Library of Congress

Washington DC

April 15-16, 2024

Fred Moore, President
Horison Information Strategies
Horison.com

HORISON
Information Strategies

AI, ML, and Big Data Fuel Shift to Secondary Storage

Key Trends - 2025 and Beyond

Primary storage (highly active, online, critical) indefinitely stores highly active, response time critical data, approximately 20% of all data. Note: *typically, low latency devices such as SSD and HDD.*

Secondary storage (persistent data) indefinitely stores lower activity data on highly economical, secure, storage mediums that don't need to be accessed as frequently as data in primary storage**.** Note: *typically tape and nearline HDD.*

Secondary Storage is the Largest and Fastest Growing Storage Category ~ 80% of all Digital Data

- **~8.4 ZB Stored on SSD, HDD and Tape by 2025 (cagr. 25-30%)**
- **~80% of Installed Capacity (6.72 ZB) is Lower Activity, Archival, Cold - Secondary**
- **Over 80% of Data Created is Unstructured**
- **AI, ML, Big Data, Edge/IoT, Virtual and, Gaming and Robotics Filling the Archives**
- **The Archival Copy is Usually the Only Copy of Data**
- **Retention Periods Over 100 Years are Common**
- **The Active Archive Becomes a Crucial Tier**
- 
  Servers and HDDs Consume >80% of IT Energy
- **AI GPUs consume 4x as much power as servers used for cloud apps**

Source: Horison Information Strategies, LLC

New Apps, Social Media and AI Creating Data at Warp Speed…

AI, Social Media and New Apps Fill the Archives

Time to Reach 100 Million Users	
Threads	5 days.. Instagram App
ChatGPT	2 months (AI)
Instagram	2 yrs, 4 months
WhatsApp	3 yrs, 4 months
Snapchat	3 yrs, 9 months
YouTube	4 yrs, 1 month
Facebook	4 yrs, 7 months
Twitter	5 yrs, 2 months
Pinterest	5 yrs, 6 months
LinkedIn	7 yrs, 11 months

Social media is a collection of unregulated media platforms.

Every "tweet", "accept", "swipe left or right", "post", "friend", "tag", or "like" creates some amount of stored data.

New Apps Reach Archival Status Quickly.

How Much Data?

How Much Storage?

Where is it Stored?

How Much Value?

How Much Data?

How Much Data?

Source: UBS, Yahoo Finance, Horison Information Strategies, LLC

VMF (Vertical Market Failure) Considerations

Risk Factors to the Current Secondary Storage Model

- **The Zettabyte scale secondary storage market (cold, archive) has become the exclusive domain of few suppliers.**
- **IBM is the only (1) tape drive developer/supplier controlling the entire tape ecosystem specifications.**
- **Fujifilm and Sony are the only (2) LTO tape media suppliers.**
- **HPE, IBM, Quantum and Spectra are the primary large-scale tape and library suppliers.**
- **Seagate, Toshiba and WD are the only (3) remaining HDD suppliers.**
- **(Tape - WD), (HDD - Seagate and TDK) are the only R/W head manufacturers.**
- **Will current HDD and tape development roadmaps keep pace with demand?**
- **HSDCs leverage their bargaining and buying power to drive down prices impacting vendor margins, R&D investments.**
- **In the event of a secondary storage VMF, sustainability challenges will become insurmountable for HDDs to address.**
- **As supplier profit margins become insufficient, future R&D funding, roadmaps, will place innovation at significant risk.**

A VMF occurs when the underlying infrastructure fails to meet the demands of the market it serves.

- ***Can Current Storage Technologies Overcome Their Challenges? Can AI help?***
- ***Will a New Novel Technology or Solution Arrive?***

(A new tape format, Ceramics, Photonics, DNA, Glass, halide, or the Space Belt)?

