The Future of Tape

Mark Lantz

Distinguished Research Scientist

Manager Tape Storage Research

IBM Research - Zurich

Outline

- The Future of Tape
  - Tape's renaissance
  - Scaling challenges in magnetic recording
  - 317 Gb/in^2 Tape Areal Recording Demonstration
  - Future scaling potential of tape

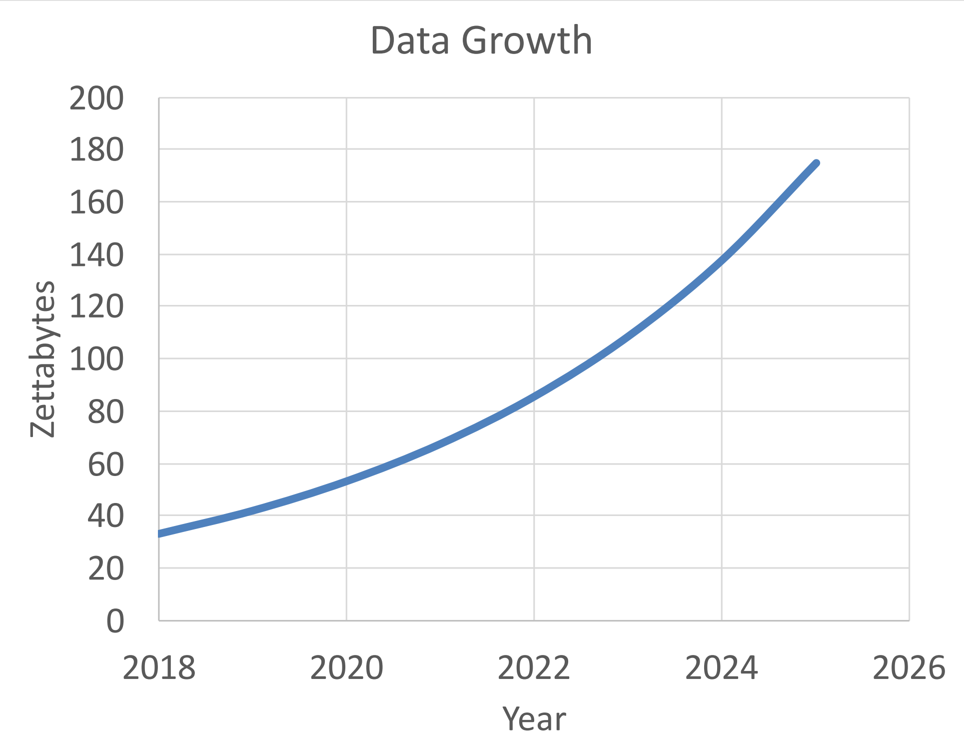
- The Future of HDD
  - Volumetric scaling
  - HAMR
  - MAMR

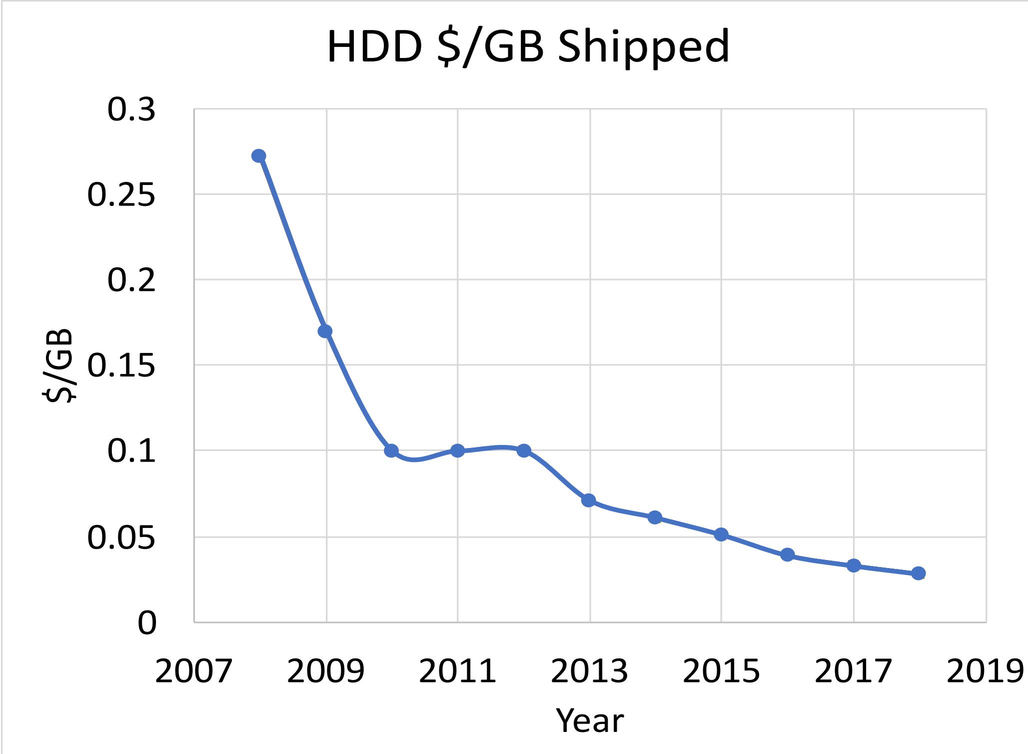
- Alternative Archival Storage Technology
  - DNA storage
  - Optical

- Summary

Tape's Renaissance

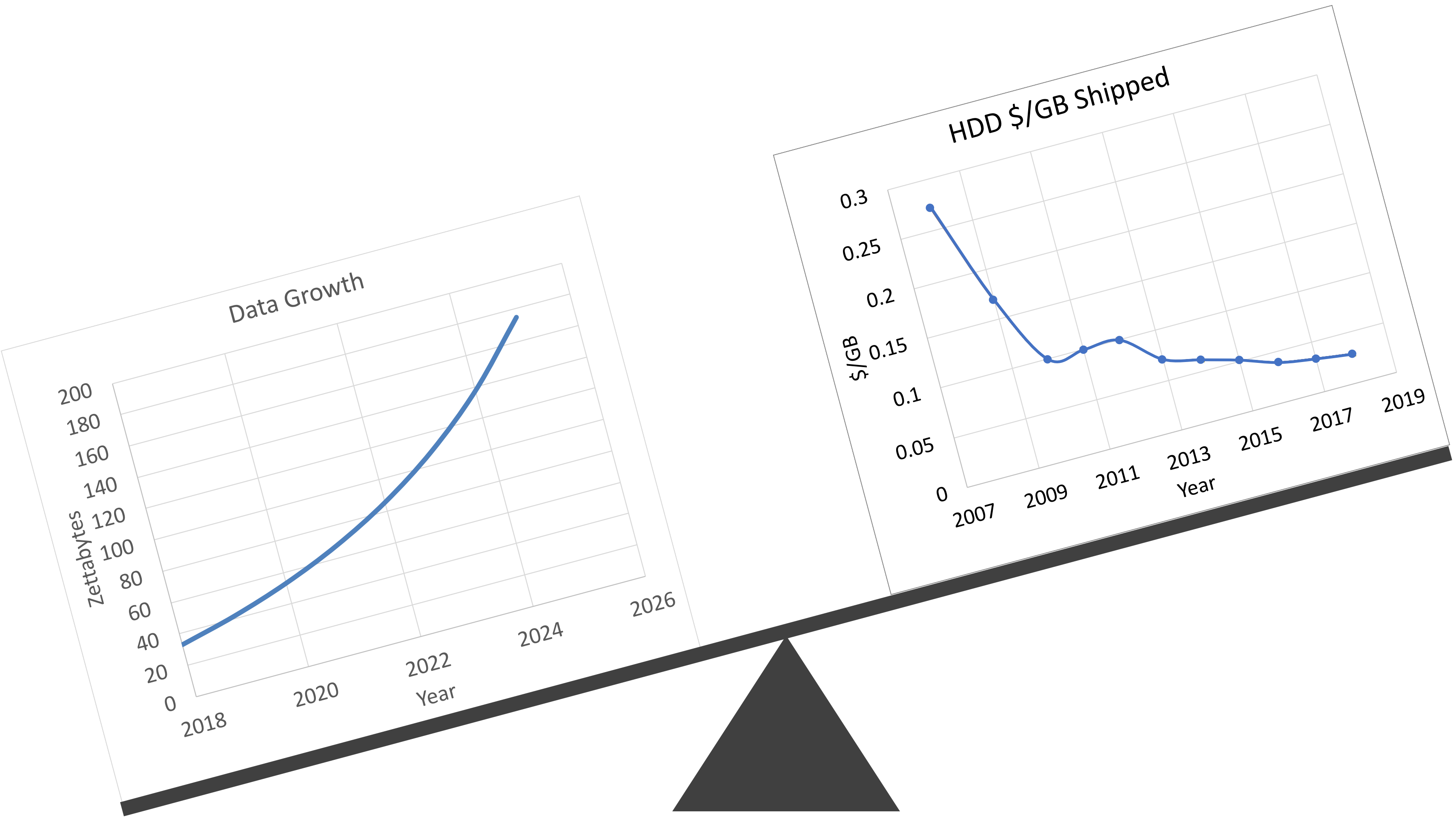
Data is growing exponentially while HDD scaling has stagnated



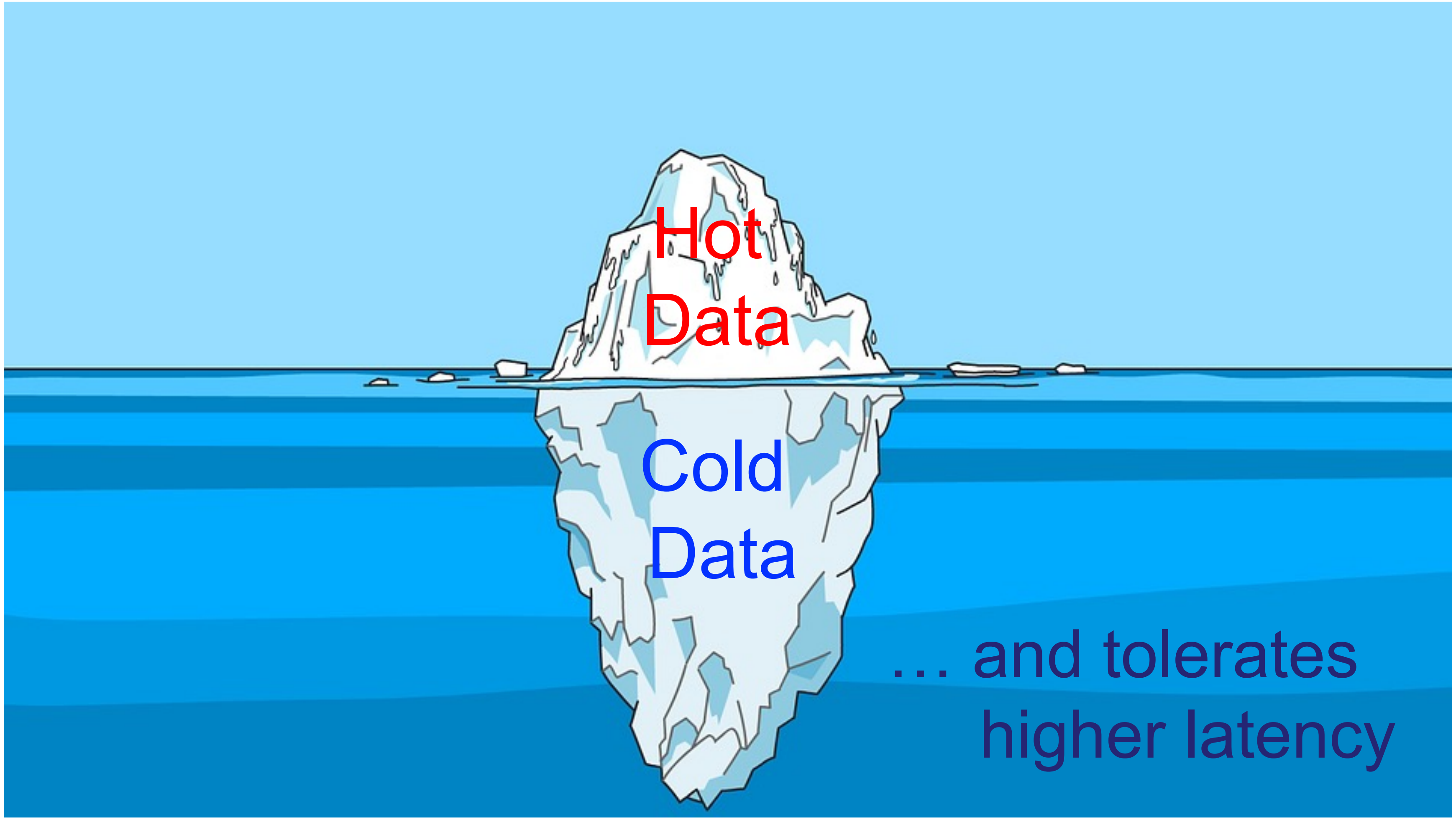


Refs: Data growth projections from IDC White Paper – #US44413318
Robert Fontana and Gary Decad, AIP Advances 8, 056506 (2018); https://doi.org/10.1063/1.5007621

The Data Center is Out of Balance



A large fraction of data is cold…



Times change!

2006

*”Tape is **dead**, Disk is Tape, Flash is Disk, RAM locality is king”*

Microsoft

2015

*”All cloud vendors will be using tape and will be using it at a **level never seen before**”*

Microsoft

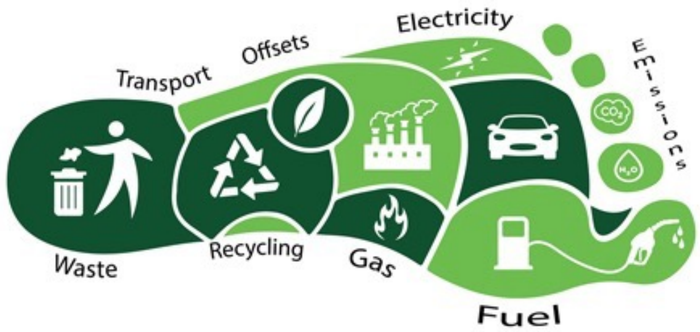
http://research.microsoft.com/en-us/um/people/gray/talks/flash_is_good.ppt
http://www.infostor.com/disk-arrays/how-do-you-store-a-zettabyte.html

Why is tape exciting again?

