



A SNIA Technology Affiliate

The Looming Need for Molecular Storage

Aaron Ogus
Distinguished Engineer
Microsoft Azure Storage

- August 4, 2022



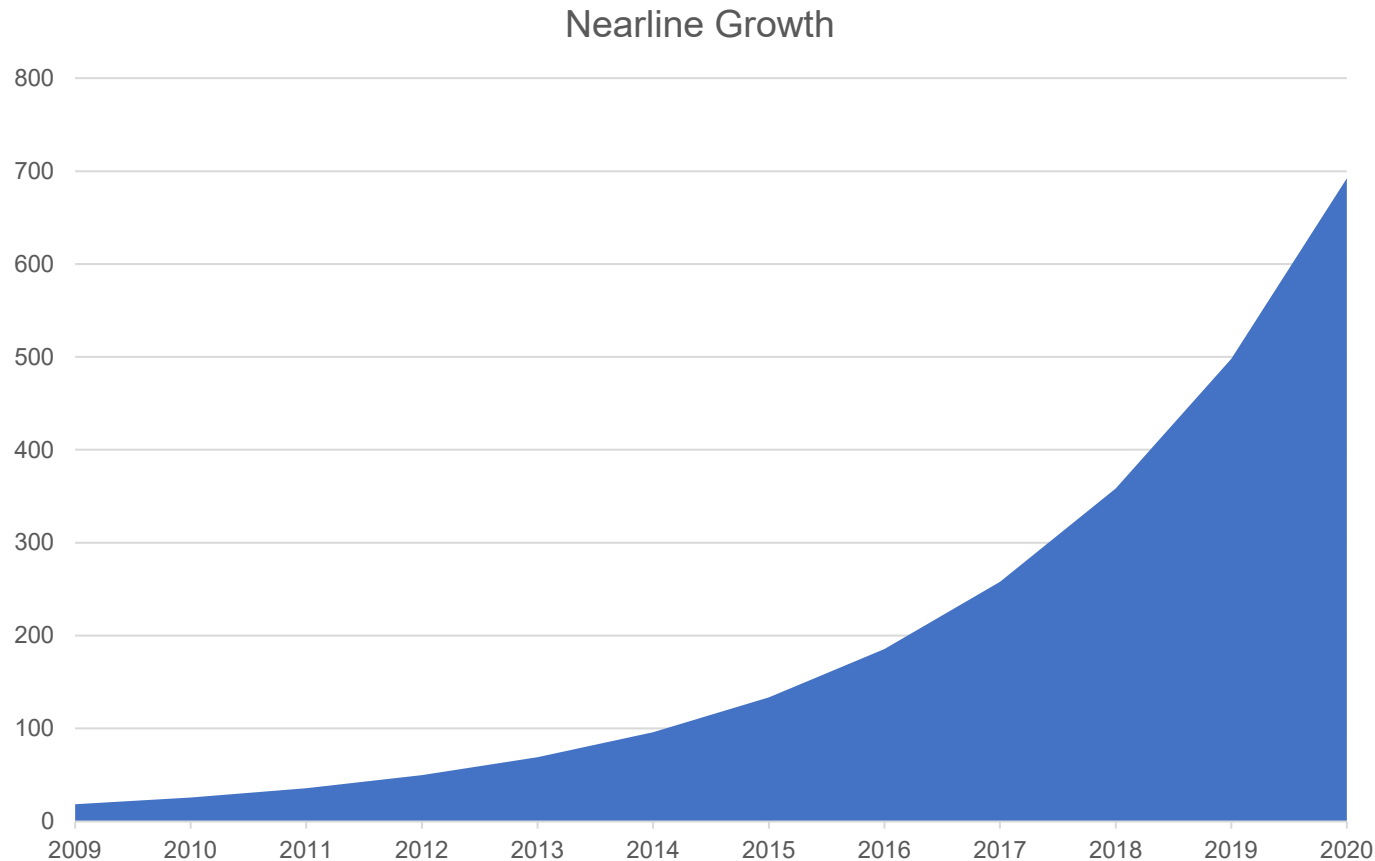
Flash Memory Summit



How do we store EBs of Data?

