Cerabyte Market Positioning

Secondary Storage Demanding a New Improved Solution

BY: Fred Moore Horison, LLC

Storage Industry Landscape

As the widespread impact of Covid-19 eases, new concerns have emerged including a fragile supply chain, prolonged inflation, higher interest rates, rising transportation, food, health care and energy prices which have painted a choppy global macro-economic picture. These trends have also caused data storage demand and forecasts to fluctuate. COVID initially increased demand for desktop, laptop, and mobile HDDs driven from a home-based workforce, but both soon fell off the cliff as the continued inexorable move to SSDs replacing HDDs accelerated. Reduced global business and manufacturing activity during the pandemic ultimately caused a decrease in demand for HDDs and cloud storage services creating the massive surplus in hard drive supply we see today. It's undoubtedly temporary, but the HDD surplus is still hurting global HDD sales. HDDs currently store about 85% of the world's digital data with about 60% of secondary storage data stored on HDDs. Current data growth projections of ~25% annually suggest a doubling of data stored every three years.

The Zettabyte Era demands advanced solutions to effectively contain petascale and exascale storage requirements. In parallel the unquenchable thirst for IT services has placed energy usage, carbon emissions and the overall environmental impact of data centers squarely in the bull's eye for most corporations to address. Data centers currently consume an estimated 3% of the world's electrical output. Magnetic tape is currently the most economical and environmentally friendly storage solution available, but even with the new capabilities and low cost of tape, the majority (~60%) of the world's lower activity, cold or archival data lies dormant for years on more costly to operate HDDs which require constant power and cooling to maintain reliable operations. Without a new storage solution to effectively manage and contain Zettabyte era demands, HDDs and magnetic tape will continue to compete indefinitely for the enormous secondary storage opportunity with neither solution being ideal. A sustainable and cost-effective long-term storage solution remains an unsolved problem.

What's Fueling the Rapid Growth of Secondary Storage?

Secondary storage is persistent storage designed to indefinitely store less active data on more economical, secure storage mediums that don't need to be accessed as frequently as data on primary storage. TrendFocus projects installed storage capacity across all enterprise data centers is expected to reach ~8.4 ZBs by 2025 with at least 80% (~6.72 ZB) classified as archival or "cold" data making secondary storage by far the largest data storage market. By 2030, total installed capacity is projected to reach ~26.3 ZBs suggesting ~21.04 ZBs of total secondary storage requirements. These projections

TRENDFOCUS
Enterprise Storage Installed Capacity - Total Exabytes

Installed Storage Capacity	2025	2030
eSSD Total EB	847	1,956
Captive eSSD Total EB	151	565
Hyperscale Nearline Total EB	4,584	18,181
OEM/Other Nearline HDD Total EB	2,120	4,016
Performance Enterprise HDD Total EB	106	62
Tape Total EB	588	1,517
Total Ent EB Installed	8,396	26,296

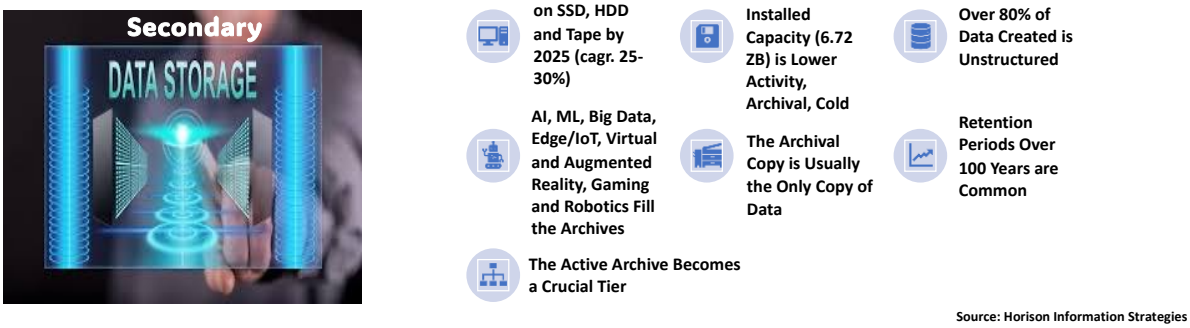
indicate that most of the world's cold data will continue to reside on costly, energy consuming HDDs unless enterprises begin to take advantage of a new secondary storage solution. The era of AI, along with ML, Big Data, Edge/IoT, Virtual Reality, Augmented Reality, Gaming and Robotics are both filling and activating the archives and unlocking the value of previously untapped data. As a result, a fast-emerging requirement for a new Active Archive Tier has arrived.