Cost Savings

@ hyperscale HDD Cost 3.7x Tape*

Sustainability

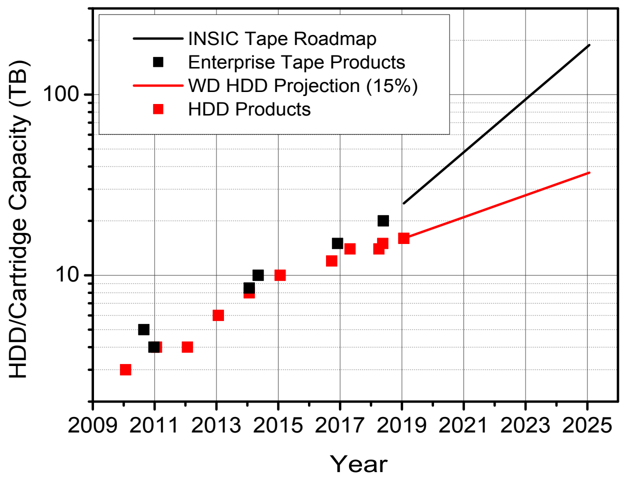


27PB Tape Archive has 96% Lower CO_2e than HDD

Security

Airgap, Encryption
Quantum Safe

Scaling



INSIC Tape Capacity Scaling
40% CAGR to 2029

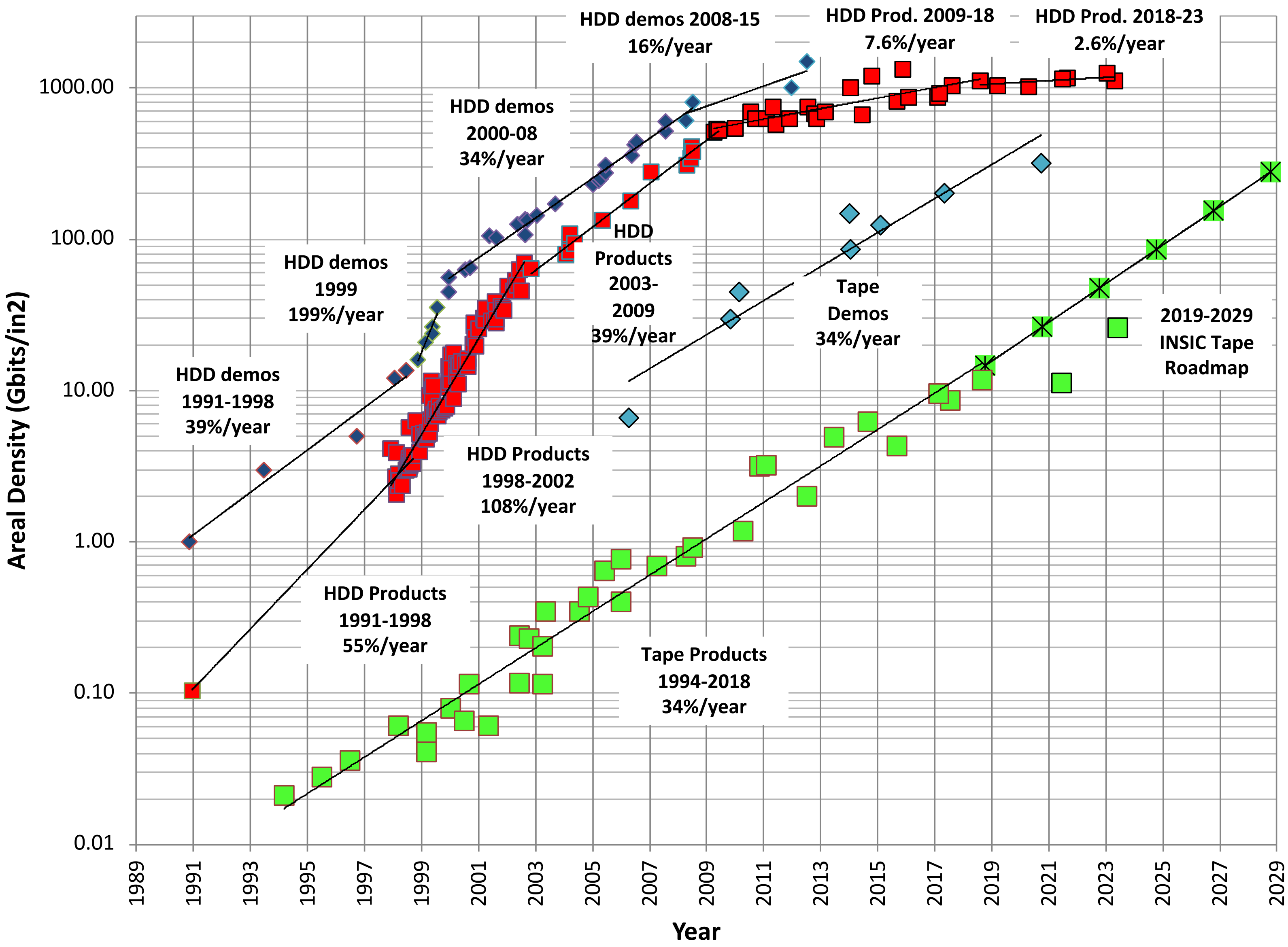
(*Reference: M.S. Azure 2016)

Tape Scaling

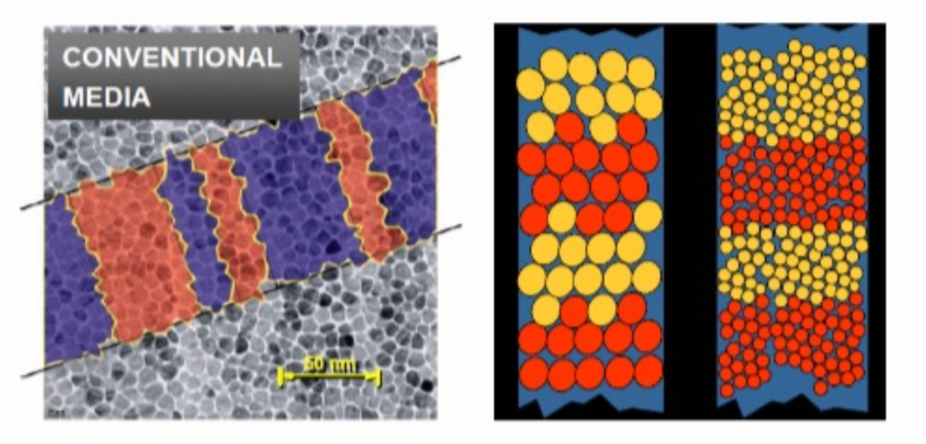
Product Year	IBM 726 1952	LTO9 2021	**TS1170** **2023**	Demo 2017 Sputtered Tape	Demo 2020 SrFe Tape
Capacity	2.3 MB	18 TB	**50 TB**	330 TBytes	580 TBytes
Areal Density	1400 bit/in^2	11.9 Gbit/in^2	**26.1 Gbit/in^2**	201 Gbit/in^2	317 Gbit/in^2
Linear Density	100 bit/in	545 kbit/in	**555 kbit/in**	818 kbit/in	702 kbit/in
Track Density	14 tracks/in	21.9 ktracks/in	**47 ktracks/in**	246 ktracks/in	452 ktracks/in

Areal Density **>18.6M** x

Areal Density Scaling



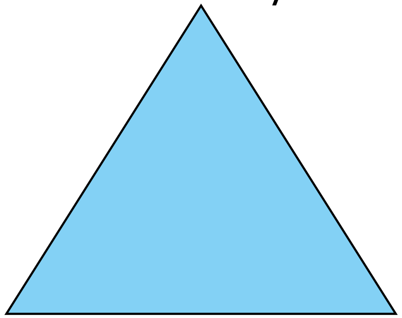
Superparamagnetic Limit



Magnetic Media “Trilemma”:

Small particles (V)

$$SNR \propto 1/V$$



Thermal Stability

$$E_B \propto K_u \cdot V$$

Writability

$$H_0 \propto K_u$$

$$H_0 < Head\ Field$$

HDD has reached the limit of (known) materials to produce larger write fields.

Tape operates at areal densities far from the superparamagnetic limit and has the potential to continue scaling for many years

Technologies to go beyond the superparamagnetic limit:

- Energy Assisted Magnetic Recording (HAMR) / (MAMR)

2023 Storage Bit Cells and Extendibility

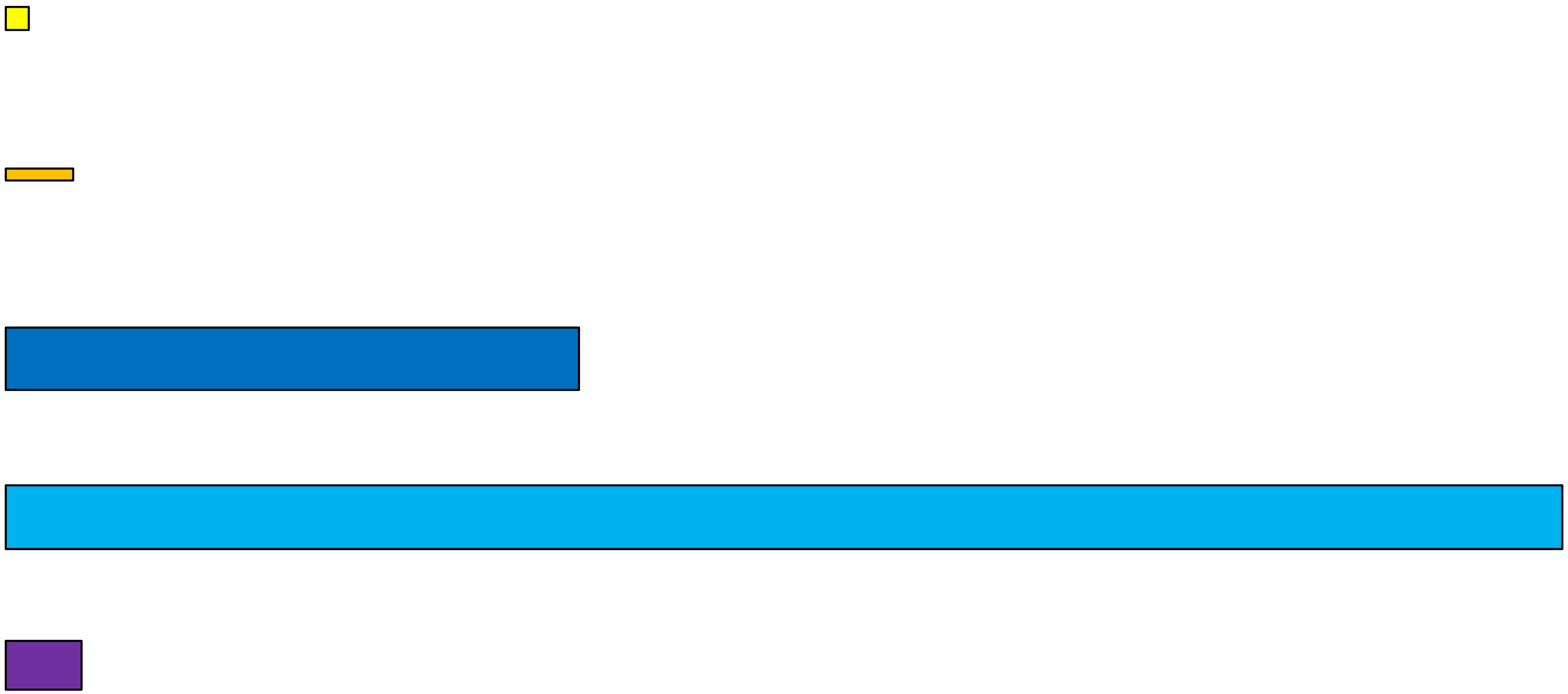
Flash (3 bits)
2150 Gb/in^2
17.3 nm x 17.3 nm

HDD
1260 Gb/in^2
~49 nm x ~10 nm

Jag7 Tape
26.1 Gb/in^2
540 nm x 45.8 nm

LTO9 Tape
11.9 Gb/in^2
1150 nm x 46.6 nm

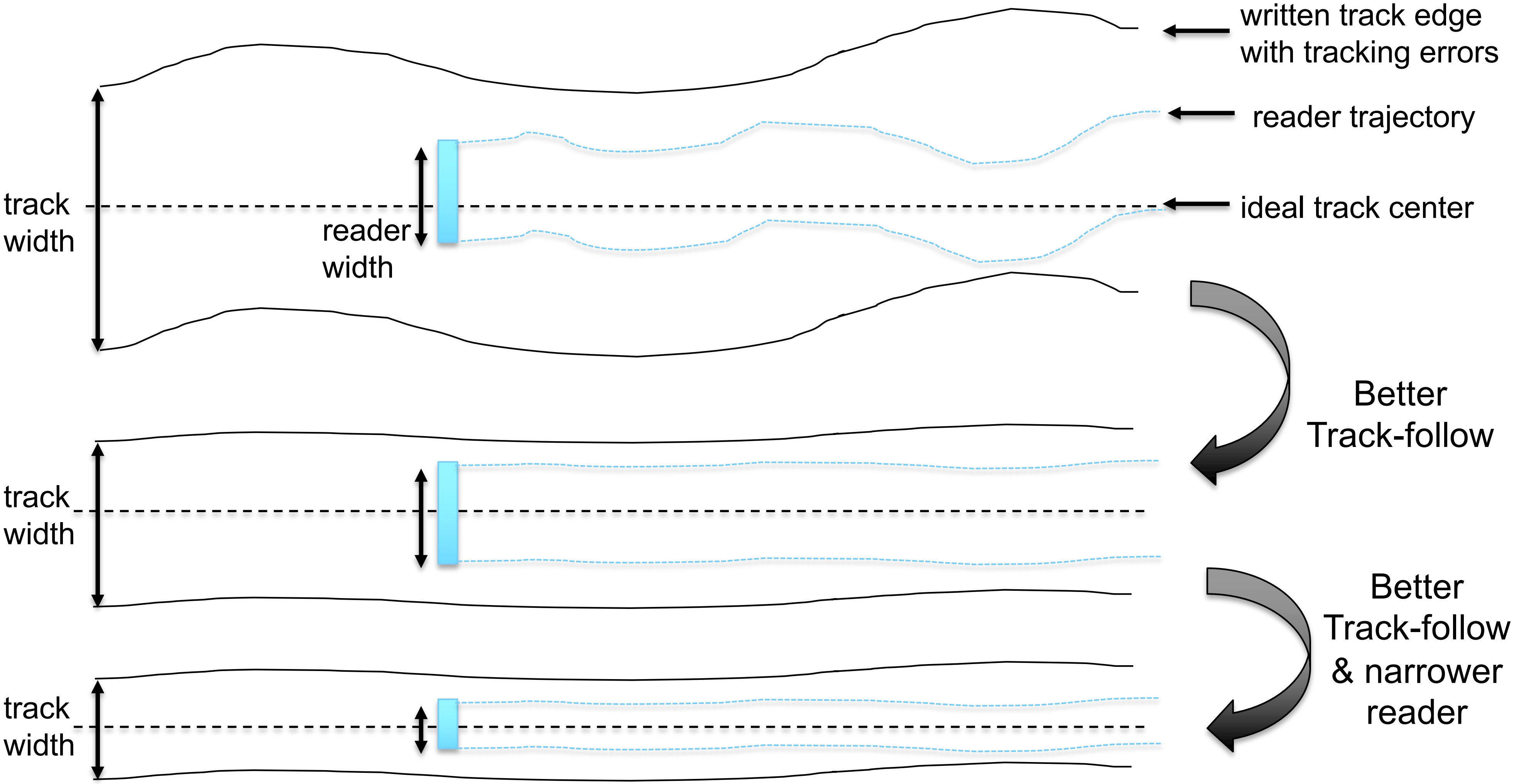
SrFe Demo
317 Gb/in^2
56.2 nm x 36.2 nm



→Most potential for future scaling of tape track density

Track density and servo control

The track following servo system enables us to measure and control the position of the read/write heads on the tape in order to write the data tracks in the correct location and then scan the read transducers along the center of the tracks during read-back operation.



