more contributed by thinner (longer) tapes.

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!!

60+ YEARS

15 YEARS

≈50 YEARS

2015	LTO 7	3, 584 tracks	60,000 GB (6 TB)
2014	3592 JD	5,160 tracks	10,000 GB (10 TB)
2013	T10K D	4,608 tracks	8,500 GB (8.5 TB)
2012	LTO 6	2,176 tracks	2,500 GB (2.5 TB)
2011	T10K C	3,584 tracks	5,000 GB (5 TB)
2011	3592JC	2,560 tracks	4,000 GB (4 TB)
2010	LTO 5	1,280 tracks	1,500 GB (1.5 TB)
2008	T10K B / 3592 JB	1,152 tracks	1,000 GB (1 TB)
2007	LTO 4	896 tracks	800 GB
2007	3592 JA	896 tracks	500 GB
2005	LTO 3	704 tracks	400 GB
2003	LTO 2	512 tracks	200 GB
2000	LTO 1	384 tracks	100 GB
1999	DLT4	168 tracks	40 GB
1995	MTC Magstar 3590	128 tracks	10 GB
1992	MTC 3490E	36 tracks	800 MB
1985	MTC 3480 T	18 tracks	210 MB
1951	Open Reel Tape	9 tracks	200 MB

CAGR 39%

CAGR 12%

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**

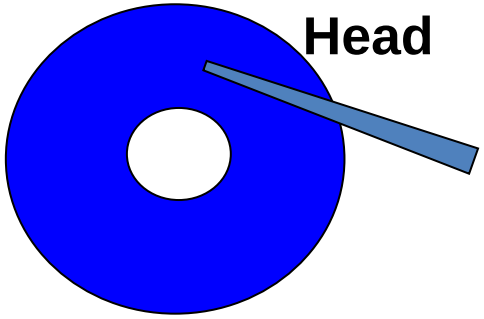
2015

- 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

＜HDD＞

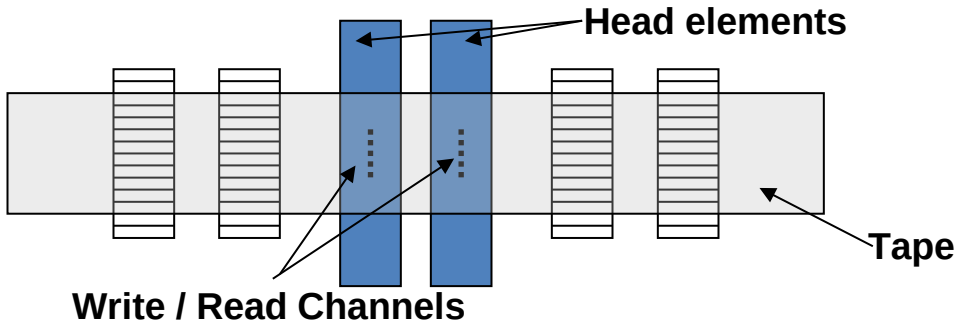


Transfer rate = Linear density * rotations
(5.4k to 15k rpm, 7.2k is current standard)

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

＜Tape＞

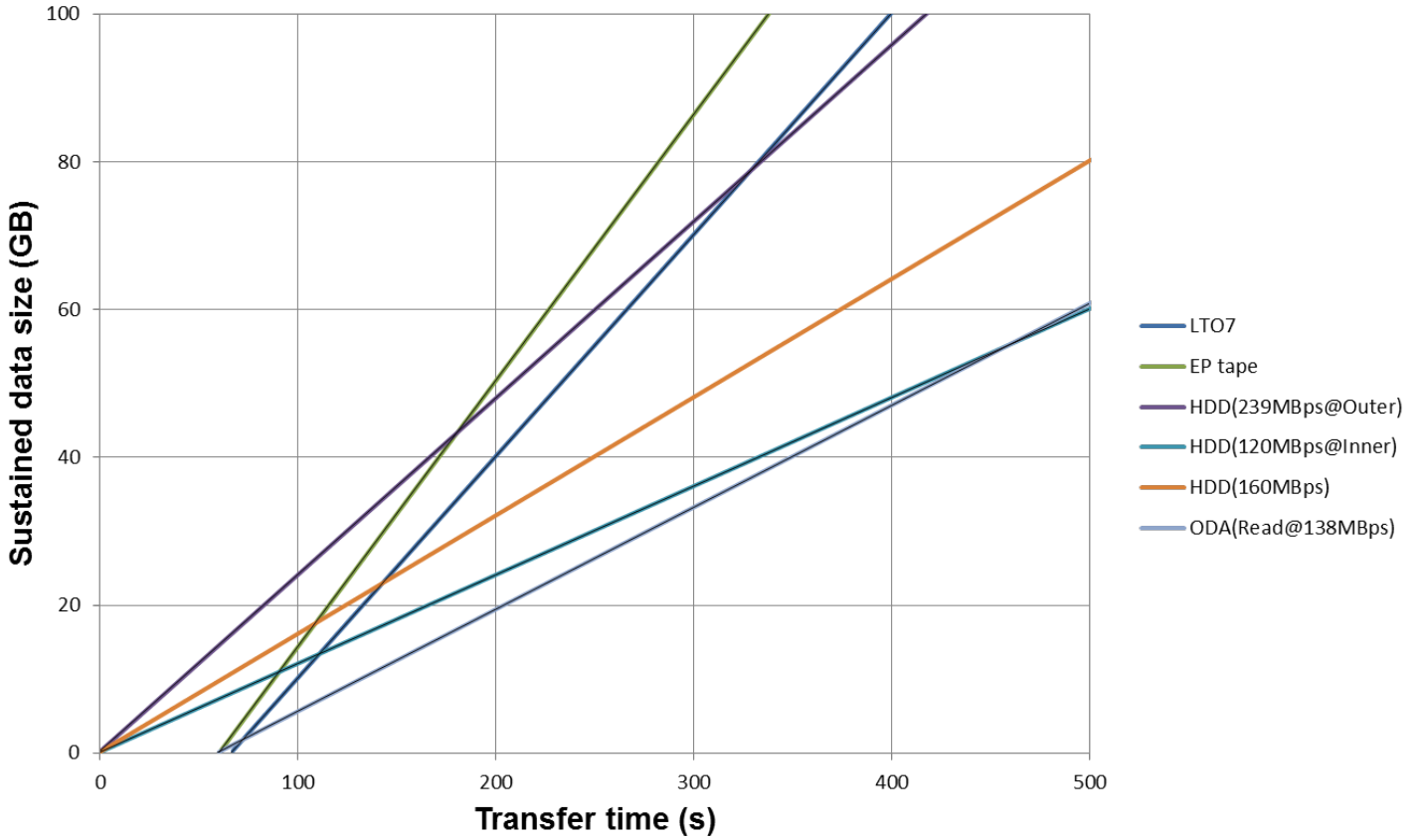
Tape overhead view



Transfer rate = Linear density * tape speed * number of Channels

Still have room for Growth!