Source: Horison Information Strategies, LLC

The Data Storage Industry Gets Ready for AI

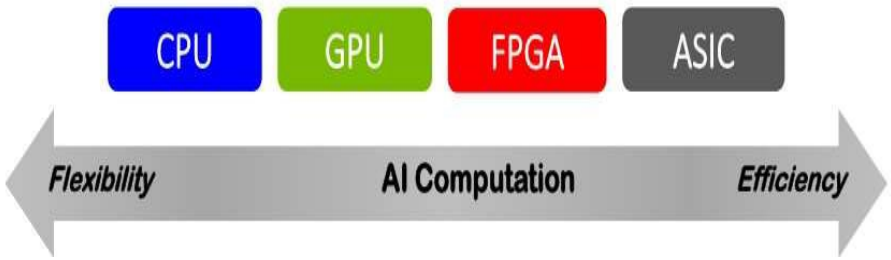
AI Enables Machines to Learn From Raw Data, Gain Experience, and Perform Human-like Tasks

Raw Data

21	11	38	73	02
91	48	54	27	88
33	76	68	47	35
66	94	12	59	37

Purpose Built Computing

Programmable Acceleration Options



- Vast sets of raw data are gathered (mined) from existing data to build AI applications and data sets; most data is stored as unstructured.
- Raw data is transformed into a format that algorithms process to train the AI model requiring massive volumes of data and storage capacity.
- The final AI model encapsulates all the knowledge gained during training.
- Businesses require easily accessible and searchable data, anticipating the need to be re-processed, re-trained, or monetized in new ways - ***hence the active archive arises.***
- Purpose built computing becomes critical for compute intensive AI.
- AI is Compute Intensive Consuming *Massive* Amounts of Energy and Water (Approximately. = Ireland's annual consumption**!**).

Source: Horison LLC.

What Can AI Do for Storage Management?

Initial Challenges for AI to Address

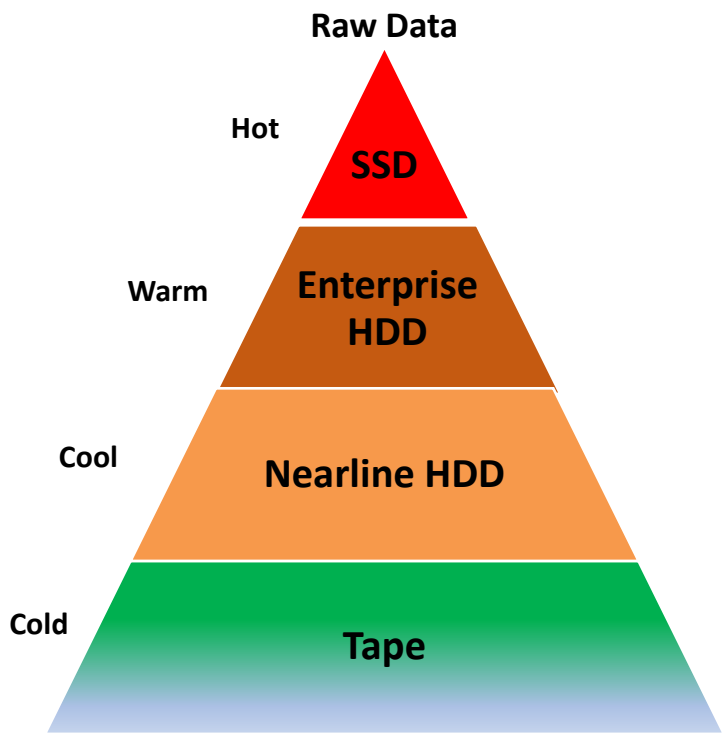
- Management of data across its lifecycle as it moves from one location to another.
- Optimizing capacity, performance, carbon footprint and cost across tiers.
- Improve storage HDD space utilization (typically less than 50%).
- Determine optimal format should data be stored (block, file, object)?
- Determine data retention, EOL and disposal policies.
- Predict storage and performance needs, proactive backup.
- Determine the best data protection method for stored data.
- Add, reduce or reconfigure storage resources.
- Learn to manage the Space Belt – 9,494 active satellites collecting data.
- Anticipate, intercept and provide proactive defense against cyber threats.
- Since AI for storage management is a new technology, AI algorithms will improve over time as more information is gathered.

Source: Horison Information Strategies, LLC

*Will AI Solve This Enormous Storage Management Problem?