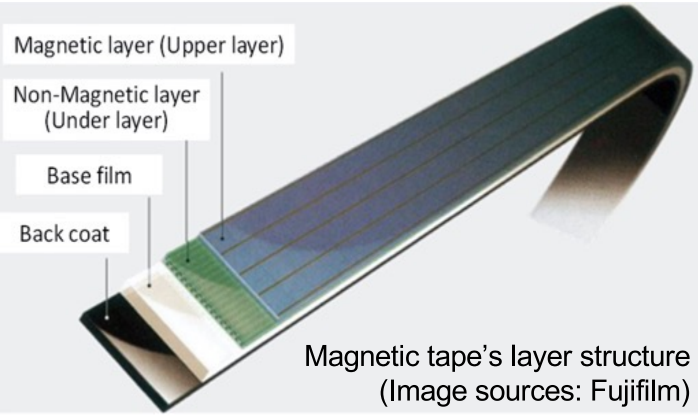
Scaling Strategy

Focus on track density scaling with modest linear density scaling

- For track density scaling we need to:
  - Improve track following
  - Reduce the reader width from ~1000nm to 29 nm

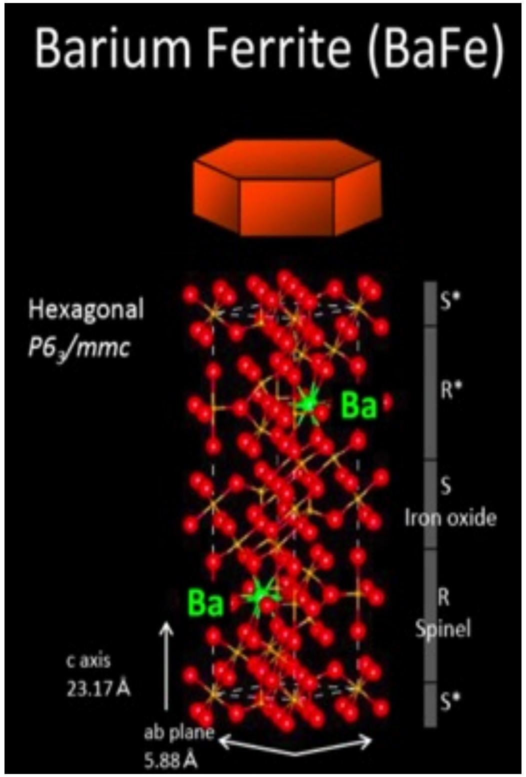
- Ultra narrow reader results in a dramatic loss in read back signal that must be compensated by:
  - **improved media technology** → require improved writer technology
  - improved reader technology
  - improved signal processing and coding

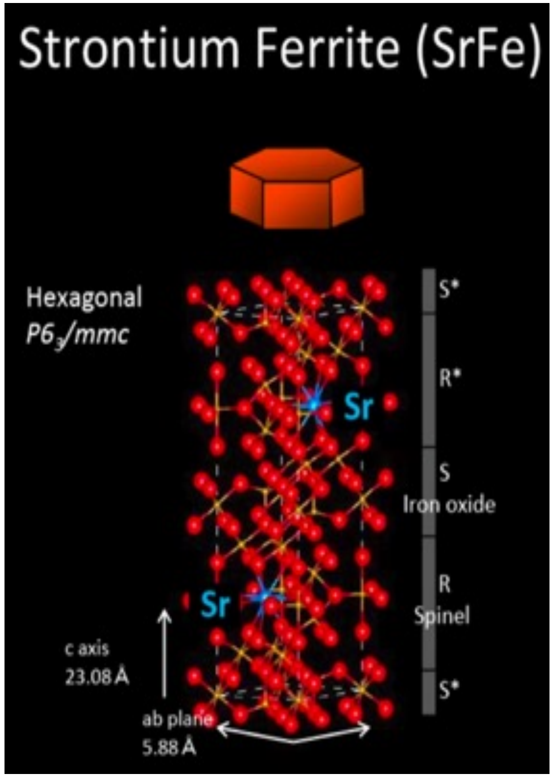
Media Technology: BaFe versus SrFe



Magnetic tape's layer structure
(Image sources: Fujifilm)

BaFe and SrFe are both stable oxides



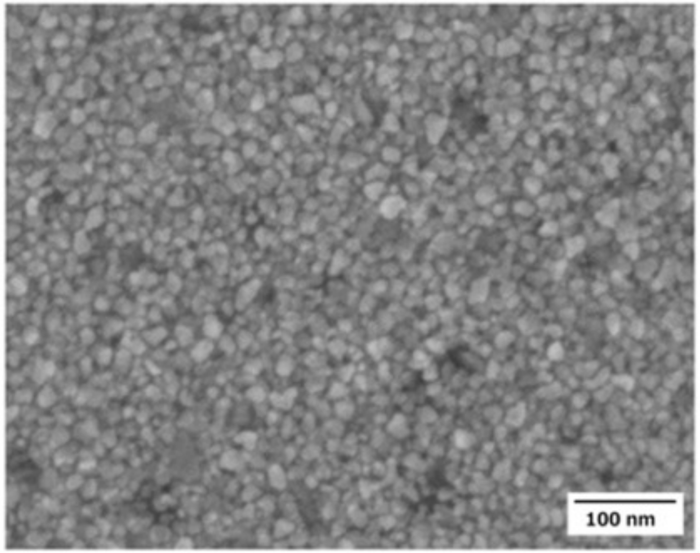


	BaFe	SrFe
M_s (Am^2/kg)	72	92
K_u (10^5 J/m^3)	3.3	3.5

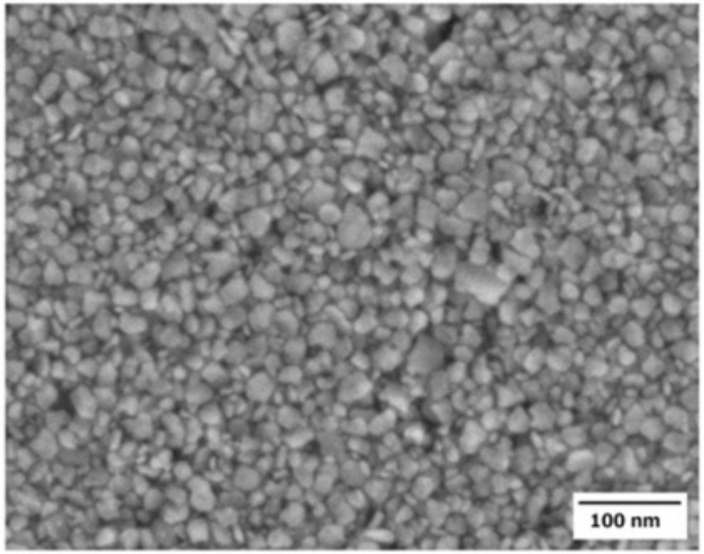
→ Enhanced magnetic properties of SrFe enables scaling to smaller particles

Particle Size and Orientation

New Strontium Ferrite **950 nm³**



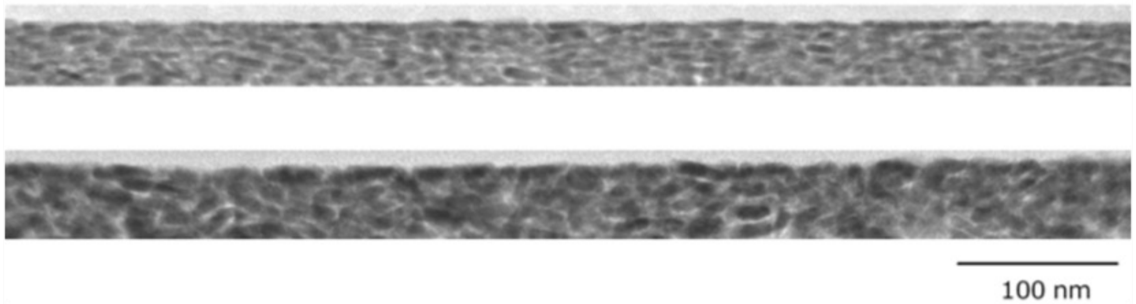
JE Barium Ferrite 1700 nm³

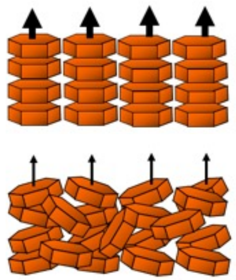


Perpendicular Orientation

New SrFe

JE BaFe

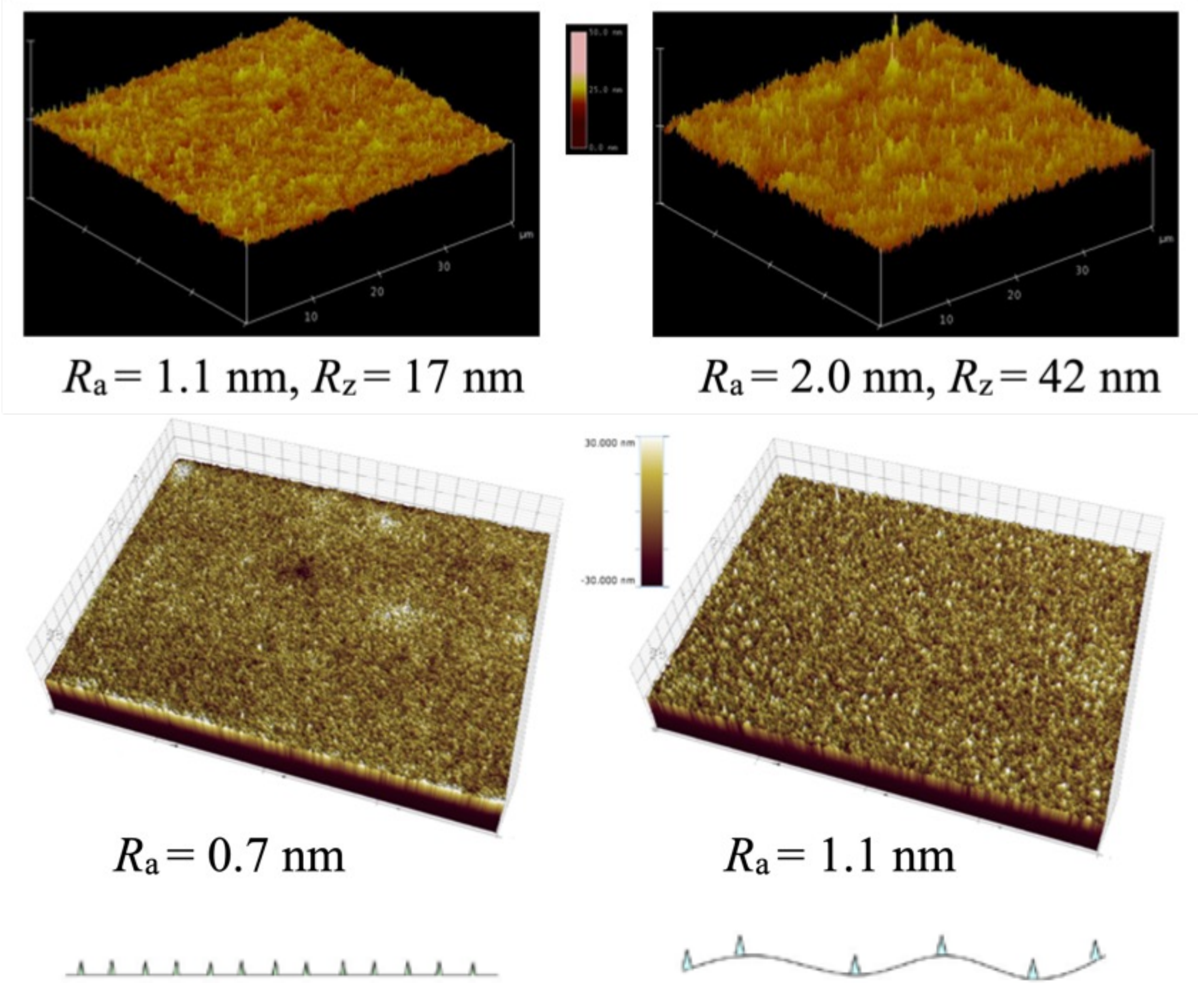




→ Smaller particle size and improved orientation provide more SNR

Smooth Media Surface Technology

New Strontium Ferrite | JE Barium Ferrite



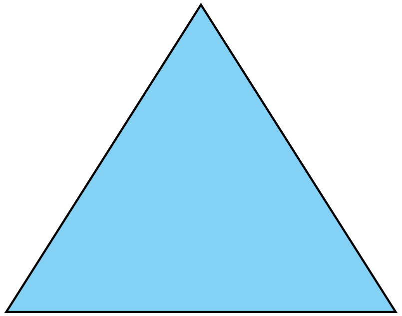
→ Smoother surface reduces spacing and increase SNR, but increases friction

Enhanced Write Field Head Technology

Magnetic Media "Trilemma":

Small particles (*V*)