Tape Advantages 2: Transfer rate (Cont'd)

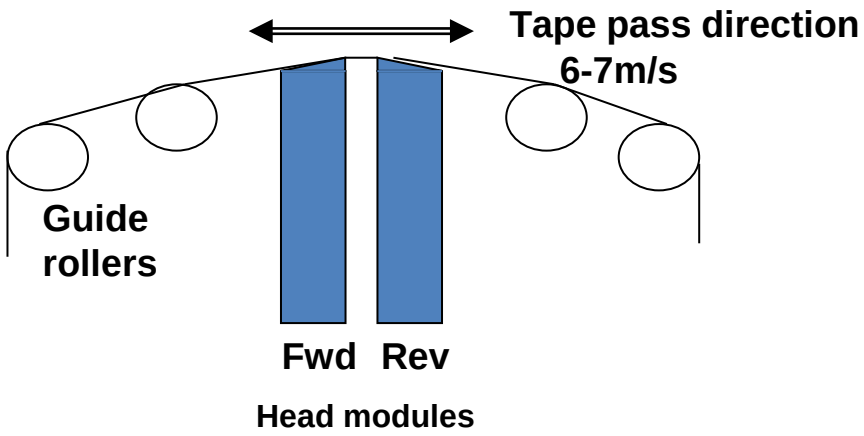


- **Tape is slower than HDD for smaller size data when a media needs to be loaded.**
- **But for bigger size data and/or sequential transfer, tape is faster or impact of loading time would be smaller.**

Tape Advantages 3: Reliability

1. Verify-after-write

Tape side view



- **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
  http://www.hgst.com/tech/techlib.nsf/techdocs/9FE5F7F0917244B888257E58006C9906/$file/Ha10-Ultrastar-HDD-DS.pdf

- **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

Storage	Hard error rate	Hours to reach possible hard error rate (50 devises)
SATA Enterprise HDD	1E-15	3
SAS/FC Enterprise HDD	1E-16	30
LTO7	1E-19	23,000 (960 days)
Enterprise tape	1E-19 / 1E-20	>=27,000 (1,140 days)

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

Source: "Tape: Comparison of LTO and Enterprise", Instrumental, Inc. http://www.instrumental.com, April 19, 2013
http://www.seagate.com/files/www-content/support-content/documentation/product%20manuals/en-us/enterprise/Savvio/10K.5/100628563f.pdf,
http://www.spectra.com/pdfs/lto_ultrium.pdf,
https://www.spectralogic.com/index.cfm?fuseaction=home.displayFile&DocID=2513

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, http://highscalability.com/blog/2014/2/3/how-google-backs-up-the-internet-along-with-exabytes-of-othe.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:

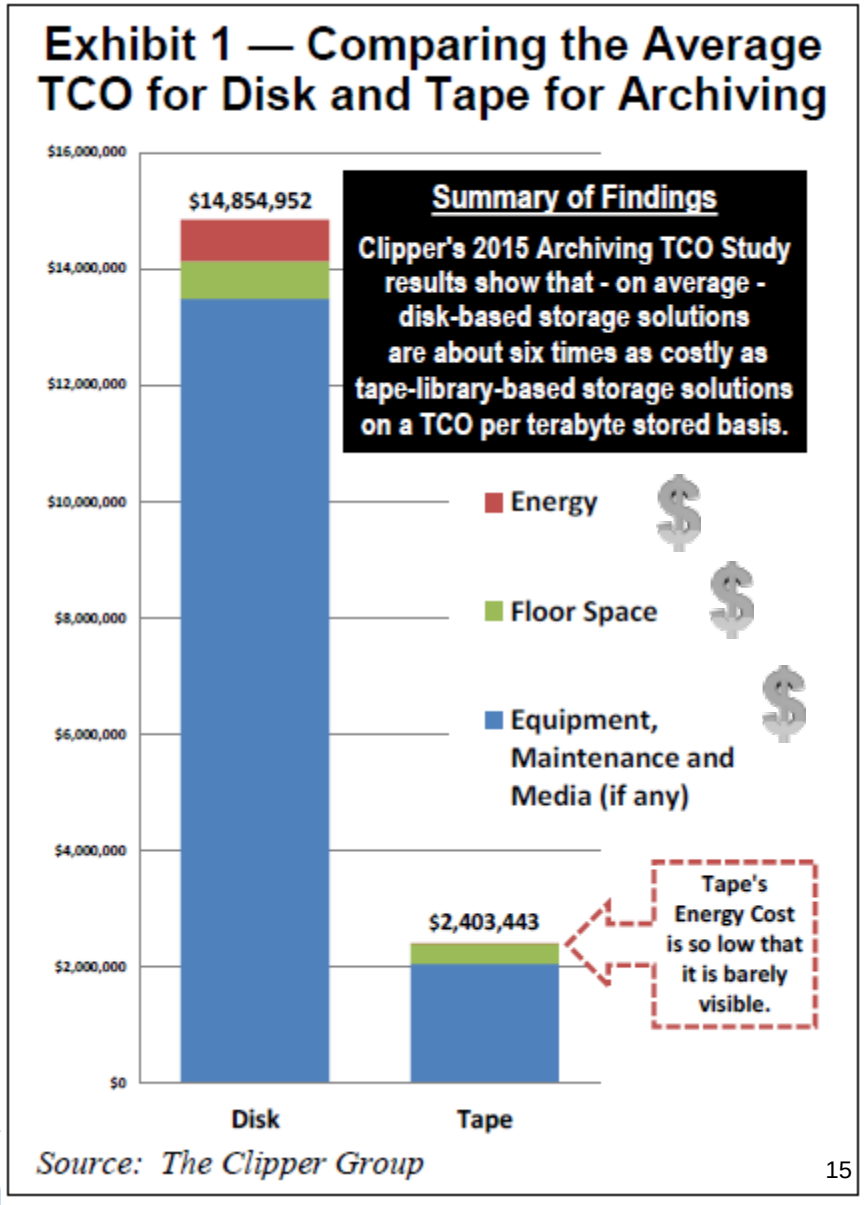
SATA Disk vs. LTO Tape Library

- 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)