- Literally Miles and Miles of Storage Racks
- Hundreds of MW

Data Headed to the Cloud



In 2010's HDD shifted from Consumer to DC. By 2030 almost all HDD will be nearline in the cloud.

Underlying data growth signal is unwavering, and projects to over **7 ZB** per year by 2030 <not shown>

Data CAGR is 40%
HDD Capacity CAGR < 20%

Efficiency Improvement 2008->2022



2006: 500GB HDD, 3 Replica, 4 HDD/Server

2008: 1 TB HDD, 3 Replica, 12 HDD/Server

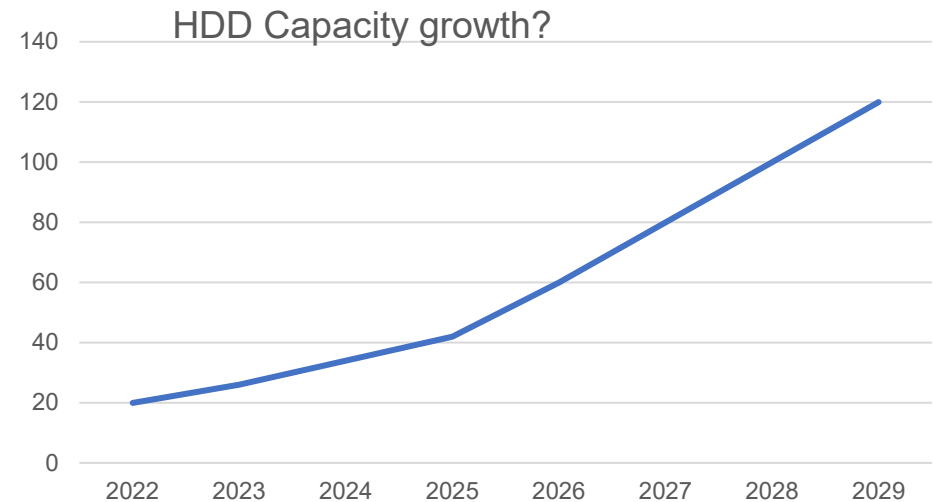


2021: 20 TB HDD, <1.3 Replica, 80+ HDD/Server

> 99.5% reduction in online storage
cost over 15 years

HDD Roadmap

- HDD Suppliers are shifting to MAMR and HAMR
- Only HAMR has legs to 60 TB+
- HAMR might run out of steam around 100TB
- Optimistically if we assume HDD growth to 230TB+, we will require 5% of current US Power capacity to be dedicated to spinning HDDs by 2042, and 60% by 2050.

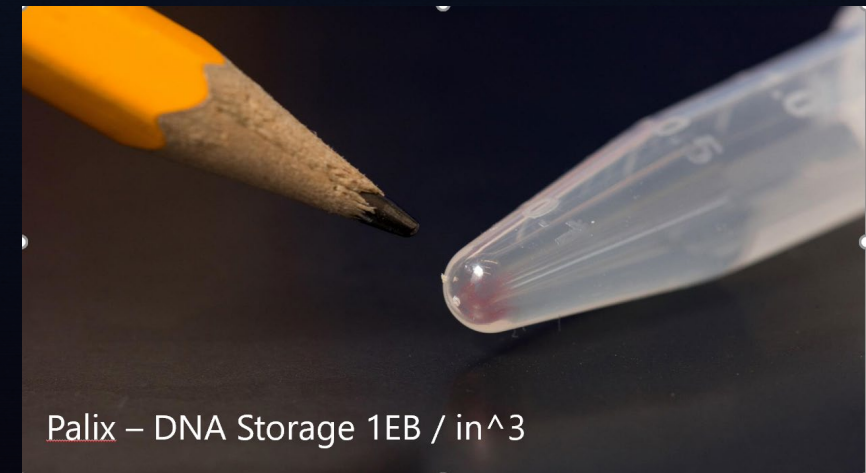
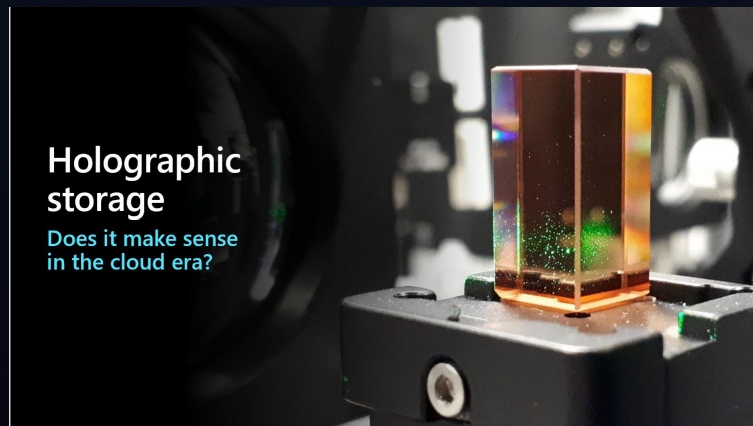