$SNR \propto V$



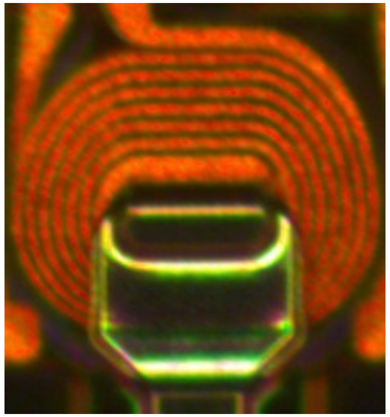
Thermal Stability

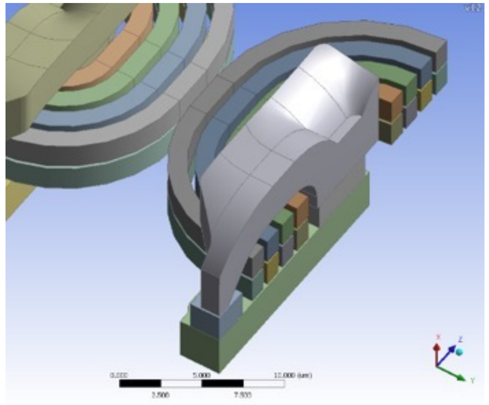
$E_B \propto K_u \cdot V$

Writability

$H_0 \propto K_u$

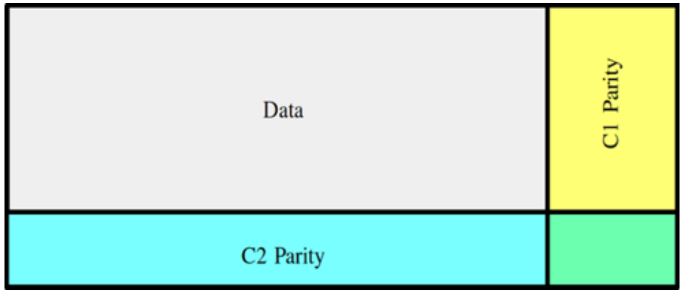
$H_0 < Head\ Field$

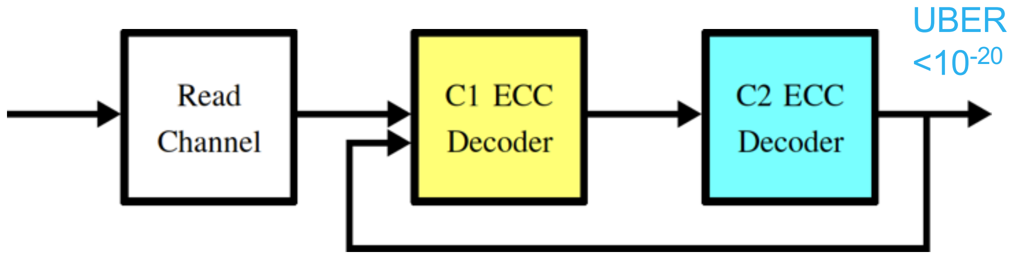




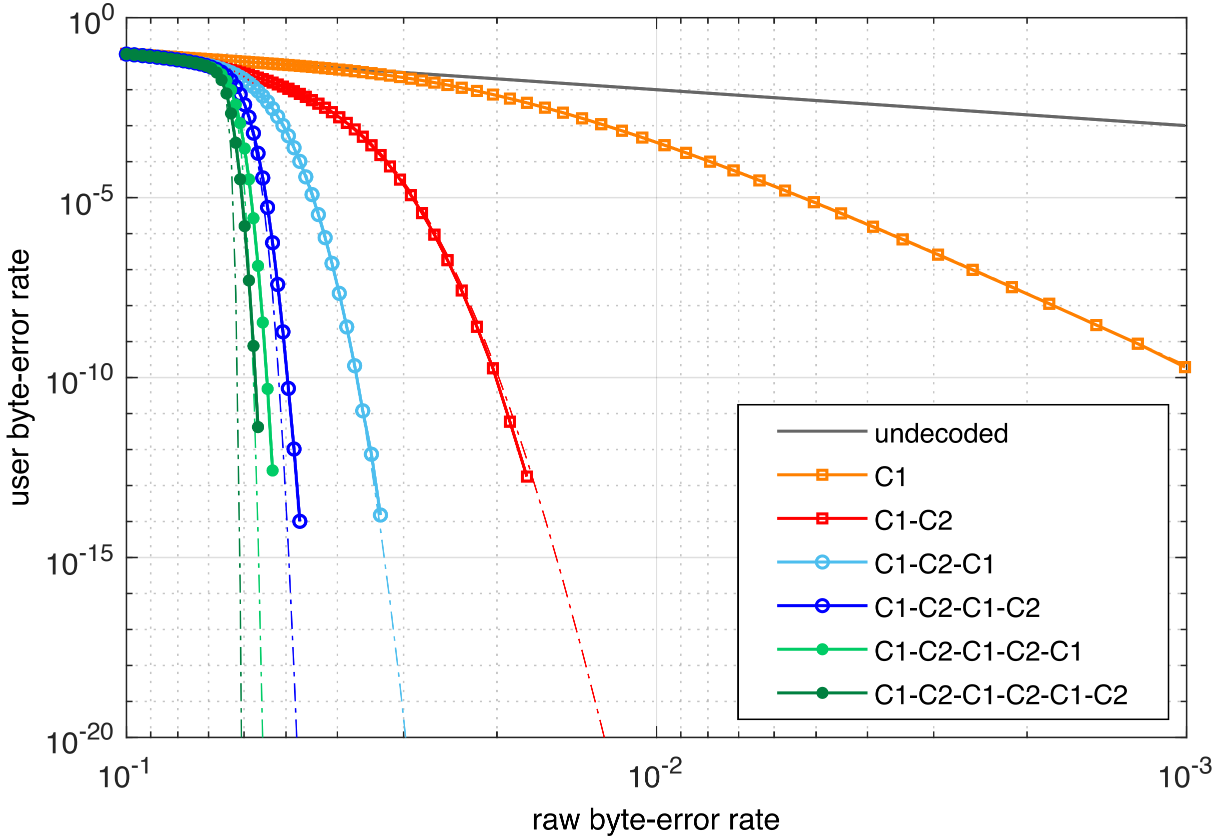
IBM developed a **new high moment (HM) layered pole write head** that produces much larger magnetic fields enabling the use of smaller magnetic particles with higher coercivity

Iterative decoding



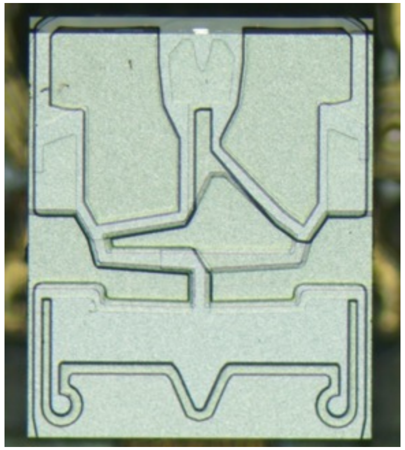


- Iterative decoding provides improved error correction performance with each iteration
- Enables high reliability with less SNR from the head/media
- With two C1-C2 iterations we achieve an **uncorrectable byte-error rate of 10^{-20}** with a byte error rate of $\approx 4.5 \cdot 10^{-2}$ at the output of the detector

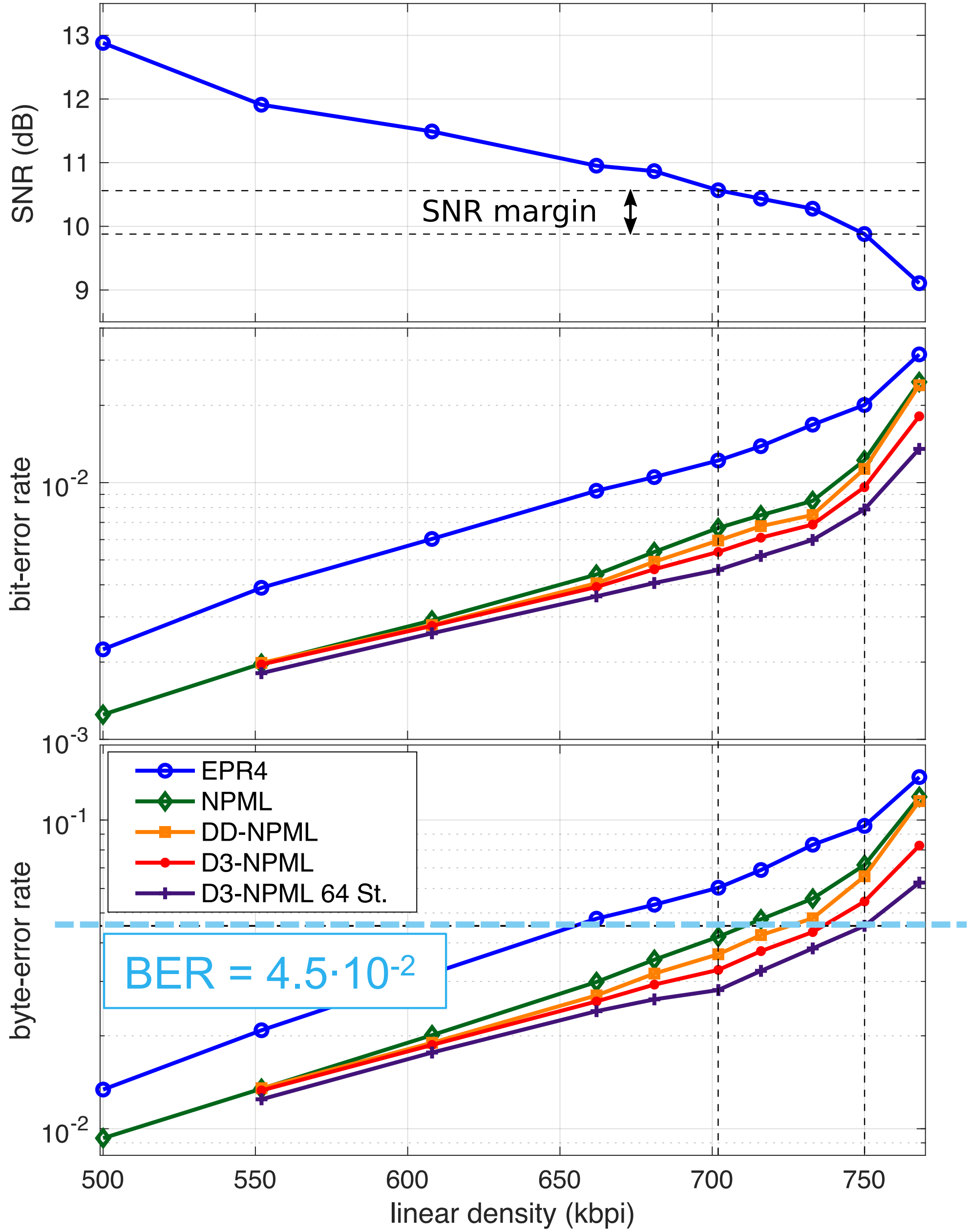


Recording Performance of SrFe

- High-moment tape writer
- 29nm-wide TMR reader (HDD)



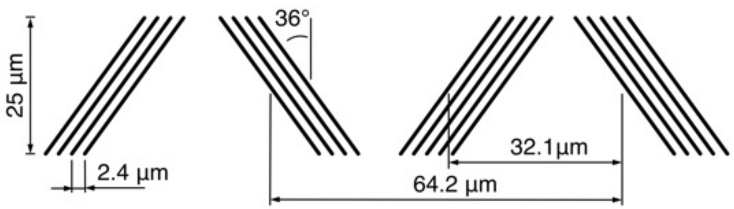
HDD slider



SrFe tape supports a **linear density of 702kbpi** with a 29nm reader and provides an operating margin of ~ 0.7dB SNR

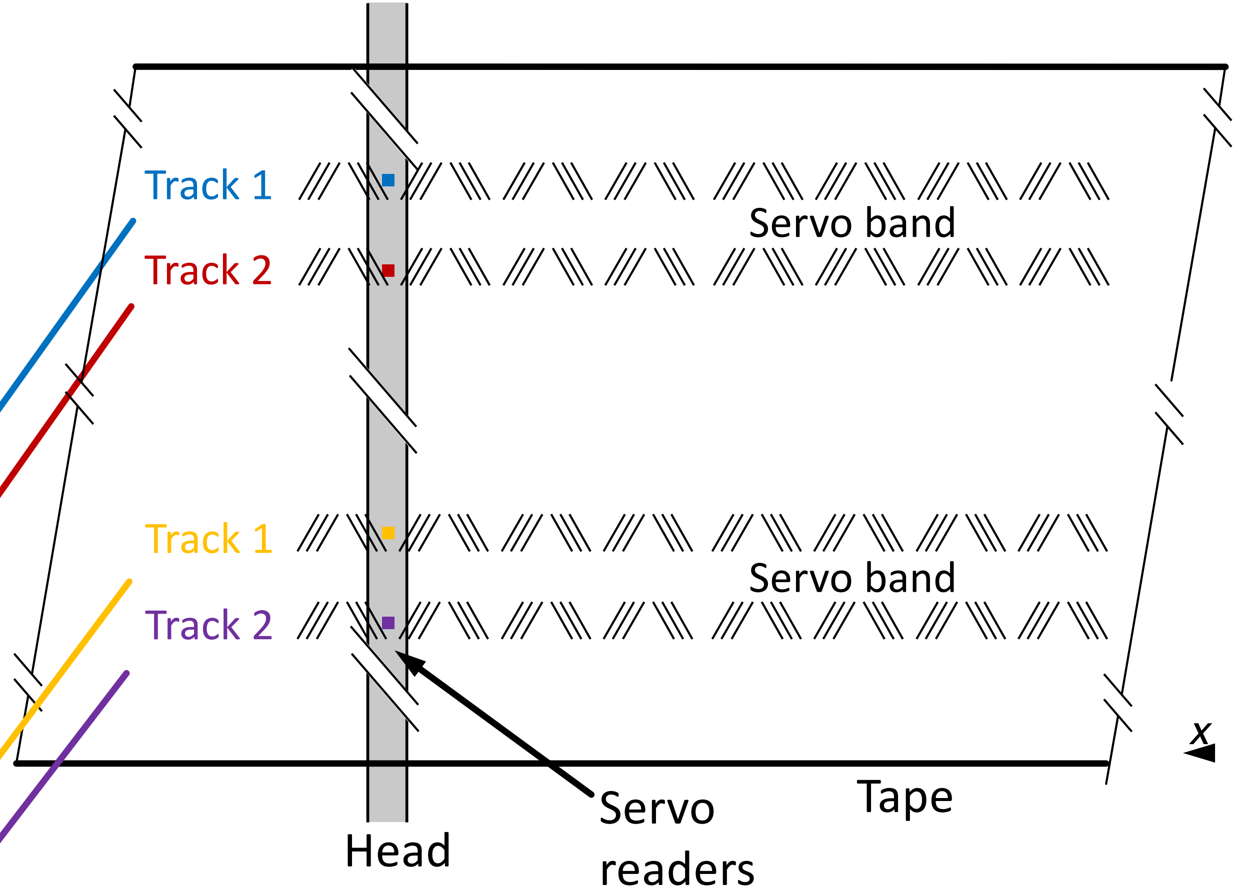
New Servo Pattern and Quad Channel Averaging