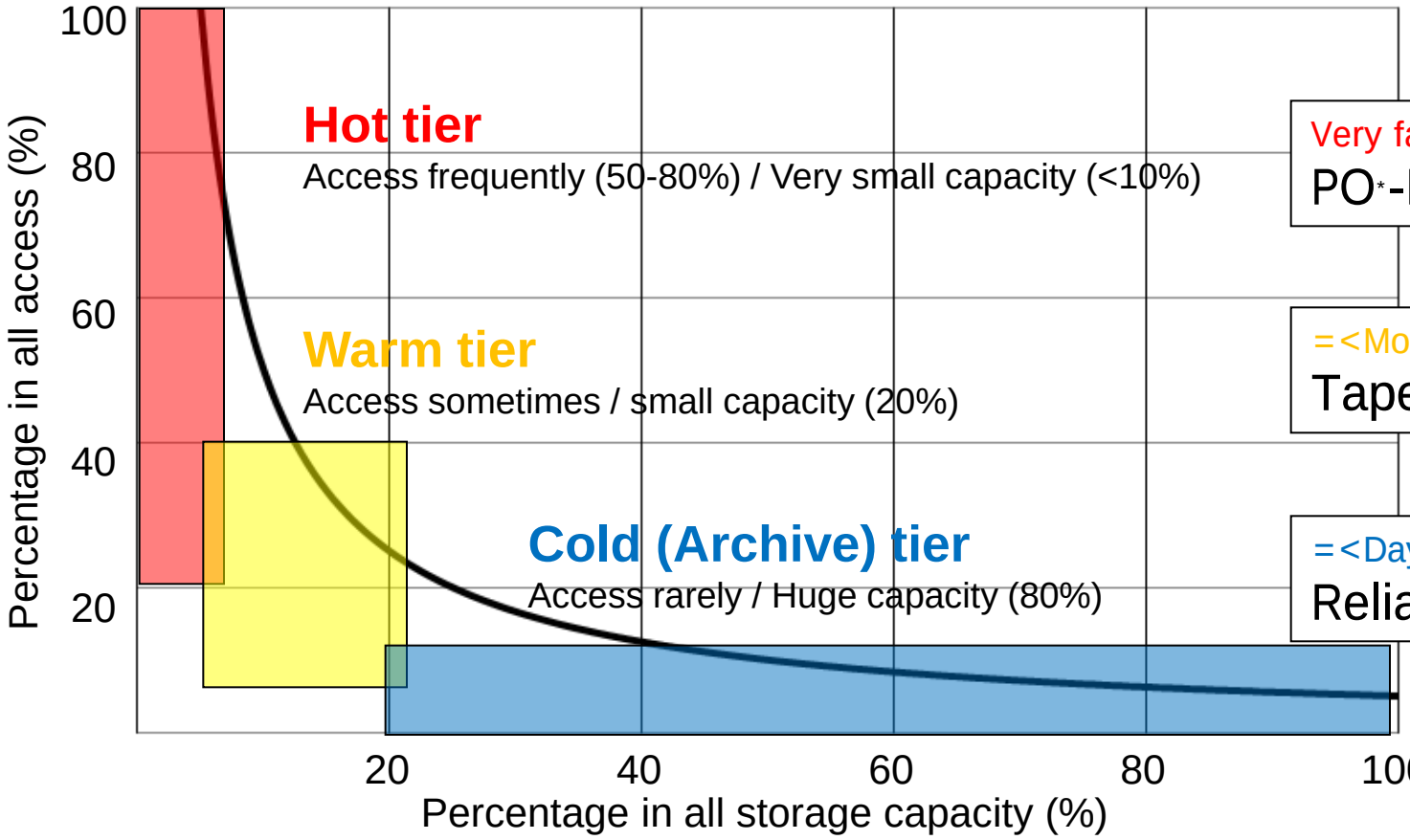
The Results

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

Source: The Clipper Group, 2015
http://www.clipper.com/research/TCG2015006.pdf



Tape's new role for cold data



Very fast access time
PO*-HDD to Flash

=<Moderate access time
Tape→CO**-HDD

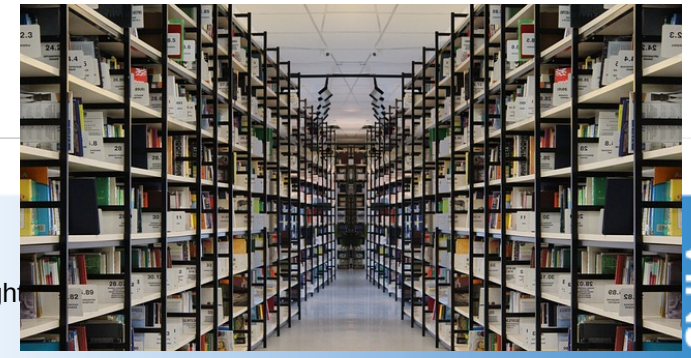
=<Days access time
Reliable/Inexpensive

*PO: Performance Optimized
**CO: Capacity Optimized

Backup

Mon Tue Wed Thu Fri Sat

Archive



Tape use case -Salesforce.com-

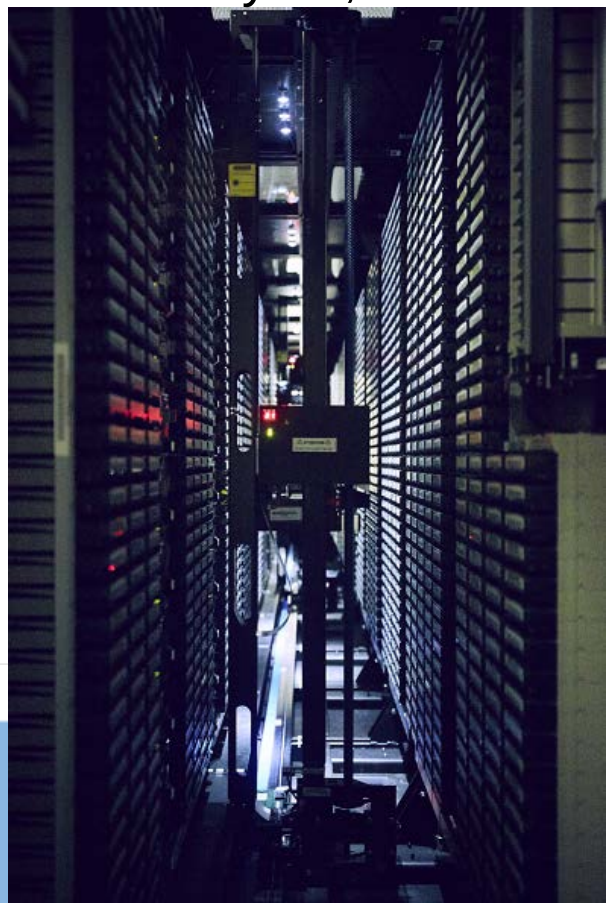
- Salesforce is a leading company in SaaS industry. Tape backup is utilized as a last resort.
- They *backup all data to tape* at each data center, on a rotating schedule of incremental and full backups.
- The backups are *cloned over secure links to a secure tape archive at different facilities.*
- 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://help.salesforce.com/apex/HTViewSolution?id=000005412&language=en_US
https://help.salesforce.com/apex/HTViewSolution?urlname=Data-Recovery-Service-and-Cost&language=en_US
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?

Research Institute	CERN	NCSA
	KEK	NERSC
	NASA	NIH
	QBI	ORNL
Media and Entertainment	Endemole	MLB
	Fox Sports	Red Bull Media House
	Los Angels Lakers	NBC
	NHK	Fuji TV
Meteorology	German Weather Service	Météo France
	Deutscher Wetterdienst	ECMWF
Engineering & Manufacturing	Bombardier	Qualcomm
Cloud-based Service Provider	Google	T3 Media
	Salesforce.com	NetSuite
	Oracle	IBM

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!!

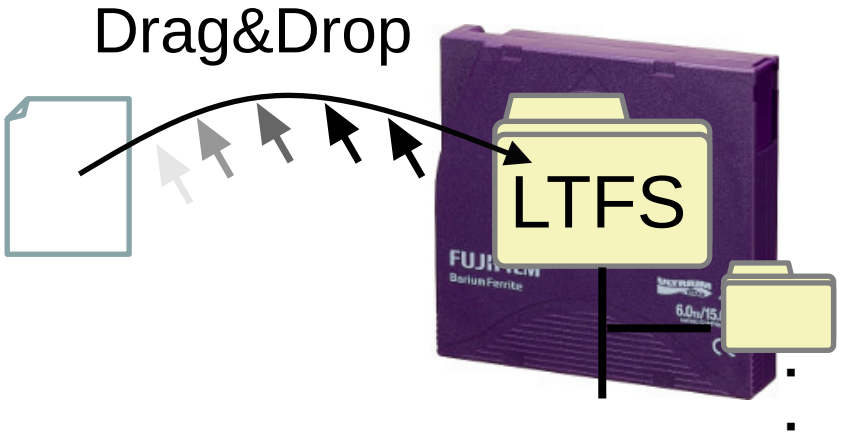
[Quick abstract in 2014 and 2015]

Year	Moth	Topics
2014	Mar	Fujifilm announced Dternity a storage and archiving solution with leveraging tape technologies. (Cloud, Tape-NAS, Media vault)
	May	Sony Sputtering Tape Technology Able to Store 185TB in LTO Cartridge
		IBM and Fujifilm answers to Sony, Native 154TB into LTO cartridge
	Sep	Sept, 2014 IBM announced new LTFS(*) Library Edition (LE) version extends tape library support to third-party tape libraries.
		New LTO roadmap extends to generation 10 (48TB / 1,100MBps)
	Oct	IBM TS1150 released (10TB / 360MBps)
2015	Mar	Oracle updated its HSM with LTFS support and providing an OpenStack swift interface
	Apr	IBM and Fujifilm demonstrated 220TB tape cartridge
	Dec	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) LTO generation 7 released (6TB / 300MBps)