"Most Data is Typically Stored in the Wrong Place"

Today's Storage Hierarchy



ACTUAL Installed Capacity by Device 2025		
SSD	998 EB	11.8%
eHDD	106 EB	1.2%
Nearline HDD	6,704 EB	80%
Tape	588 EB	7.0%
Total Installed Capacity 8,400 EBs or 8.4 ZBs		

How Can We Get From *Actual* to *Optimal?*

OPTIMAL Installed Capacity by Device 2025	
SSD	15%
eHDD	1%
Nearline HDD	24%
Tape	60%
Total 100%	

- 80% of all Data is Archival, or Cold and Should be Stored in Secondary Storage (i.e., Tape)
- However Over 75% of Archive Data is Stored on HDDs
- *Moving 60% of the worlds (low activity, inactive) data from HDDs to tape can avoid 79 million tons of C02e

Source: Horison, LLC

A New Storage Stack Evolves to Avoid a Carbon Crisis

Mid-Century Installed Capacity Projections From 2030 to 2050

Zettabyte Era 2016 $1x10^{21}$

Yottabyte Era 2043 eta.. $1x10^{24}$

Cagr. ~27.5%

← By 2035 AI Assumes Most Human Storage Management Tasks →

Note: Installed Capacity Projections Built From TrendFocus Growth Trends

2030

Installed Capacity ~26.29 ZB

- NVM (Flash) 9.5%
- HDD 85% Nearline, Online
- Tape 5.5%
- Green Sustainable Long-Term Storage ?

2040

Installed Capacity ~251.83 ZB

- NVM (Flash) 15%
- HDD 48% Primarily for Backup, Active Archive
- Tape Legacy <5%
- Green Sustainable Long-Term Storage ? 32% *TBD*

2050

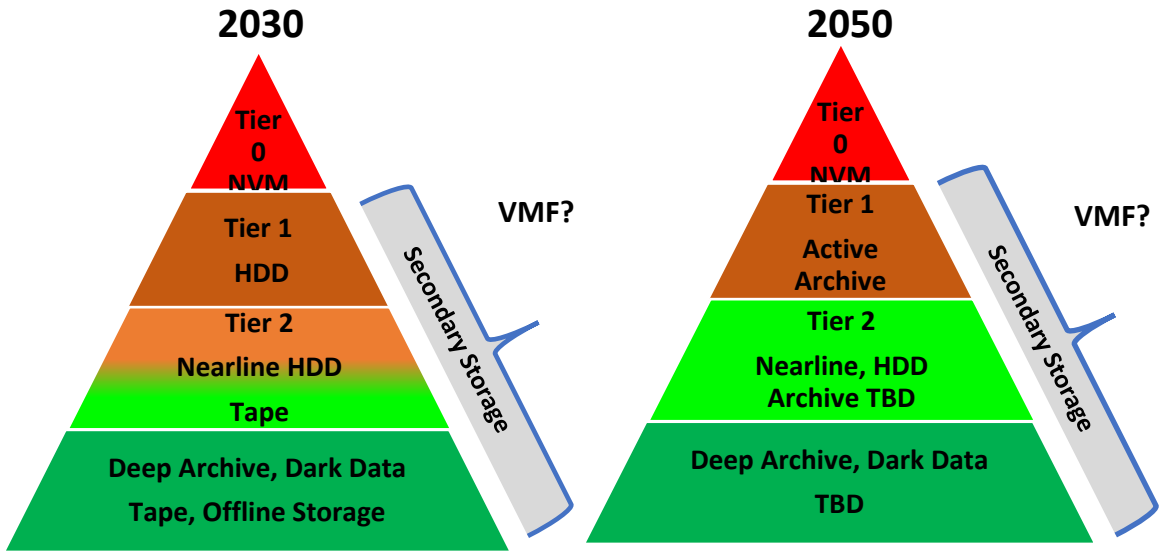
Installed Capacity ~2.41 YB

- NVM (Flash, tbd) 18%
- HDD 24% Primarily for Backup, Active Archive
- Tape Legacy <3%
- Green Sustainable Long-Term Storage ? 55% *TBD*

Long Range Scenario

~80% of All Data Stored Will be Active Archive, Archive or Deep Archive, as *Green Technology* Becomes Critical Through 2050.

- AI, ML, Big Data Analytics, Edge/IoT Computing, Virtual Reality, Augmented Reality, Gaming and Robotics Fill and Activate the Archives (Secondary Storage).
- Sustainability Challenges Become Untenable (VMF).
- HDD and Tape Endure But New Immutable, Random Access, Removable (Air Gapped), Sustainable, Secondary Storage Solution Will Be Required To Avoid a VMF.