Big Data, AI, ML Fuel Shift to Secondary Storage

Key Archive Trends - 2025 and Beyond

Secondary (persistent storage) is designed to indefinitely store less active data on highly economical, secure, storage mediums that don't need to be accessed as frequently as data in primary storage.

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



TrendFocus indicates magnetic tape capacity shipped growing at 19% annually, but still not effectively capturing the optimal share of the secondary storage demand. Moving inactive data from HDDs to tape, deleting unused, unidentified, and unwanted data still hasn't generated widespread momentum from storage administrators as all types of low activity data piles up on costly, high capacity Nearline HDDs.

A Vertical Market Failure is Unfolding for Secondary Storage

A VMF (Vertical Market Failure) in Secondary Storage exists when the underlying storage technology infrastructure becomes insufficient to address secondary storage market demands. The slowing rate of HDD and tape technology roadmap development in recent years, product delays, coupled with HDD and tape storage supplier consolidations have become alarming trends. This suggests that a new secondary storage solution not currently available will be required to address future secondary storage demands. Avoiding a VMF is critical to the survival of any ecosystem. Without a new vastly improved solution, the threat of a VMF for the vast secondary storage market becomes a likelihood, painting an unhealthy future for most of the world's data.

VMF Risk Factors – To the Current Secondary Storage Model

- IBM is the only tape drive developer/supplier controlling the entire tape ecosystem specifications
- Fujifilm and Sony are the only two tape media suppliers
- IBM, HPE, Quantum, and Spectra are the only large-scale tape library suppliers

- Seagate, Toshiba and WD are the only three HDD suppliers
- Western Digital is the Only Magnetic R/W Head Supplier
- The ZB+ WW secondary storage market has become the exclusive domain of few suppliers
- The few remaining storage suppliers limit competition enabling customers to "be held hostage"
- IP litigation risks, as evidenced by recent lawsuits, can impact and delay roadmap progress
- The LTO roadmap product specifications have been relaxed and delivery dates pushed out by the LTO Consortium (led by IBM) raising concerns about the extensibility and viability of the current secondary storage ecosystem
- The HDD industry is also facing roadmap challenges to increase capacity while any performance gains (IOPs) for HDDs appear to be negligible
- HSDCs (Hyperscale Data Centers) now contain over 60% of the world's installed storage capacity exerting tremendous influence on supplier development and pricing – there over 700 HSDCs WW
- The race to zero ($0.00/TB) fueled by HSDCs places supplier margins, future R&D funding, roadmaps, and innovation at significant risk and is crucial factor driving VMF
- The ability of the current secondary storage model for HDDs and tape to efficiently meet these future demands is unlikely and both will continue to struggle and fall behind
- A new disruptive storage solution is needed to overcome the increasing limitations of HDD and tape technologies

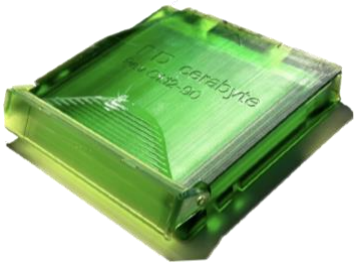
Cerabyte – A New Storage Solution Emerges to Address Mounting Secondary Storage Challenges

Finding a new technology to effectively address the slowing tape and HDD roadmap progress and fill the missing gaps in the storage hierarchy has not yielded a compelling solution until recently. Fortunately, a much better solution is under development. Cerabyte, a Munich Germany startup founded in 2020 is developing a patented data storage medium that addresses the remaining un-addressed gaps in secondary storage unfilled by HDDs and tape. Cerabyte uses ceramic nano-layers that are 50-100 atoms (10 nm) thick to store data protected against most data storage media threats: fire, flood, electrical surges, EMP, etc. Data can be safely stored in a wide temperature range of -273°C (-460 °F) to 500°C (930°F) even in a corrosive or acidic atmosphere and can run at hotter room temperatures than any conventional data center technology creating the most reliable data storage media ever developed. All magnetic media has limited lifetimes, and the data on it can wear out or degrade over time, so it must be periodically migrated. HDDs have a ~5-year lifetime, LTO tape drives are typically replaced every 8-10 years and tape media has a ~30-to-50-year lifetime. Cerabyte offers permanent data storage as ceramic media life projects to 1,000 years or more.