***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/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

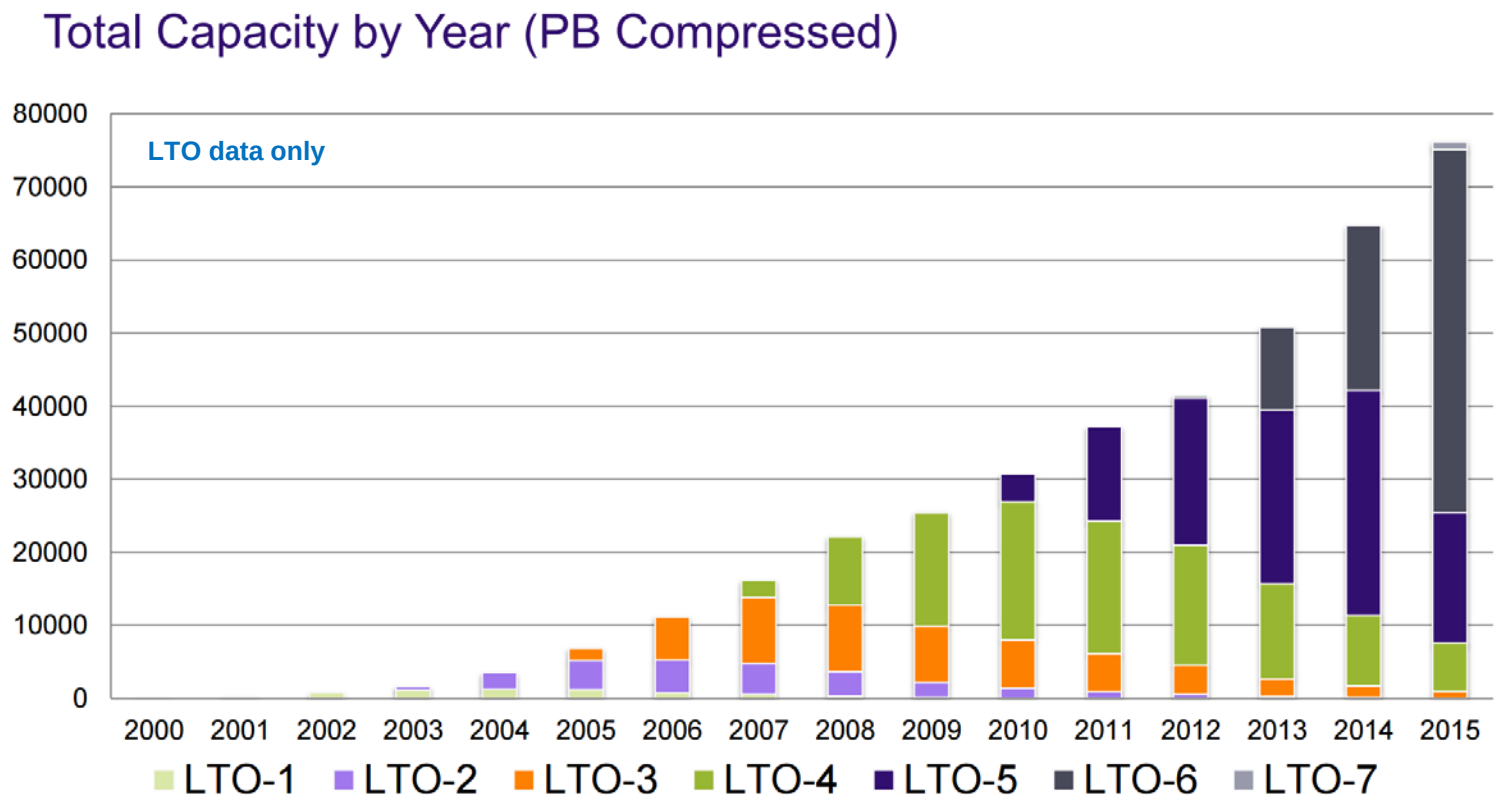
LA Kings has selected Tape NAS for active-archive and deep archive.

dternity

http://www.activearchive.com/sites/default/files/AA%20Case%20Study%20LA%20Kings.pdf

Demand for tape

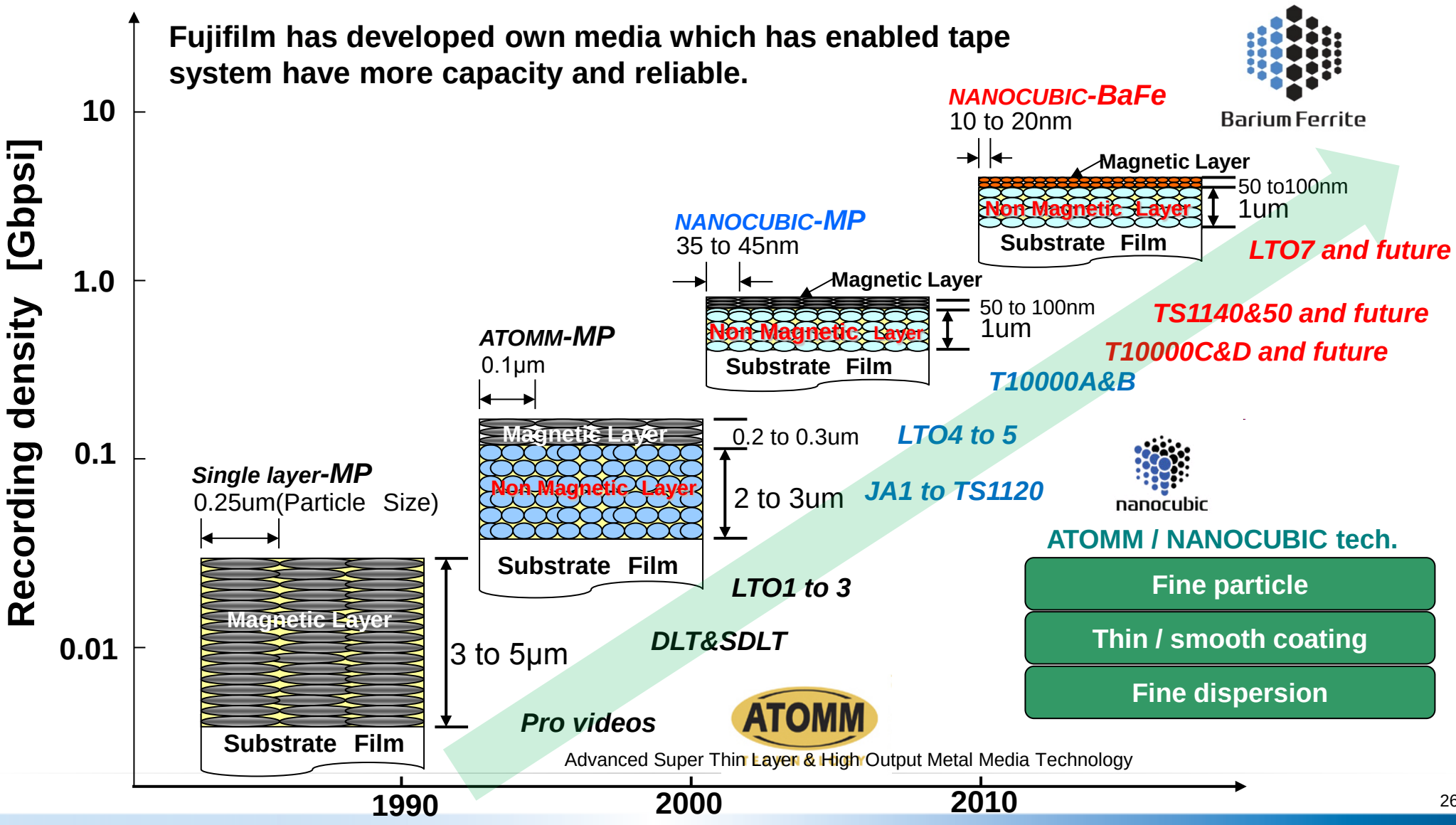
- **Demand for tape has continued to increase!**