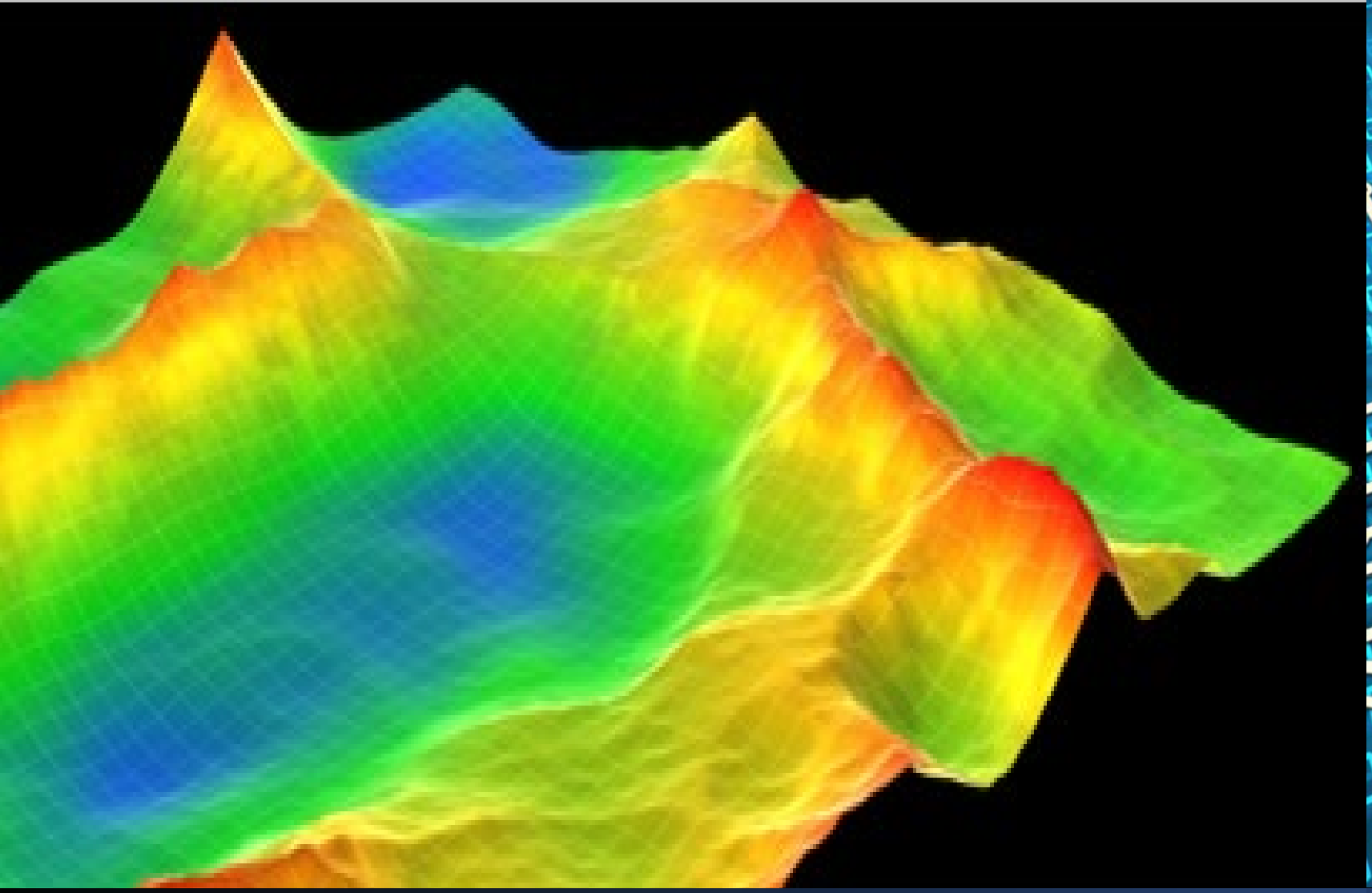


Advanced Sustainable Energy Efficient Secondary Storage

Existing Energy Efficient Storage

TBD – Tape, Cerabyte, Folio, Glass..?

Source: Horison, LLC



Without AI, Effective Storage Management Will Be a *Hope* – Not a *Strategy*

Fred Moore