Assignment of specific files to platters significantly enhances performance. As an advanced, intelligent storage system, Cerabyte packs files that are expected to read together to the same platter. This minimizes the robotic access or mount time of platter travel, and the number of load and unloads. Customer application identifiers, file write times, and historical access trends are used to make informed decisions on which files should be packed together to optimize performance. Data is written and read using laser or particle beams, structuring the information in data matrices, in a similar way to QR-codes (quick-response

codes) which are two-dimensional, machine-readable, optical bar codes. The Cerabyte square-shaped data carriers can be stored in a robotic library much like today's robotic tape libraries but once a ceramic cartridge is mounted, other files on the data carrier can be accessed at disc-like (ms.) speeds. The shape of the media inside the cartridge indicates the data surfaces are square-shaped, single accessible sheets, stacked in the cartridge, both sides are ceramic-coated, and the data is permanently written in the form of holes or no holes (zeros & ones) into the ceramic layer, analogous to quasi-punched cards in nanoscale.

Media Stacked in a Cartridge **Ceramic Medium** **Data Coded in Matrix**



Cerabyte plans to have a demo system with a single write and read head unit achieving 100 MB/s write/read speeds and a storage capacity of 1PB per 19" rack in late 2024. The first product for corporate (on-prem) archiving systems is scheduled to be launched in 2025 with 500 MB/s write/read speeds and a capacity of 5 PB/rack up to ten 19" racks. In 2026-27 a 20-rack system for cloud data centers is planned to be launched with 1 GB/sec write/read speeds and a capacity of 10-30 PB/rack, which will increase over time and will achieve capacities attractive for hyperscalers by end of the decade.

Development Roadmap Until 2030

Demo Systems	**On-Prem Systems**	**Cloud Systems**	**Hyperscaler Systems**
1 PB/rack	5 PB/rack	10-30 PB/rack	60-100 PB/rack
100 MB/s	500 MB/s	1+ GB/s	2+ GB/s
<90 sec to first byte	<30 sec to first byte	<15 sec to first byte	<10 sec to first byte

To summarize, as expected, the workload for a typical archival storage system is heavily write-dominated and the I/O operations (reads) are typically dominated by small file accesses heavily favoring random access storage. Clearly there is variability in the workload for both reads and writes, both within and across different data centers; and the read bandwidth needs to be provisioned to at least handle the burstiness. Note: If the archival file(s) stored on tape requires random access, it must be moved from tape to HDDs (or SSDs) to enable access by an application.

Note: Archival Data is Mostly Write Once*

Immutable Data Properties
Active Archive - WORM (Write Once, Read Many)
Archive - WORSe (Write Once, Read Seldom)
Deep Archive - WORN (Write Once, Read Never)
****Cannot be deleted, modified or overwritten***

Cerabyte Creates a New Tier for the Active Archive - Faster Access to Archival Data

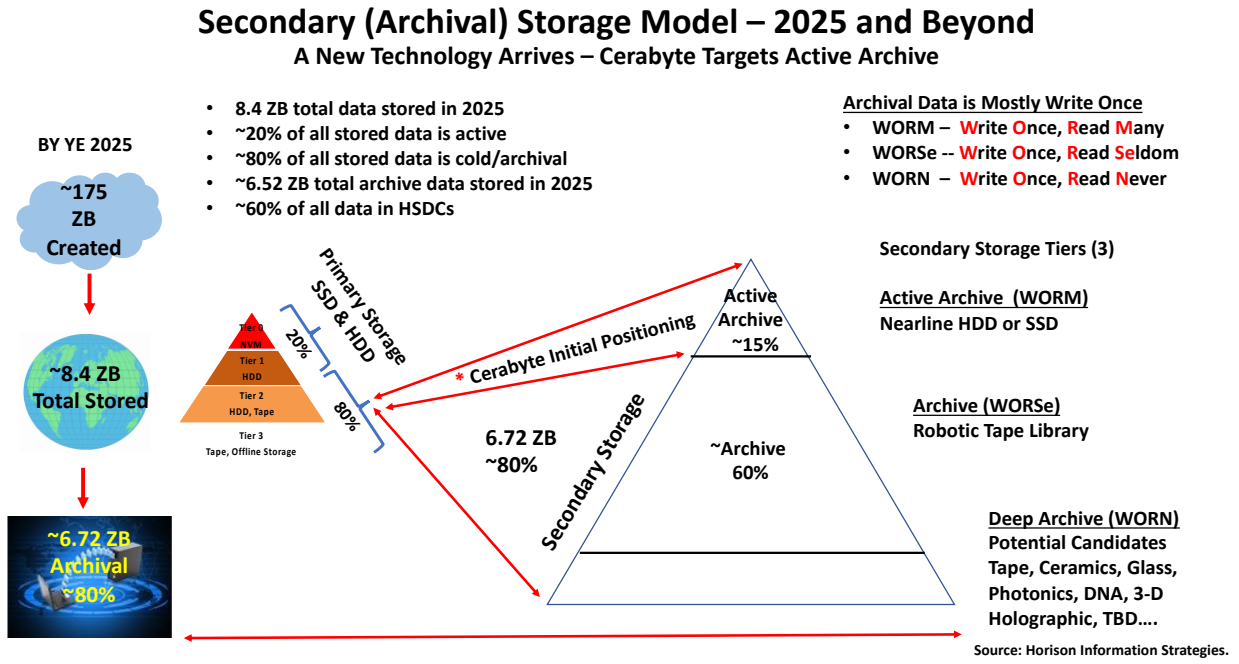
The difference between an archive and an active archive is performance. An active archive is a high-value, scalable storage architecture that gives the user faster access to archival data than traditional tape systems. Today's active archive typically consists of an HDD array serving as a cache front-end to a large capacity tape library. When tape data is *read* by a server, it is also *written* into the HDD (or possibly Flash) cache for future accesses at HDD-like ms. access speeds making the archive more active. This requires two I/O operations (a read and a write) to get the data from tape into the cache buffer. With Cerabyte random or sequential access occurs directly from the Cerabyte media requiring no additional I/O's for data placement adding to overall system performance.