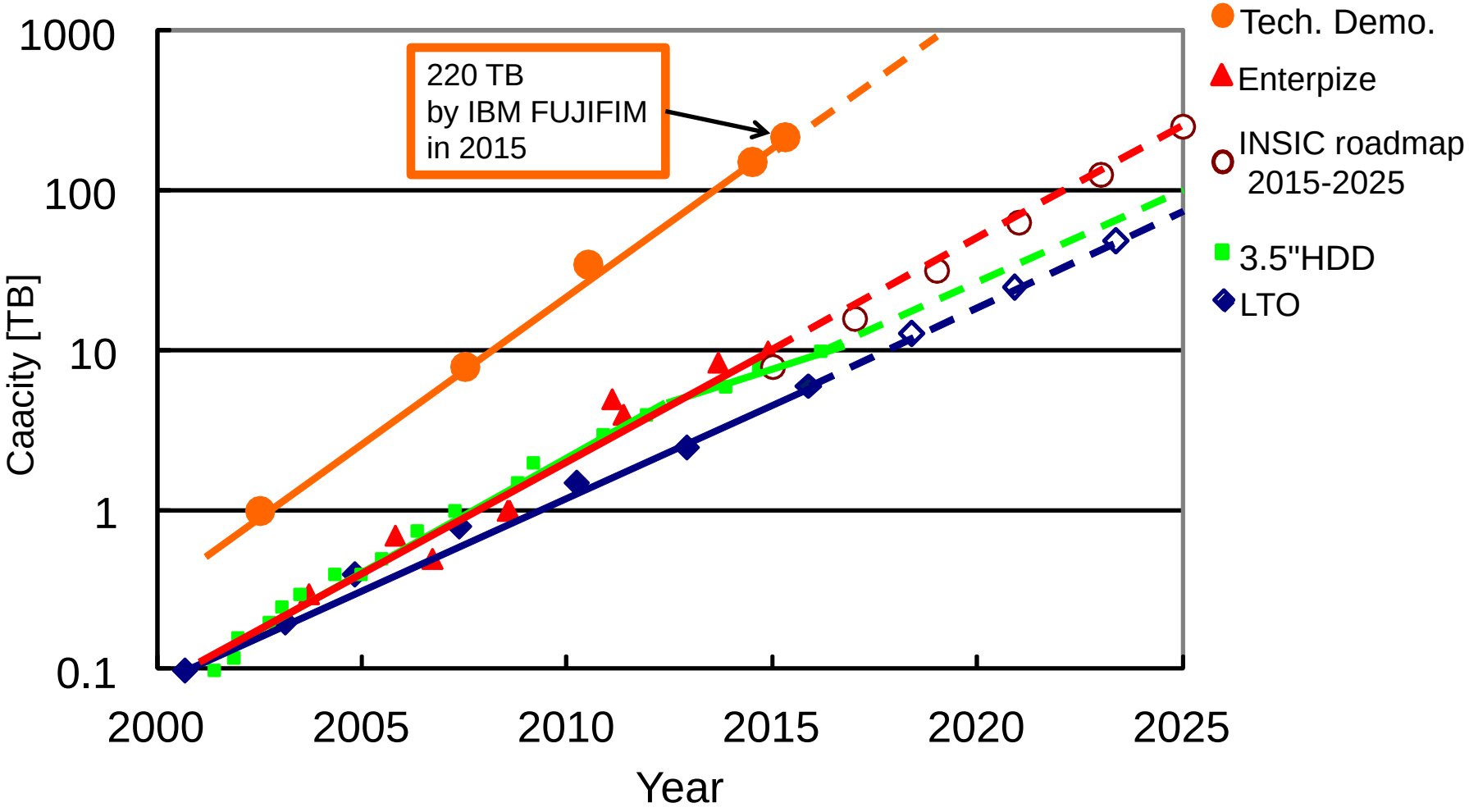
Source: http://www.lto.org/wp-content/uploads/2016/03/LTO_Media-Shipment-Report_3.22.16.pdf

Tape Technologies

Fujifilm media development history



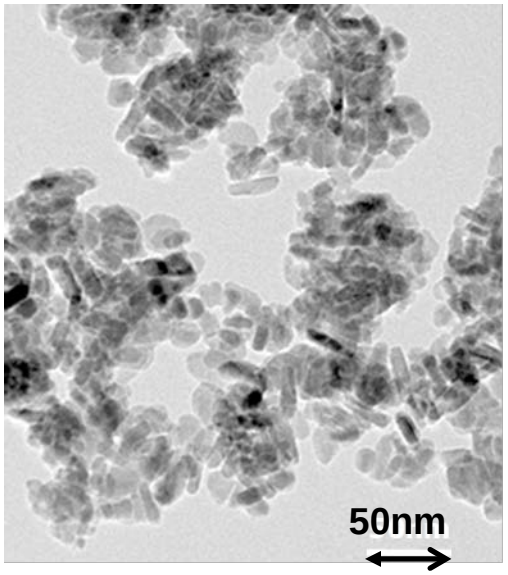
Cartridge Capacity trends



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.)
http://www.insic.org/news/2015%20roadmap/15_25roadmap.html
http://idema.org/wp-content/plugins/download-monitor/download.php?id=2456

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.



Barium Ferrite

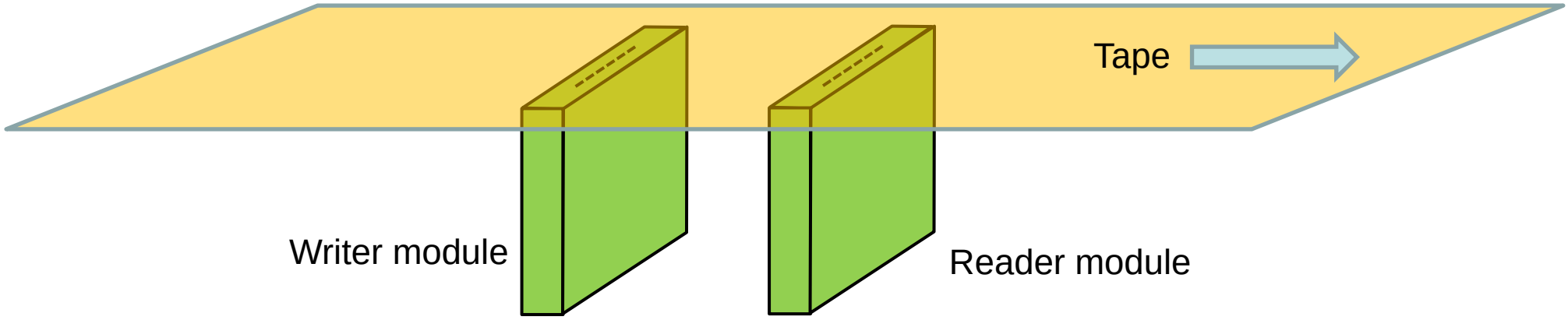
Long term reliability of Tape system

- 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.