Something has to change:

Data growth has to slow,
Generating capacity needs to vastly increase
Storage technology needs to change

Storage Research – MSR

Azure Storage and MSR collaborate on research and commercialization of new media types for storage.



Molecular Storage is the Panacea for density

Highest Storage Density of Any Technology

Problem: Read and Write Rates

Electronic Molecular interface... based on medical research

Can we build fast Molecular Storage?

Where is most of Humanities Data Stored?

HDD shipped 1ZB for the first time in 2021

1,000,000,000,000,000,000,000 bytes 10^{22}

500 MW/ZB

Power if on 20TB HDDs @10W each would be: 50Million HDDs and 500MW.

What about Human Brains?

8 Billion Humans

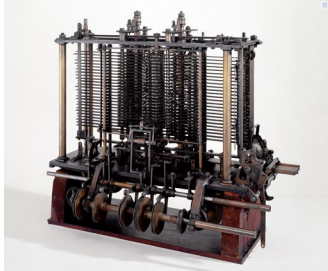
2.5PB per brain (estimated)

8 MW/ZB

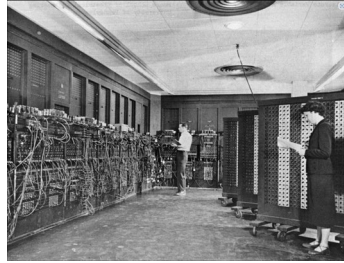
Power @ 20W per brain: 160GW

2 Yottabytes in Brains (About 1000x data stored on HDD)

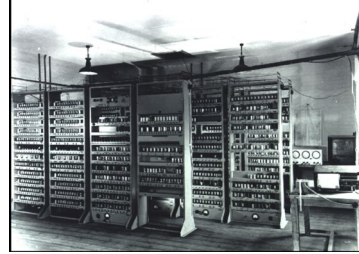
Compute Evolution



Analytic Engine
Designed 1820
Metal Gears
Mechanical



ENIAC
1945
Vacuum Tubes
Crystal Diodes

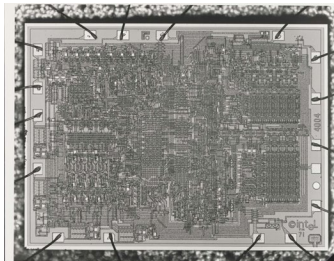


EDSAC
1949
Thermionic Valves

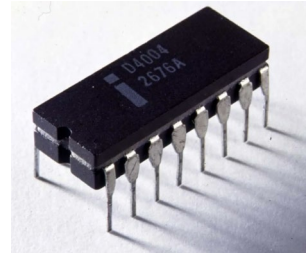


1955 Burroughs
Atlas
Transistors

1957
Univac
Transistors



Intel 4004: First computer on an IC
1971



Intel Xeon SPR
2021, 100,000,000x
improvement over 4004

What is the next technology platform shift for storage?

Magnetic Charge (Tape/HDD/Core Memory)

Electric Charge (Flash, RAM)

Optical

Holographic (Lithium Niobate and Lasers)

Molecular (DNA)

What kind of Molecular?:

DNA?

Molecular Machines?

What's Next?

Research in Molecular Simulation is Ramping:

Alpha Fold

MSRC

Resolving the Molecular / Electronic interface is critical

Possibility to leverage patterns from Biology? Molecular Machines?

Q&A