The active archive has become a critical component of the secondary storage stack in recent years. The rapid rise of AI, ML, Big Data, Edge/IoT, Virtual and Augmented Reality, Gaming and Robotics are now creating and analyzing archival data at alarming rates, often completing the analysis in a few days or weeks creating the requirement for an active archive. For an active archive, the archive becomes active or used repeatedly, for a while (hours, days…) and then returns to archival status. Compared with HDDs and tape, Cerabyte fills in the remaining gaps providing a complete secondary storage solution that HDDs and tape don't address as the chart below indicates.

Secondary Storage Requirements – Key Missing Active Archive Gaps of HDDs and Tape						
GAP	Random and sequential access	Removability / Air-Gap	Energy efficient CO_2	Long-life Media	Access time to file	Lowest Cost($/TB) and TCO
HDD	Yes - Random & Sequential	No	No	No ~ 5 years	Fast (ms)	No
Tape	No Sequential only	Yes	Yes	30-50 years for media 8-10 years for drives	Slow (secs/mins)	Yes
Cerabyte	Yes - Random & Sequential	Yes	Yes	1000+ years	Medium/fast (ms, secs.)	Yes

The Secondary Storage Model -2025 and Beyond

TrendFocus projects 8.4ZBs of installed capacity by year end 2025 with ~80% or 6.72 ZB cold archival capacity best suited for secondary storage. The secondary storage stack is evolving into three tiers; the active archive tier (mainly nearline HDDs), the archive tier (mainly magnetic tape libraries) and the yet to be determined deep archive or time capsule tier. Typically, about 15% of secondary storage data is classified as active archive data or 1.008 ZB in 2025. As archival data continues attract more attention, the size of the active archive will steadily increase from 15%.



Cerabyte Takes Several Steps on The Path to Climate Neutrality

Sustainable long-term data storage is one of the most urgent IT problems in our world. Cerabyte can reduce 99% of CO_2 emissions generated by conventional Nearline HDDs used in active archive and large cloud data centers. Cerabyte does not require energy to store data, thus saving ~2% of global electricity or roughly 250 megatons of global CO_2 emissions by 2030. Most secondary storage data is continuing to pile up on high capacity Nearline HDDs which consume energy 7x24x365 and need replacement every ~5 years. This causes more than one megaton of electronic waste dumped every year to be avoided as Cerabyte media is 100% recyclable. Cerabyte media saves energy and resources before the first byte is stored: the ecological footprint of ceramic media production is significantly lower than for HDD and SSD.

Bottomline

By offering both random and sequential access, an air gap, the longest life media, lowest cost and TCO, removable disc-like access times and significant sustainability advantages, Beyond active archiving, the Cerabyte architecture is well positioned to address the remaining Archive tiers in entire secondary storage stack including the yet to be defined Deep Archive tier (~80% of all data). Finally, the gaps in the storage hierarchy can be filled and the Zettabyte era can thrive.