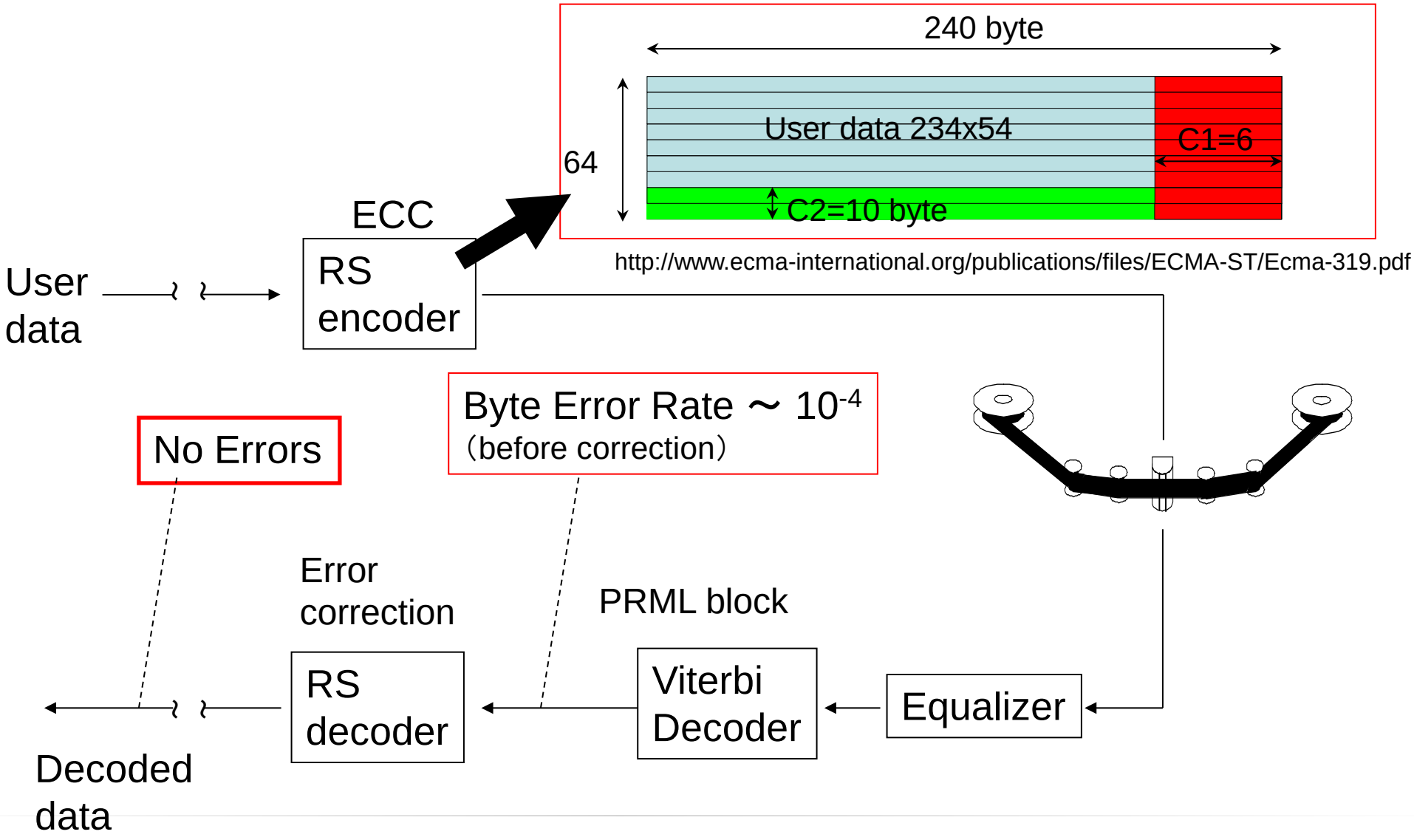
Read-after-write verification

- 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



Examples of Redundancy array

D_1 D_2 D_3 D_4 P

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

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)

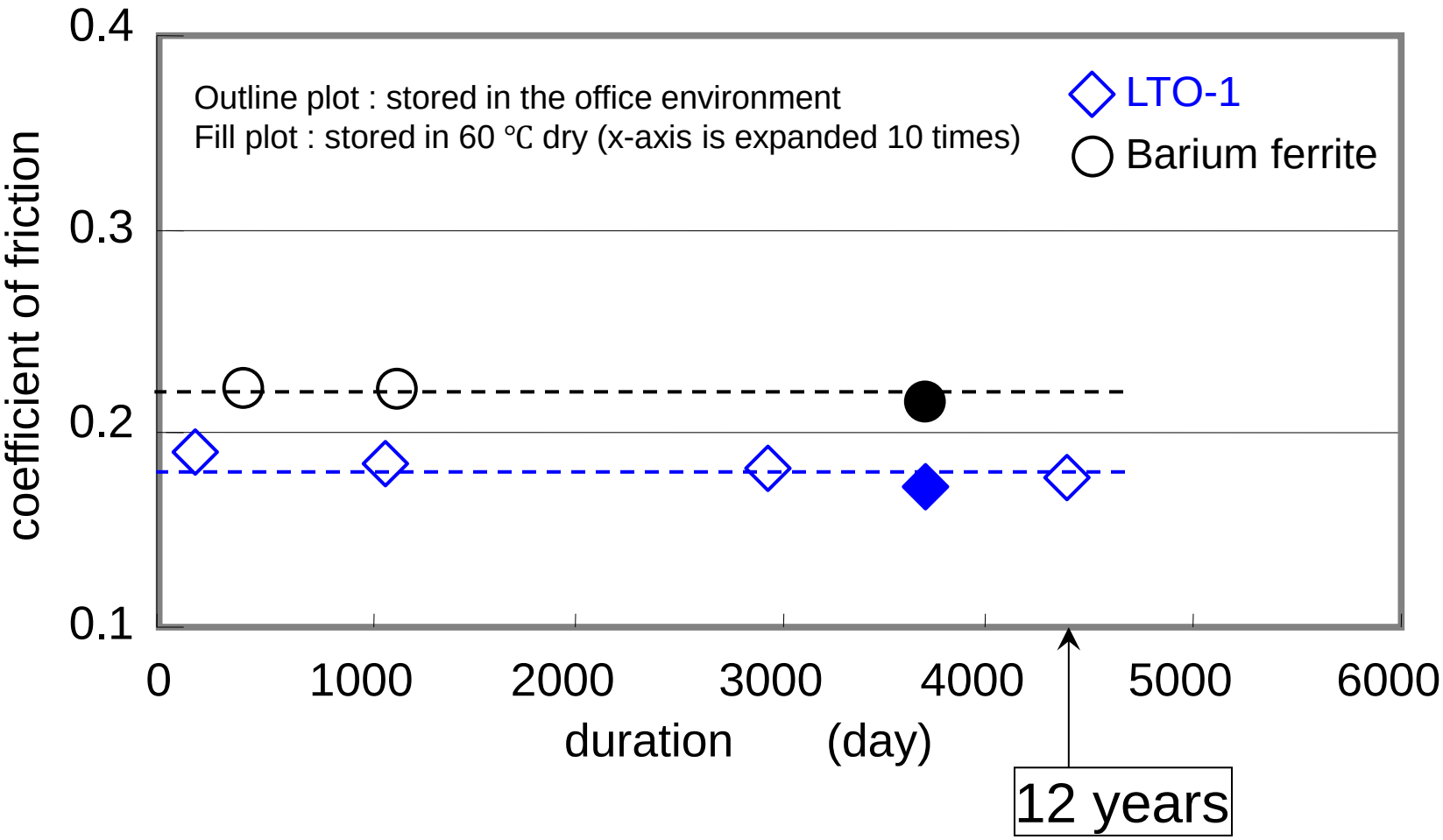
https://www.youtube.com/watch?v=eNliOm9NtCM

http://highscalability.com/blog/2014/2/3/how-google-backs-up-the-internet-along-with-exabytes-of-othe.html