New 36° Servo Pattern

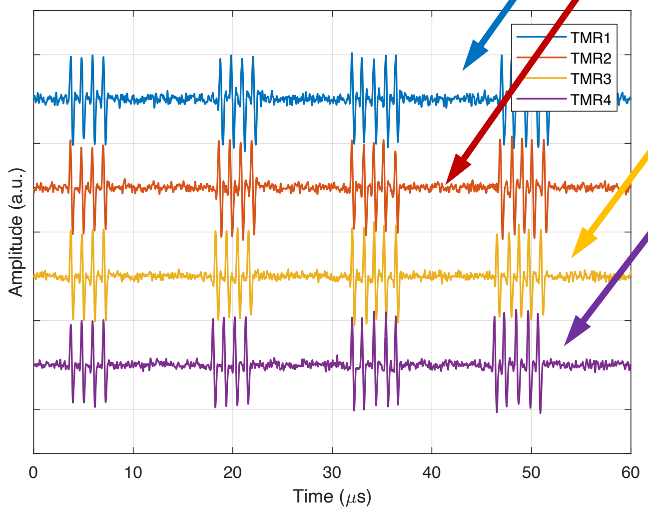


→ Increased resolution

→ Reduced delay



Quad Channel Averaging

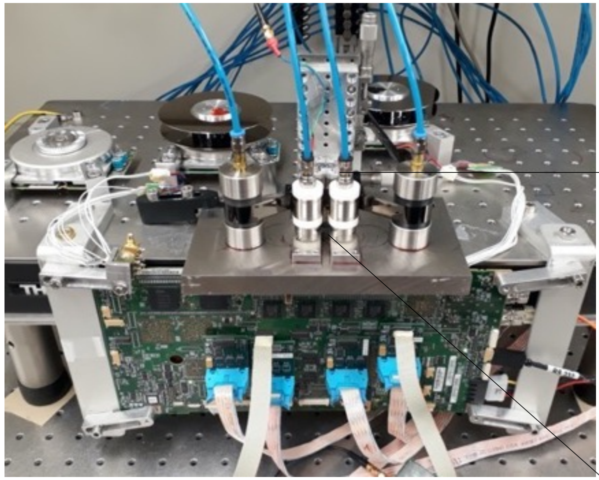


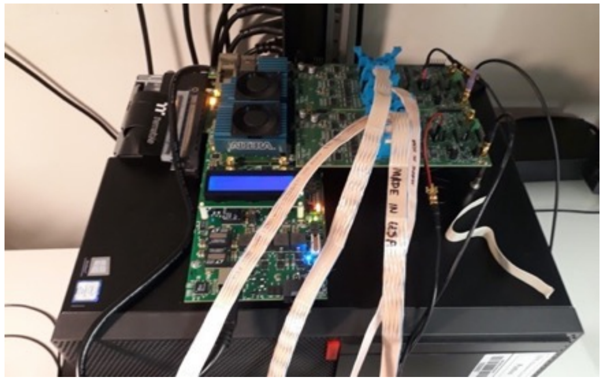
→ Reduced Noise

→ Enhanced Robustness

Tape Path, HW Platform, Actuator and Head

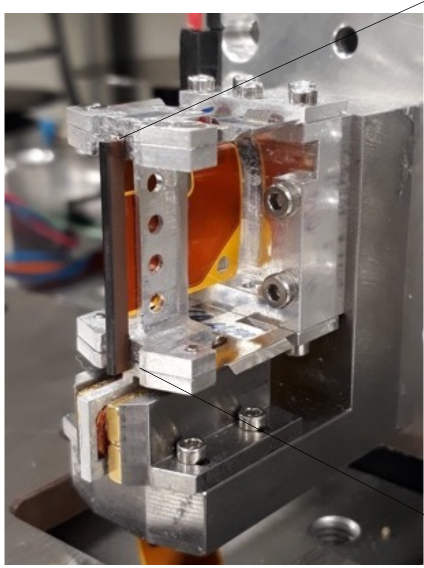
Low Noise Tape Path





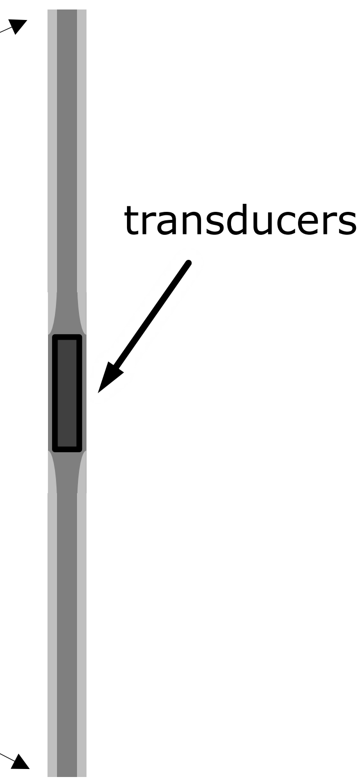
New FPGA prototyping board
→ Four servo channels
→ Reduced delay

Prototype Actuator



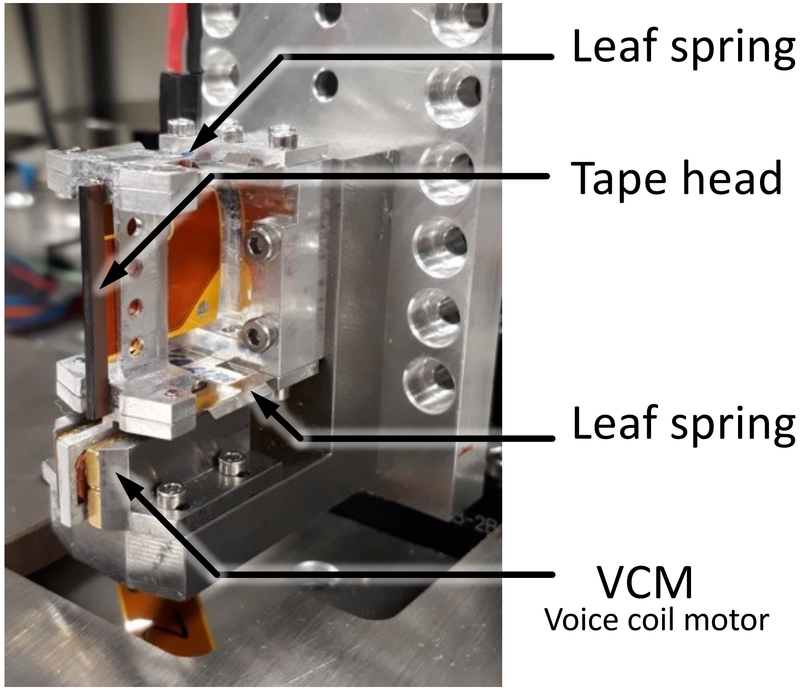
Enables:
Narrow rollers for low LTM
High bandwidth control

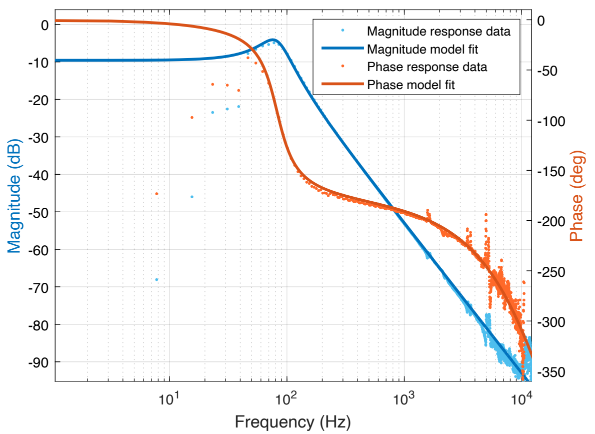
Low friction head



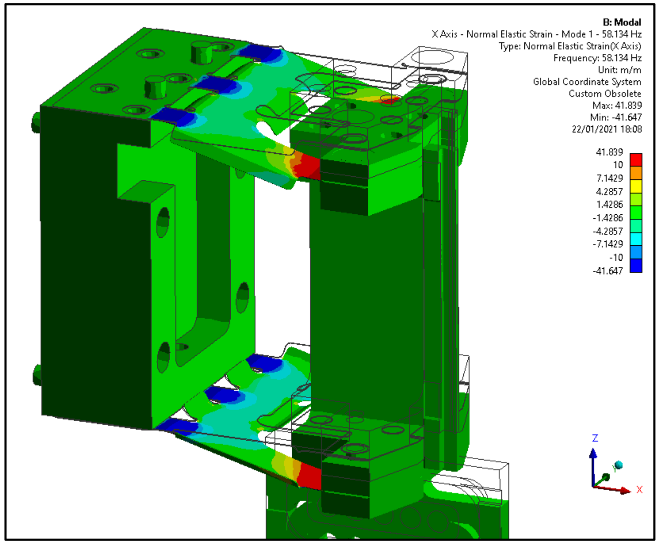
Enables:
Smooth media

Narrow Roller Actuator

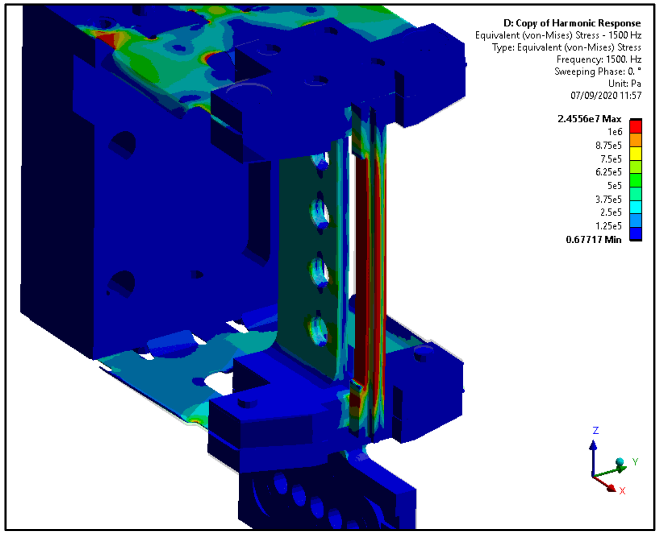




Desired motion ☺

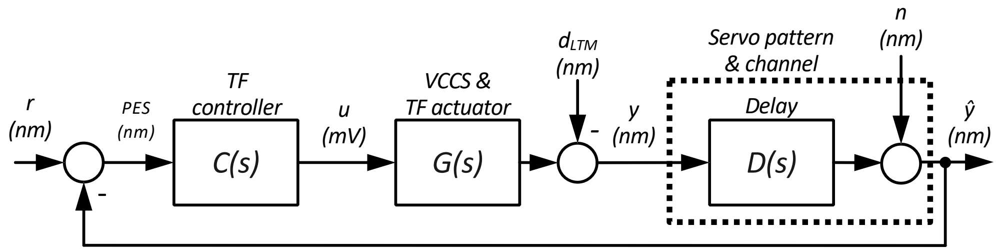


Undesired motion ☹

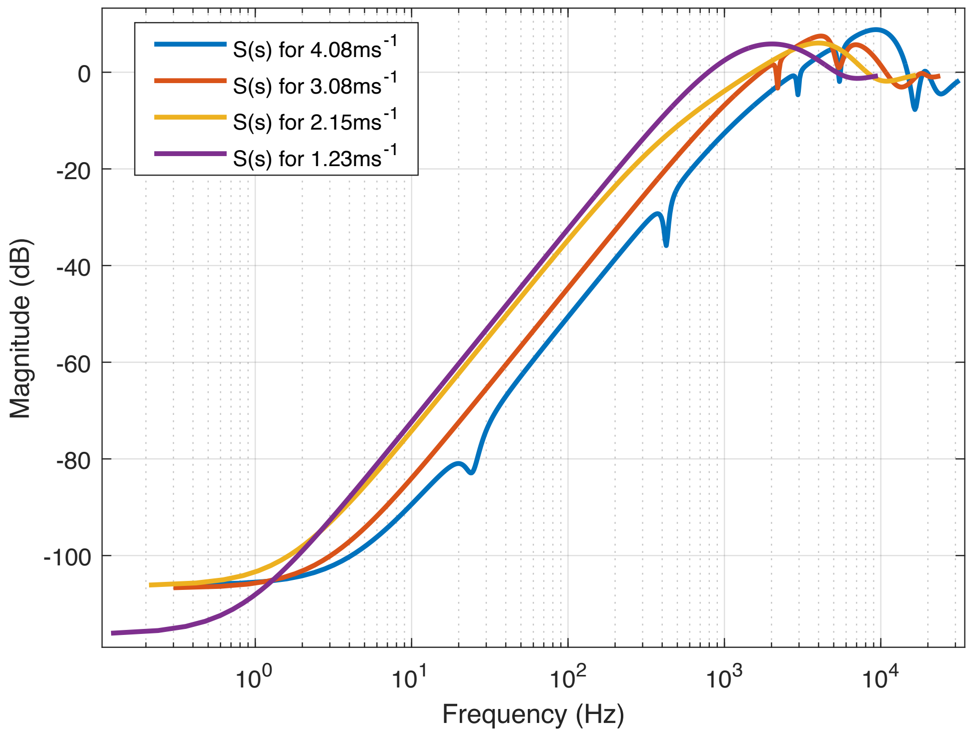


Speed Optimized Control

H_∞ model based control



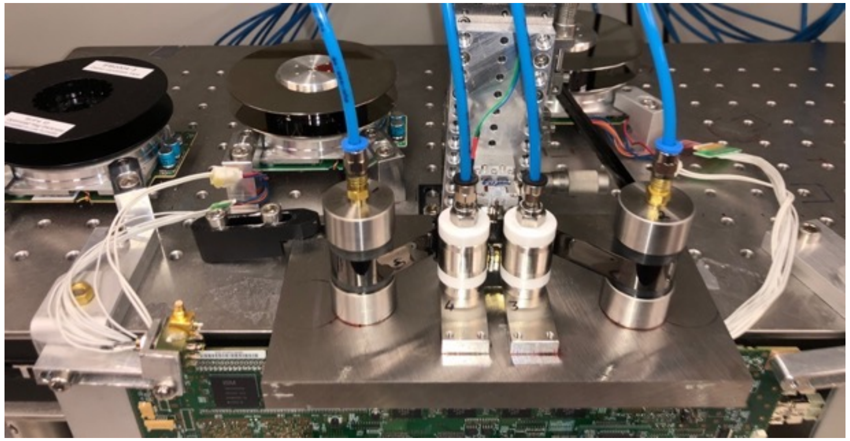
- Distrurbances shift to higher frequency with increasing tape speed
- Delay decreases with tape speed
- Optimize controller design for each tape speed
- Control bandwidth ranges from 700Hz to 3100Hz