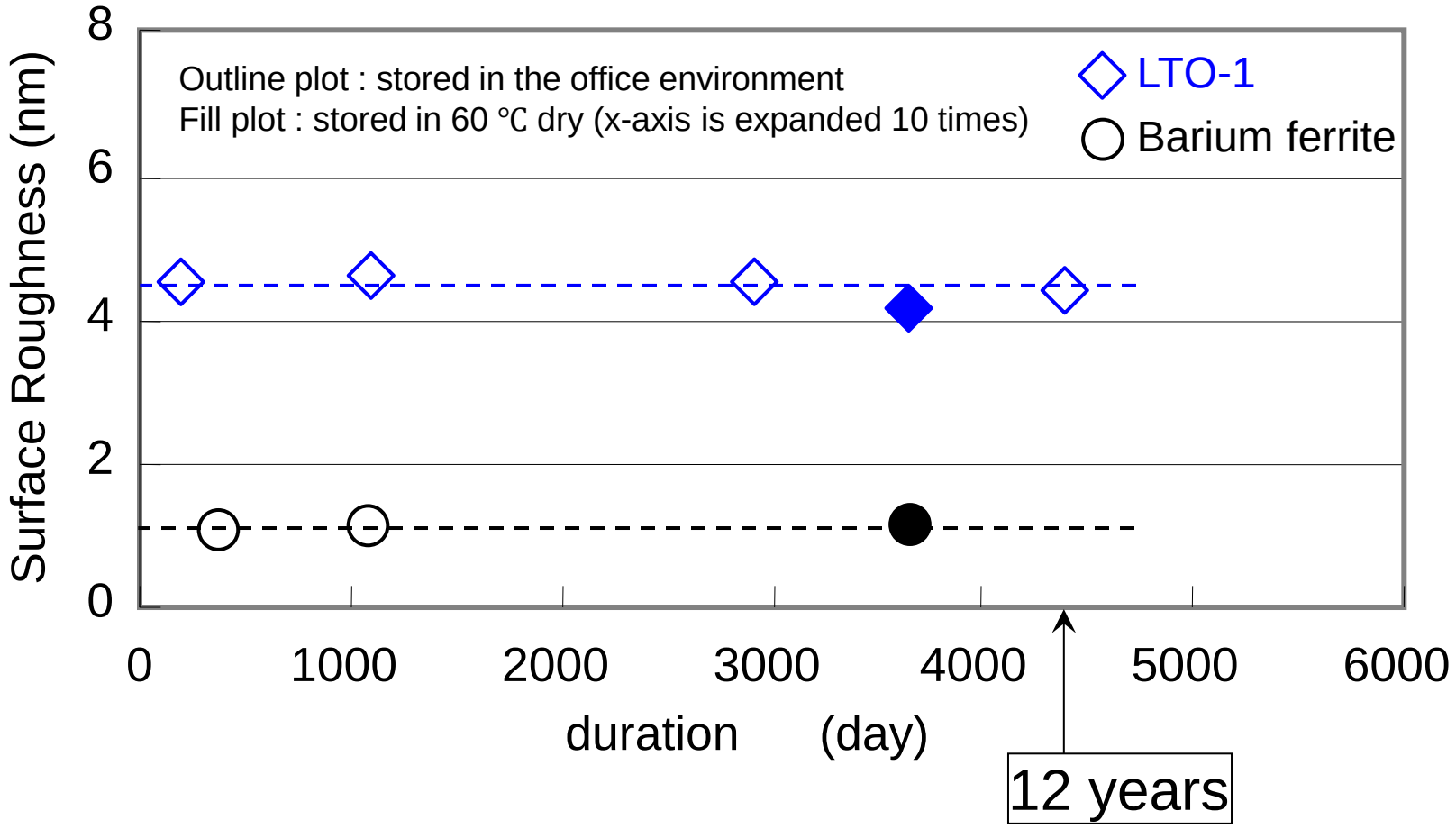
or Redundancy Array of Independent Tape (RAIT)

http://www.activearchive.com/common/pdf/AA-Case-Study-NCSA.pdf

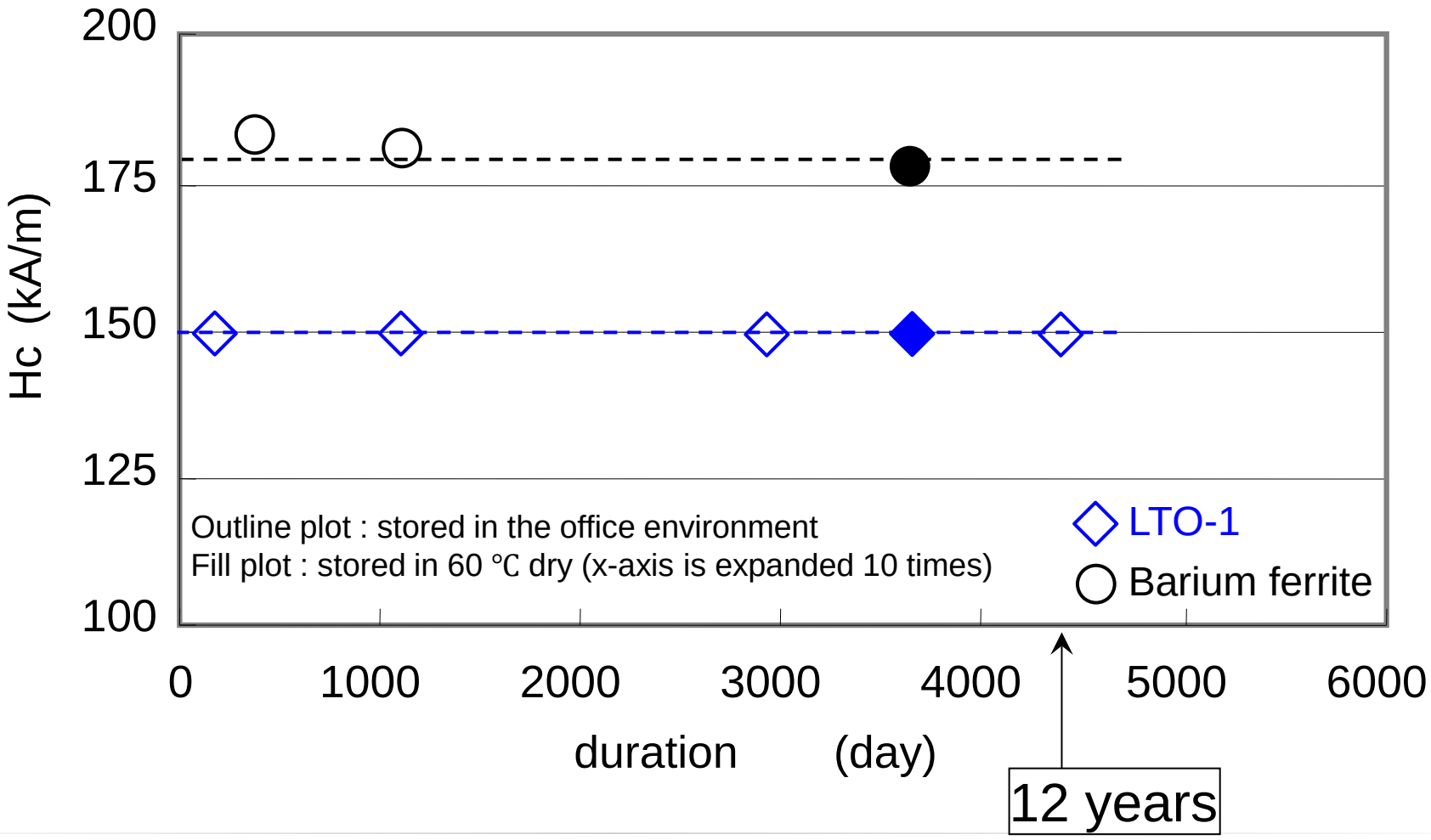
Chemical stability of tape media



Physical stability of tape media



Magnetical stability of tape media



Future Tape prospect

Tape used in 123 Gbpsi demonstration

PROPERTIES OF MAGNETIC PARTICLE AND EXPERIMENTAL TAPES

Media type	New BaFe tape	TS1150 JD media
<Magnetic particle>		
Volume (nm^3)	1600	1950
Coercivity (kA/m)	223	190
σs ($A \cdot m^2/kg$)	45	50
<Media>		
Magnetic properties		
Longitudinal direction		
Coercivity (kA/m)	146	182
SQ	0.24	0.39
Perpendicular direction[1)]		
Coercivity (kA/m)	263	214
SQ	0.87	0.66
Surface roughness		
Optical interferometry[2)]		
Ra (nm)	0.9	1.6
AFM[3)]		
Ra (nm)	1.8	2.0
Rz (nm)	27	34
Coefficient of friction[4)]	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

M. Lantz et al., "123 Gbit/in2 Recording Areal Density on Barium Ferrite Tape," *IEEE Trans. Magn.*, vol . **51**, #3101304

Tape used in 123 Gbpsi demonstration

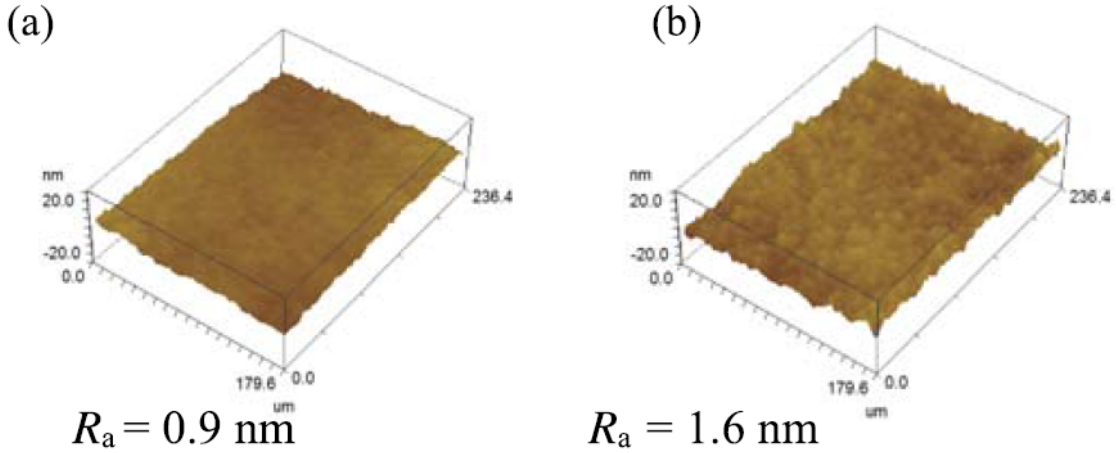


Fig. 1. Surface profile images measured by optical interferometry. (a) New BaFe tape. (b) TS1150 JD media.

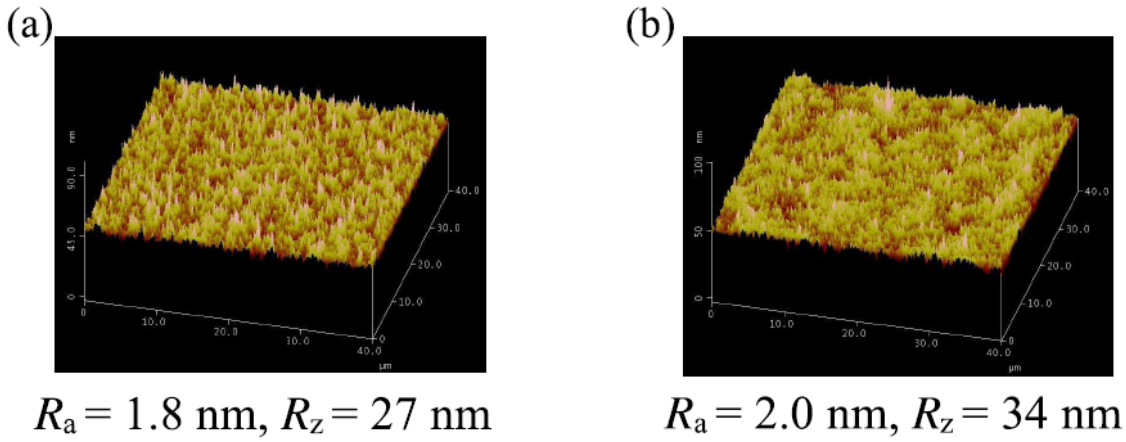
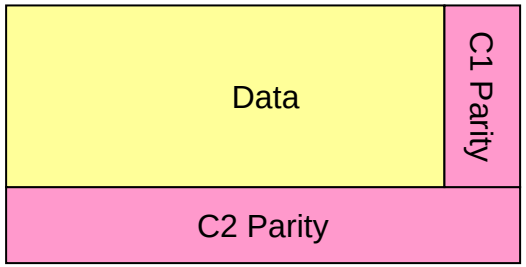
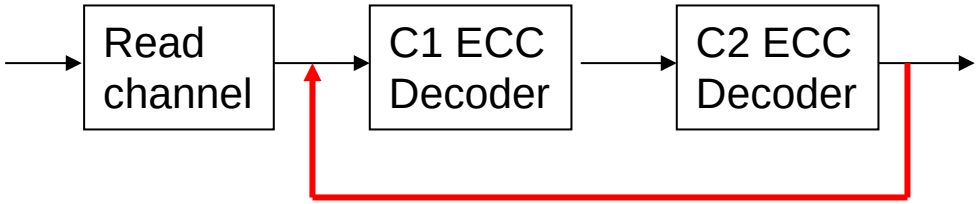
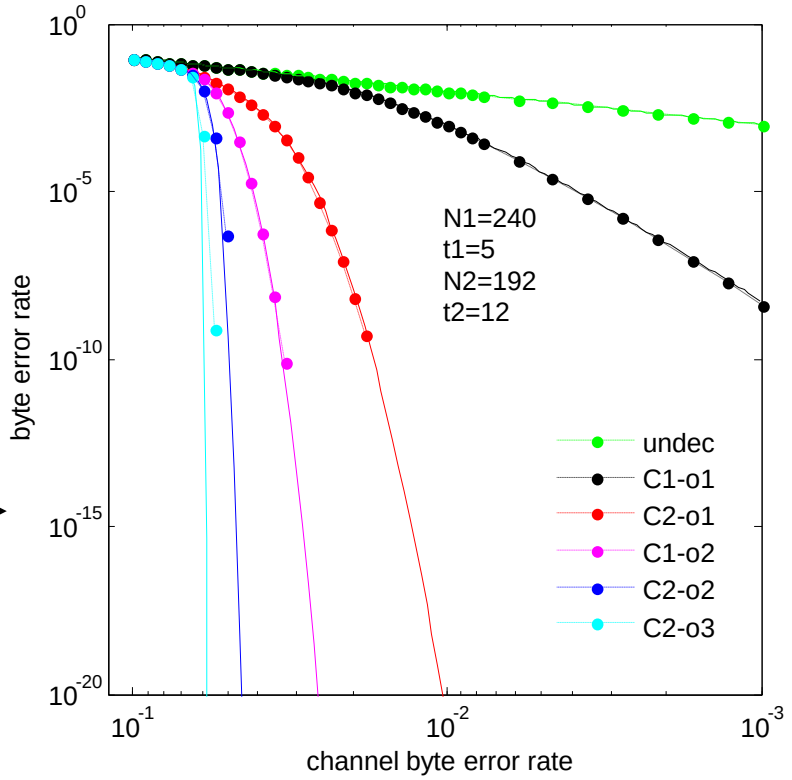


Fig. 2. Surface profile images measured by AFM. (a) New BaFe tape. (b) TS1150 JD media.

ECC used in 123 Gbpsi 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