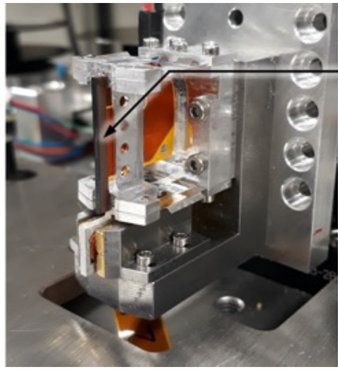


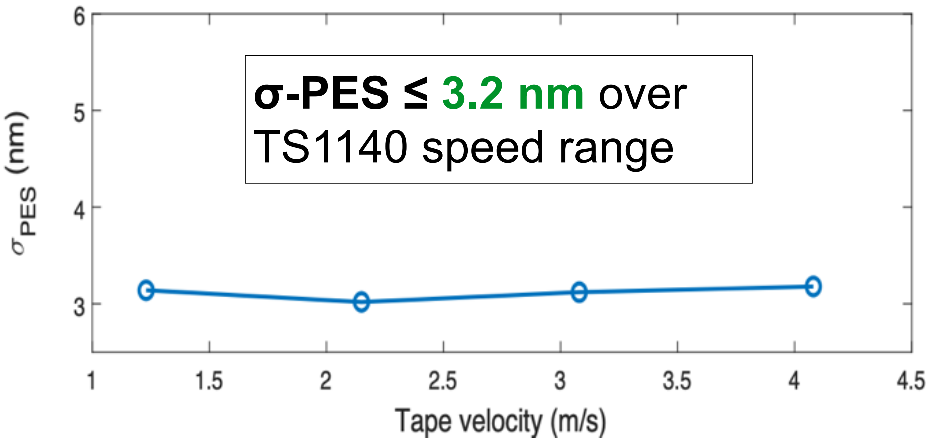
Track-following Performance on SrFe Tape

- New TBS servo pattern with 36deg azimuth angle
- New X4 FPGA platform with 4-channel averaging
- Low disturbance tape path with new prototype track-following actuator and new low-friction head
- H-inf based optimized track-folllowing controllers

Low disturbance tape path







$$Track\ Width = 2 \cdot \sqrt{2} \cdot 3 \cdot \sigma_{PES} + Reader\ Width$$ (INSIC method)

Reader Width = 29nm
σ-PES ≤ 3.2 nm

Track width = 56.2nm
Track density = 451.9 ktpi

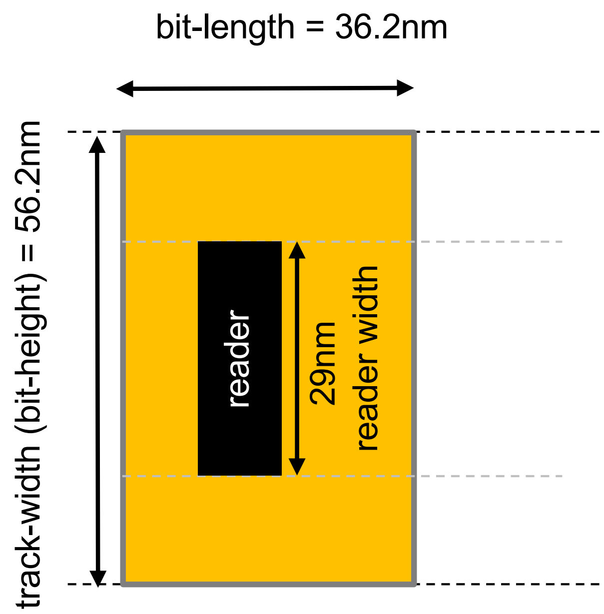
2020 Areal Density Demonstration on SrFe Tape – Main Results

Recording Result:

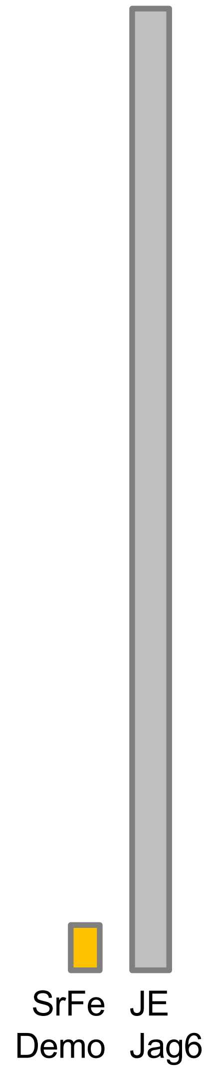
- Reader Width = 29 nm
- Linear density = 702 kbpi. Bit-length=36.2nm @ ~0.7dB SNR margin

Track-following Servo Results:

- Worst case $\sigma_{PES} \leq 3.2$ nm
- Track Width = $2 \cdot \sqrt{2} \cdot 3 \cdot \sigma_{PES}$ + Reader Width = 56.2nm
- Track density = 451.9 ktpi



Areal density = track density × linear density
= 451’994 b/in × 702’000 b/in
= 317’300’000’000 b/in^2
= 317 Gb/in^2



Bit size comparison

317Gb/in^2 Areal Density –> Potential Cartridge Capacity

29 JE cartridges (20TB)

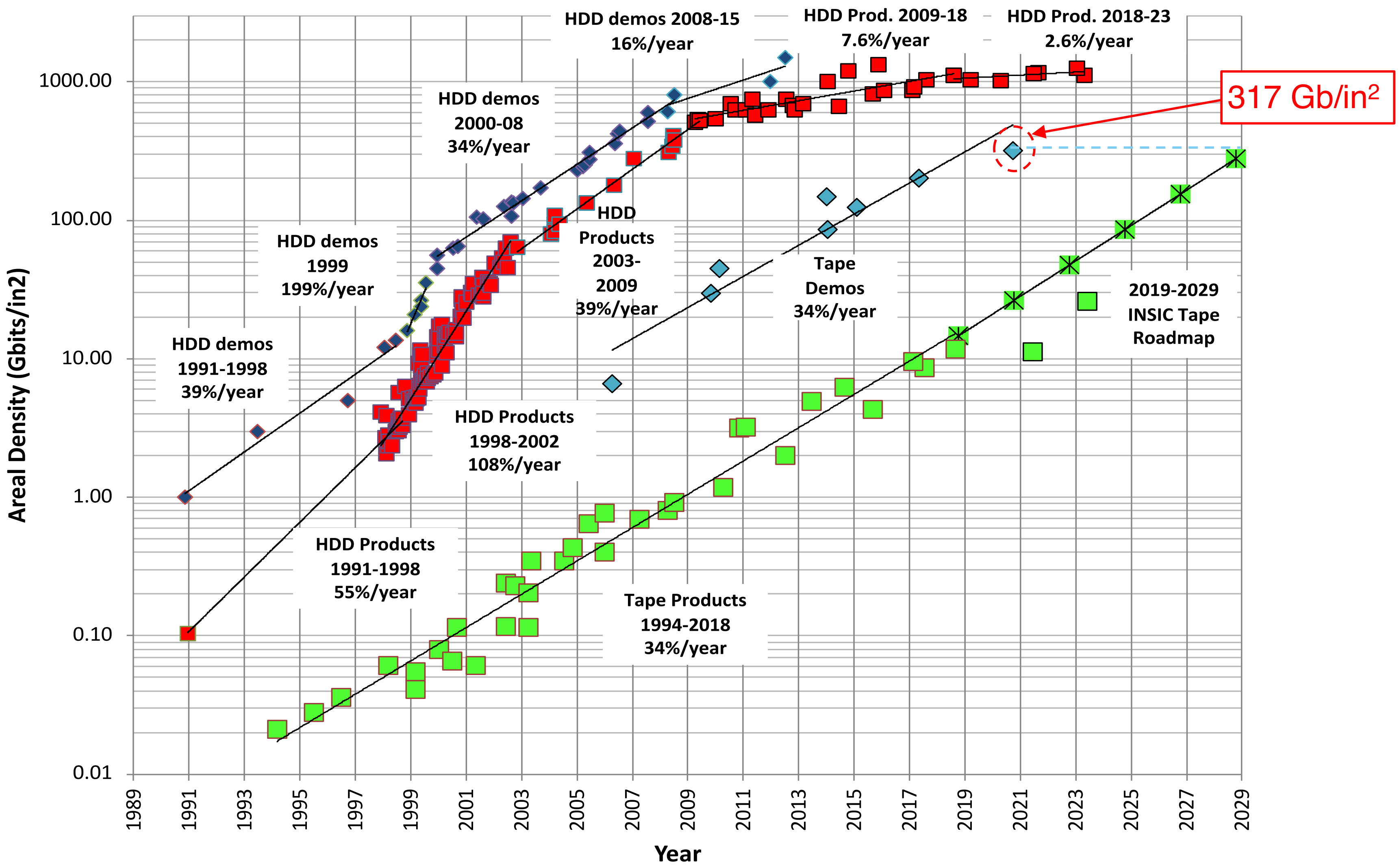
Tape	Thickness (um)	Length (m)	Length Scale Factor	Areal Density (Gb/in^2)	AD Scale Factor	Capacity (TB)
JE	4.6	1163	1	11.7	1	20
Demo	4.3	1244	1.07	317.3	27.12	**580.2**

Potential Cartridge Capacity:
580 TB !

IBM TS4500 Tape Library w/
potential 580TB cartridges: **Potential Capacity = 10.18EB**

Areal Density Scaling

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



Other Recent Tape Demos

Western Digital

IEEE TRANSACTIONS ON MAGNETICS, VOL. 58, NO. 4, APRIL 2022

Evaluation of Sputtered Tape Media With Hard Disk Drive Components

Pierre-Olivier Jubert[1], Yuri Obukhov[2], Cristian Papusoi[1], and Paul Dorsey[1]

[1]Western Digital Corporation, San Jose, CA 95119 USA
[2]Western Digital Corporation, Rochester, MN 55901 USA

"We demonstrated that such sputtered tape could achieve an areal density of **400 Gb/in^2** under shingled magnetic recording conditions."

SONY

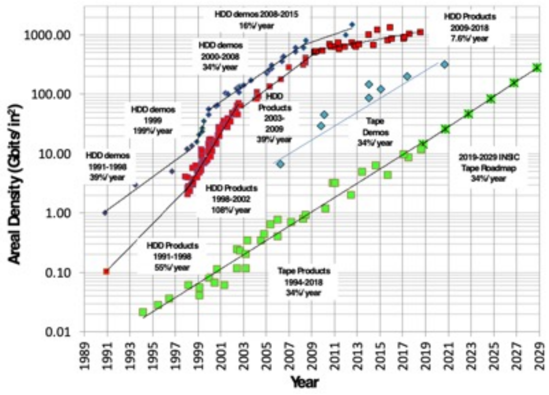
OPA-01 1

Challenge of Media Noise Suppression with Oxidation Control in Granular Structure for CoPtCr-based Sputtered Tape towards 400 Gb/in^2

Junichi Tachibana[1], Hiroyuki Kobayashi[1], Teruo Sai[1], Satoshi Kodama[1], Takashi Aizawa[1],
Atsushi Yamaguchi,[1] and Shin Saito[2]

[1]Sony Storage Media Solutions Corporation, Tagajyo, Miyagi 985-0842, Japan, Junichi.Tachibana@sony.com
[2]Department of Electronic Engineering, Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan

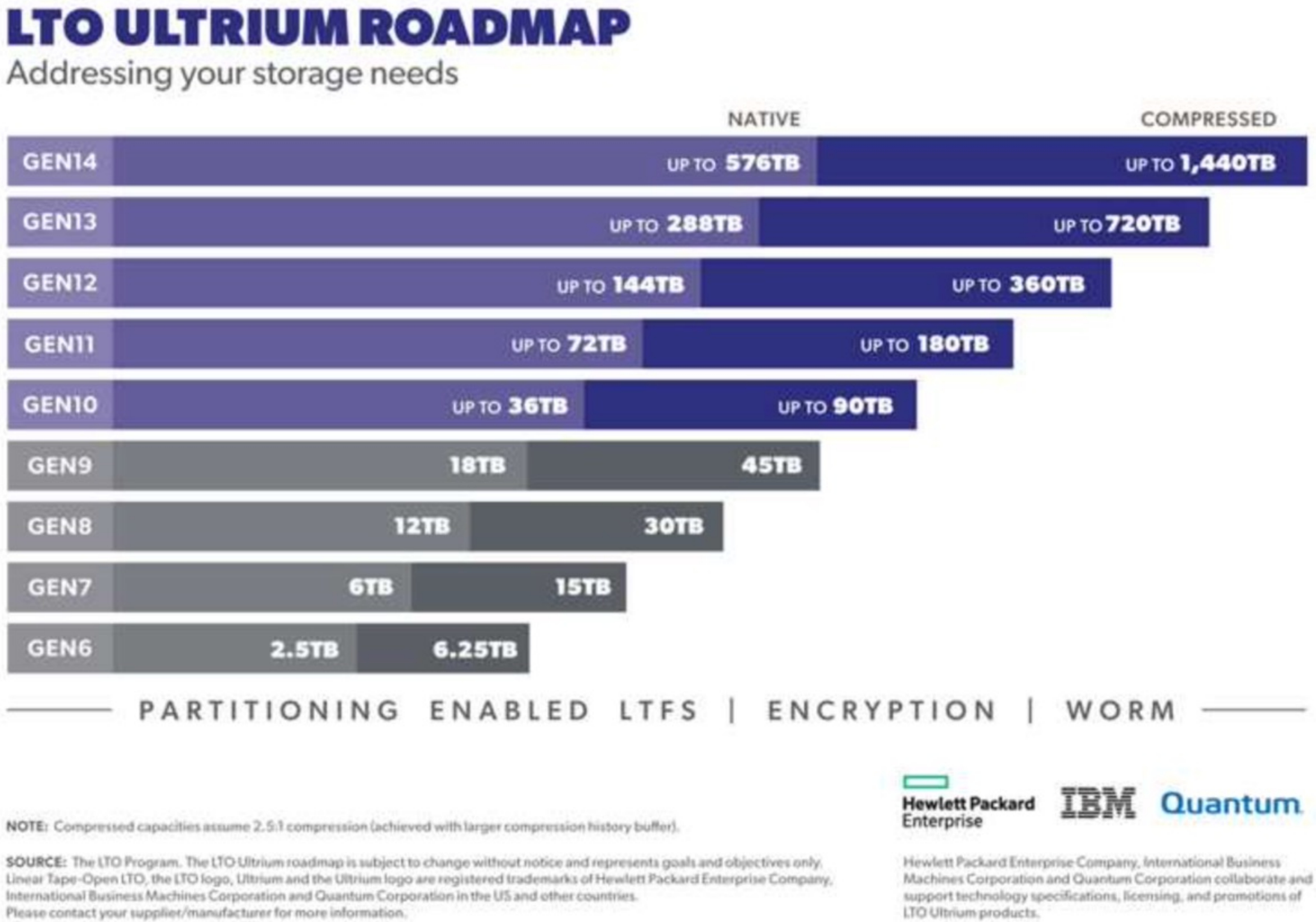
INSIC 2019-2029 Tape Roadmap



Parameter/Year	2019	2021	2023	2025	2027	2029	
1. Capacity (TB)	25	49	96	188	369	723	40.00%
2. Maximum data rate per channel (MB/sec)	14.8	19.6	17.3	22.9	30.3	30.0	
3. Maximum total streaming drive data rate (MB/sec)	475.0	628.2	830.8	1098.7	1453.0	1921.6	15.00%
4. Minimum streaming drive data rate	115.7	134.9	236.0	275.3	321.1	499.4	1.50
5. FC Speed Roadmap (MB/sec)	12800	25600	51200	51200	102400	204800	
6. Number of channels	32	32	48	48	48	64	
7. Tape thickness (um)	4.88	4.50	4.14	3.82	3.52	3.24	-4.00%
8. Data capacity reserve	3.0%	3.0%	3.0%	3.0%	3.0%	3.0%	
9. Tape length that is recordable (meters)*	1069	1163	1266	1377	1497	1628	40.00
10. Tape length total (meters)**	1109	1203	1306	1417	1537	1668	4.17%
11. Track density (TPI)	23,918	36,933	57,041	88,119	136,154	210,413	24.29%
track pitch = 2.54×10^7/tpi (nm)	1,062	688	445	288	187	121	
12. Linear bit density (Kfci)***	612	714	833	971	1133	1321	8.00%
fcmm = kfci/0.0254	24,094	28,104	32,780	38,235	44,597	52,018	
13. Areal density (Gbits/inch2)	14.64	26.36	47.49	85.58	154.23	278.01	34.23%
14. Tape speed (m/sec)	6.2	7.0	5.3	6.0	6.8	5.8	-0.65%
15. Tape width in mm	12.65	12.65	12.65	12.65	12.65	12.65	
16. ECC and formatting overhead	20.00%	20.00%	20.00%	20.00%	20.00%	20.00%	0.00%
17. Servo track and layout overhead ****	16.00%	16.00%	16.00%	16.00%	16.00%	16.00%	0.00%
18. Number of passes to write a tape	313	483	497	768	1187	1375	
19. Number of passes to end-of-life (media)	32200	34560	37093	39812	42730	45862	3.6%
20. Time to fill a tape in mins	877	1300	1927	2855	4232	6272	21.74%
21. Number of data tracks	10,006	15,451	23,863	36,864	56,960	88,026	24.29%
22. Number of data bands	4	4	8	8	8	8	
overall head span (um)	3,000	3,000	1,500	1,500	1,500	1,500	
23. Tape Dimensional Stability (ppm)	142	92	119	77	50	32	-13.77%
24. Bit Aspect Ratio (BAR)	30	23	17	13	10	7	-13.11%
25. Bit Error Rate	3.50E-20	1.79E-20	9.11E-21	4.65E-21	2.37E-21	1.21E-21	-28.57%

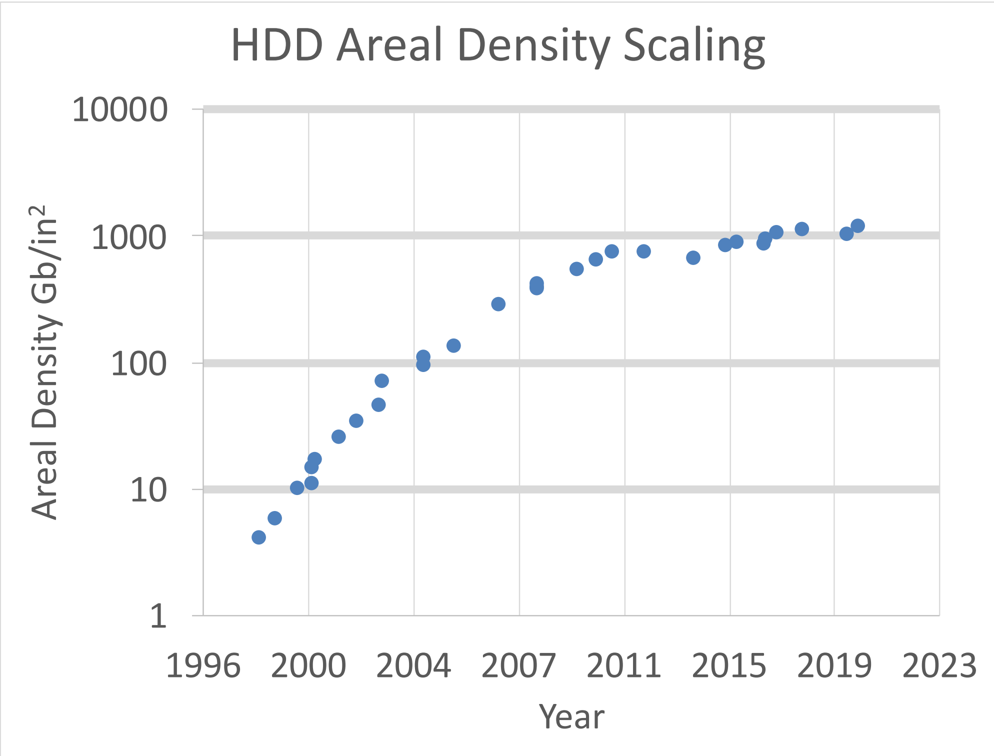
INSIC Roadmap available at: https://www.insic.org

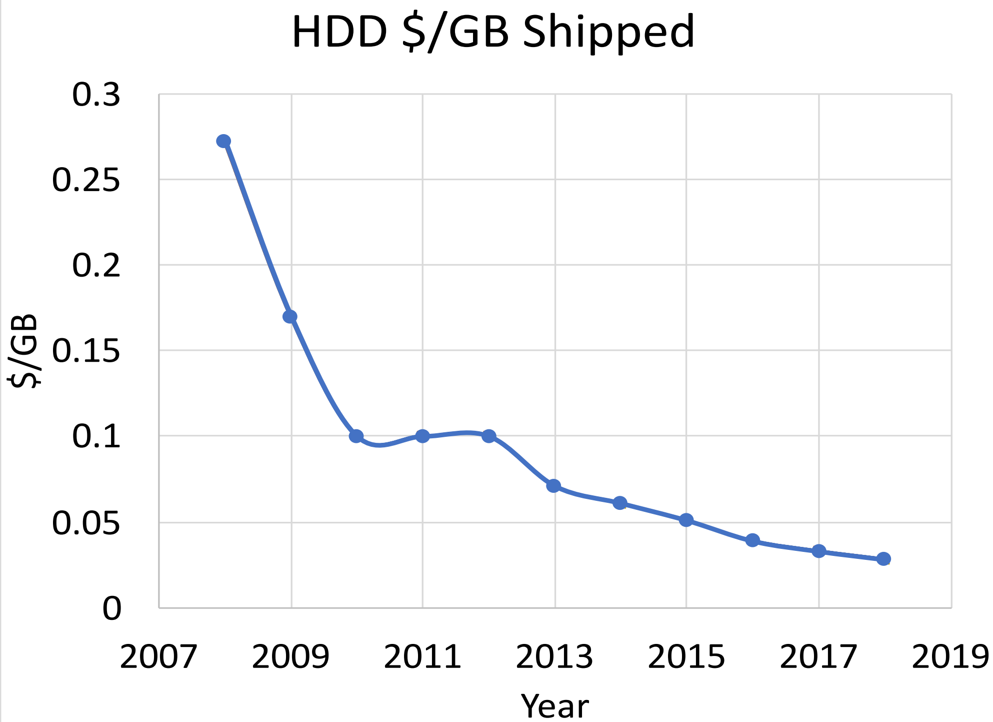
LTO Consortium Roadmap



The Future of HDD

Recent HDD scaling





Recent Capacity Scaling of HDD: Volumetric Density

- Slow down in areal density scaling partially compensated by adding more disks: conventional technology ran out of space at about 5 platters
- Helium filled drive → less turbulence → thinner/wider disks → higher capacity
  - WD 18 TB Drive 9 platters
  - WD 20 TB SMR (YE2020)
  - 20TB PMR (2021)
  - 22TB PMR/26TB SMR (2022)
- Helium: lower power, better TCO
- Working on ultra-thin aluminium-magnesium platters for 11 platter drive

Doesn't scale:

→Effectively no space for more heads and platters in current form factor

→Cost of head and platters dominate

→ IOPS / TB is decreasing

Heat Assisted Magnetic Recording (HAMR)

→Laser used to locally heat the media to lower the magnetic field required to write a bit

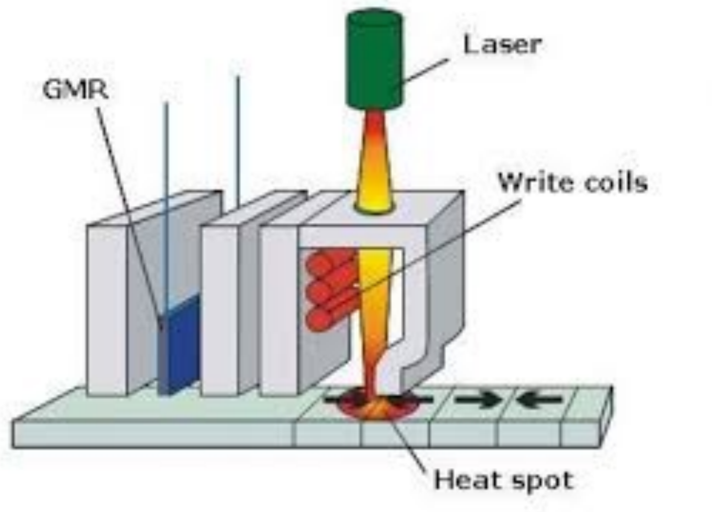
Many Engineering Challenges

- New media, glass disc, thermal stability of overcoat/lubricant, disc/media reliability
- Confinement of heat
- Data dependent track width, transition curvature
- **Life time of laser / near field transducer**
- **Cost**

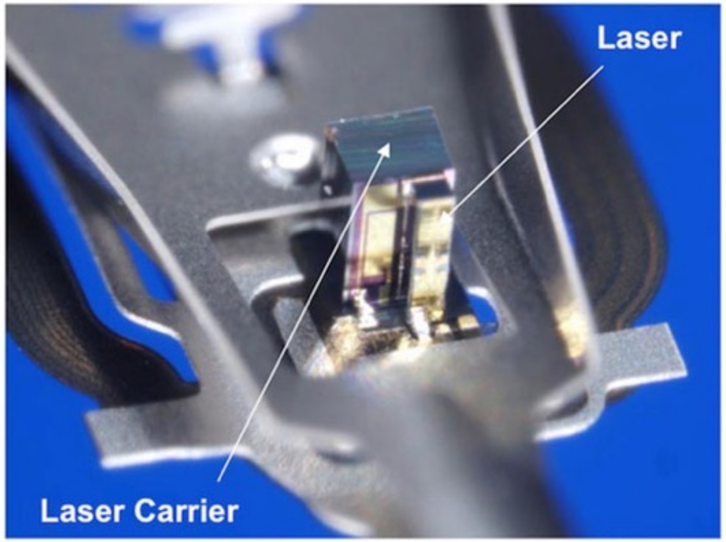
Time to market:

- **WD: probably will never be cost competitive**
- **Seagate 20TB TB HAMR drives, Dec. 19 2020:** "shipping on a limited basis…."
- Seagate shipped first 30TB HAMR drive samples 04/23, now shipping in volume, but also shipping 22TB PMR and 24TB SMR drives
- Need to get to ~40TB before it will be economical

→ **Scaling:** expect to scale to 50TB by 2028 ~13.5% CAGR



HAMR Head

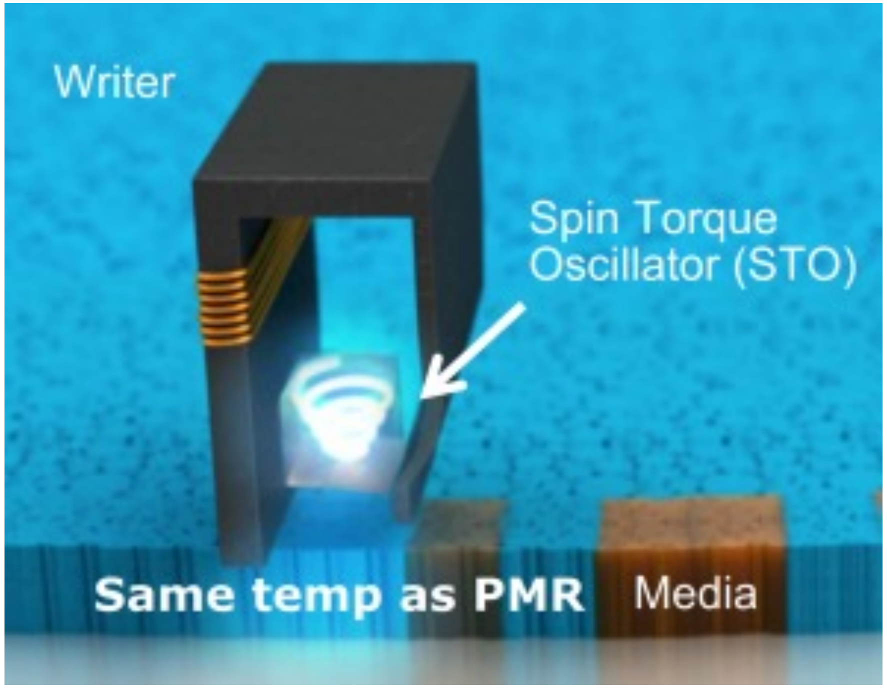


* https://www.techradar.com/news/seagate-confirms-20tb-hamr-hard-disk-drives-have-been-shipped

Microwave Assisted Magnetic Recording (MAMR): WD

→Spin torque oscillator used to locally generate magnetic fields that rotate at microwave frequencies and lowers the magnetic field required from the writer to write a bit

- Minimal additional cost
- Less disruptive than HAMR (more PMR-like)
- Does not use heat → better reliability than HAMR
- Still many challenges to scale:



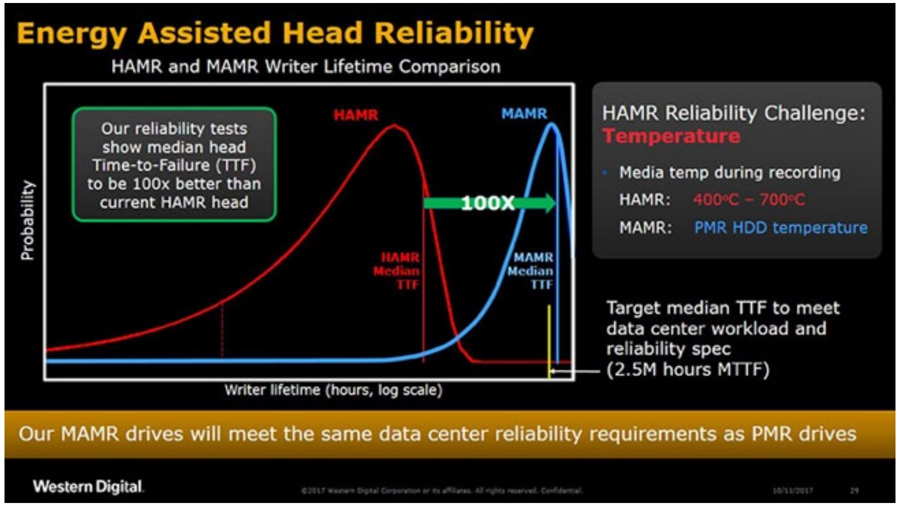
Time to market:

WD: Was supposed to be introduced in 2019,….. But nothing to date

WD now plans to use ePMR and OptiNAND technology combined with 10 platters to scale to 30TB and no longer mentions MAMR

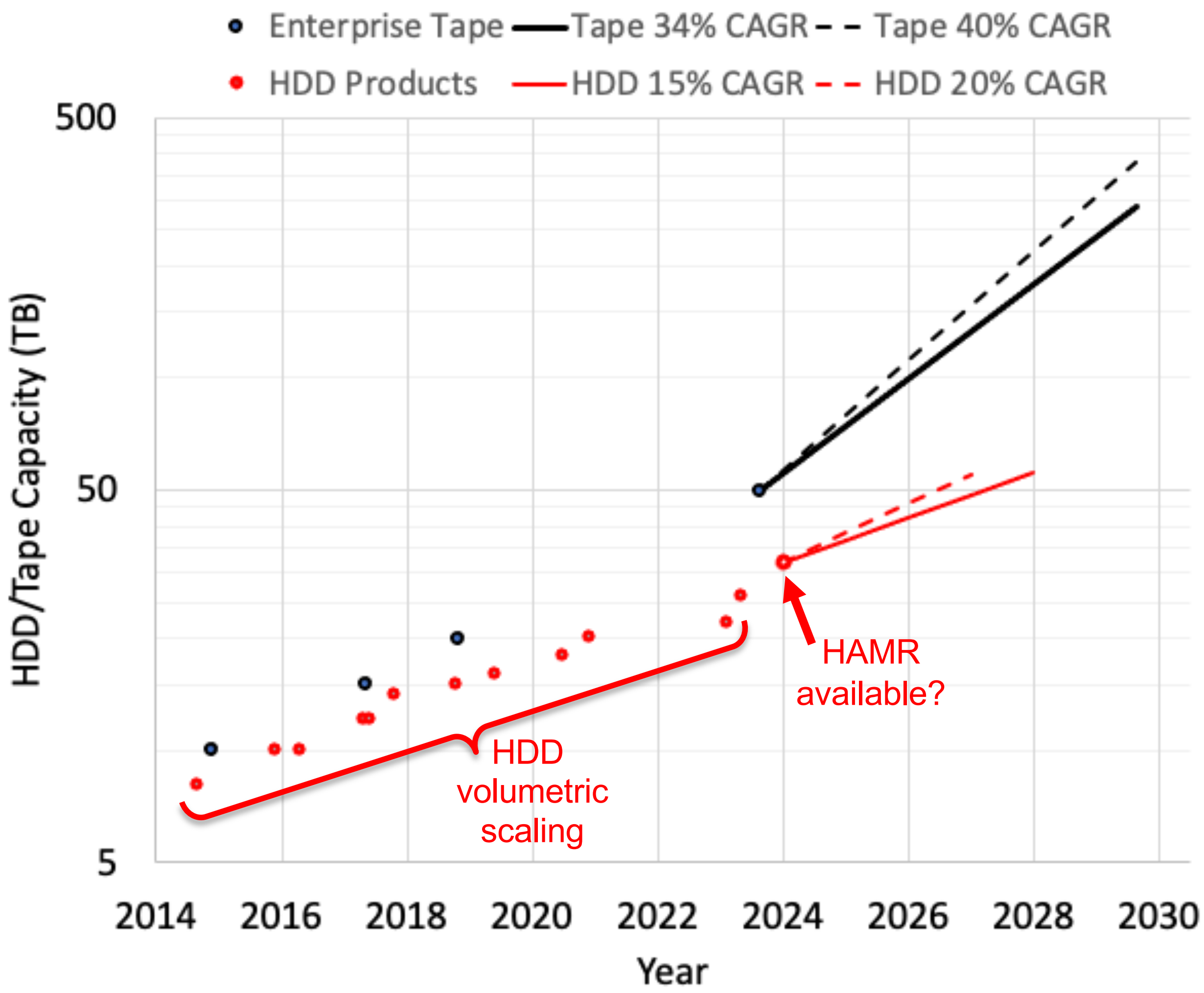
Toshiba: started shipping FC-MAMR in 2021, but only 18TB with 9 platters….

plans 30TB 11 platter MAS-MAMR in 2024



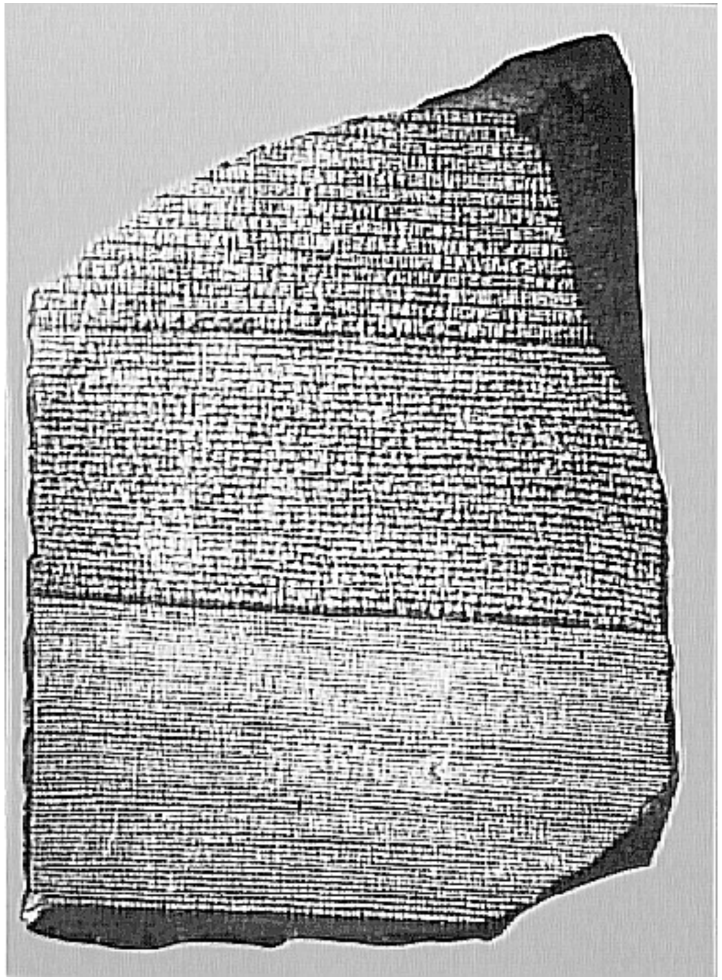
https://hexus.net/tech/news/storage/123953-seagates-hdd-roadmap-teases-100tb-drives-2025/
https://www.anandtech.com/show/17097/western-digital-30tb-hdd-plans-mamr-future-unclear

Tape and HDD Projections



The cost advantage of tape over HDD will grow exponentially for the foreseeable future!

Alternative Archival Storage Technology



DNA Storage

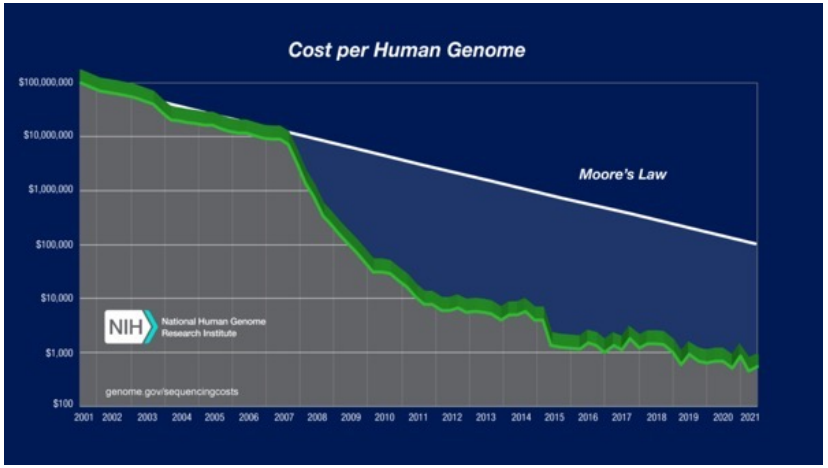
Store data in the base sequence of DNA, 4 nucleotides (A,C,T,G) → ~2 bits/base pair
Write and read using DNA synthesis and sequencing

Pros:

- high storage density: estimated hundreds of PB/gram (but only MBs - GBs actually written)
- Long DNA lifetime in controlled environment
- Technology to read DNA likely to be around for a long time

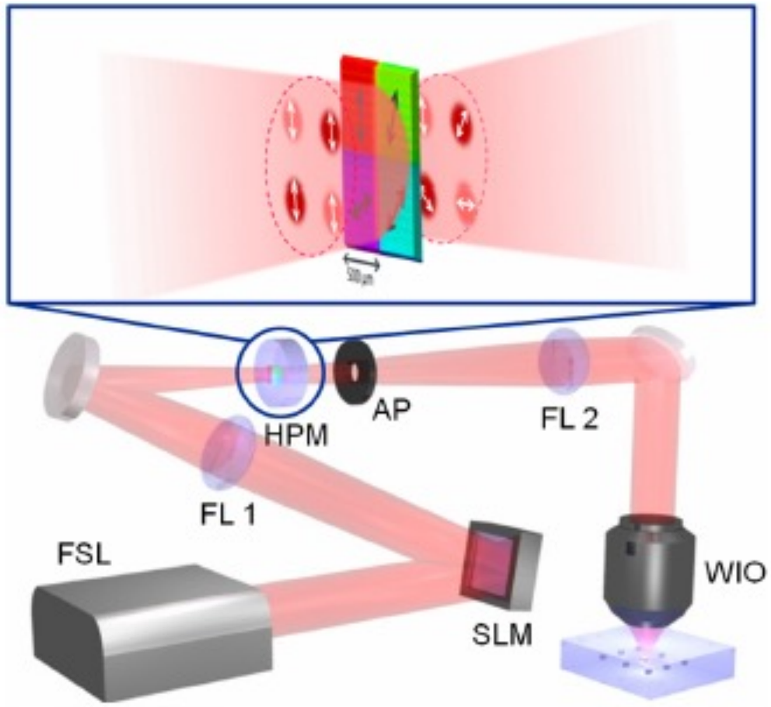
Cons:

- Cost: read ~$2000 / GB, (~10 million times more expensive than tape), cost to write >> cost to read
- Cost to synthesize (write) DNA historical scaling ~13.5%/ year, i.e. lower than scaling rate of tape → need new paradigm e.g. Catalog
- Large market for read (sequence) technology, small market for write (synthesis) technology
- Data rate: estimate ~kbytes/second (require multiple reads)
- File access at high volumetric density is a challenge, recent progress in this area



"5D Optical Storage" a.k.a. "Superman memory crystals" or Glass

Ultra fast laser pulses encode data in 3 spatial dimensions + polarity & intensity of light in a quartz



Published Demonstrations:

2013: ~47 Mb/in^2 x 3 bits/pixel x 3 layers

2018: 10 layers

2019: Microsoft stores 1978 Superman movie

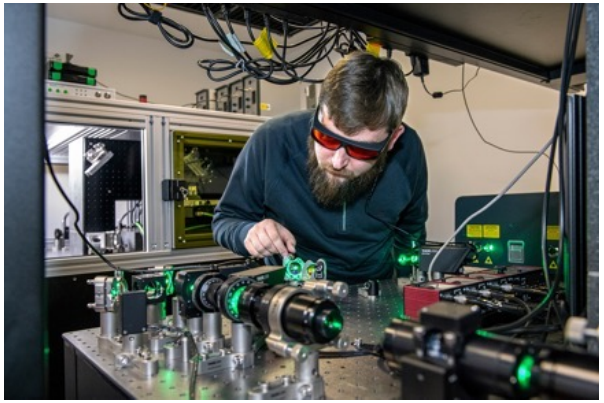
2021: 8kB/s write speed (best published to date)

2023: Microsoft ~6TB per 12cm x 12 cm sheet

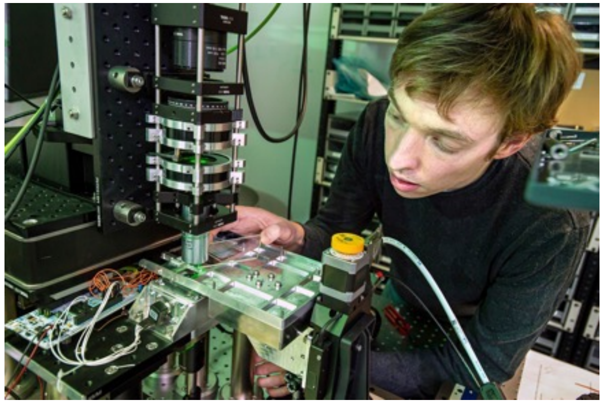
Pros:
- media is quartz (low cost)
- long media/data lifetime

Cons:
- very low data rate (8 kB/s)
- Write: femto second laser → big and ~$100k
- Read: expensive microscope + offline image processing
- No servo: how to control position for write and read

write



read



Summary:

- Tape has a sustainable roadmap for at least another decade with areal density projected to scale with a 34% CAGR and capacity with a 40% CAGR

- HDD areal density growth has stagnated with current scaling ~ 2.6% CAGR
  – MAMR projected to enable density scaling ~15% CAGR for a few generations
  – HAMR projected to enable density scaling ~15-20%, but smaller $/GB scaling

- The continued exponential growth of data combined with the stagnation of $/GB scaling of HDD is impacting the storage hierarchy
  – Tape, HDD and Flash will continue to coexist
  – Increasing use of Flash for IOPS intensive workloads
  – Increasing use of Tape for archive and active archive workloads