



Fred Moore
President
Horison.com

*Introducing a New Class of Advanced
Object Storage and Services for Cold Data*

ACTIVESCALE COLD STORAGE
FOR ENTERPRISE, WEBSCALE and
HYPERSCALE DEEP ARCHIVES

Introduction

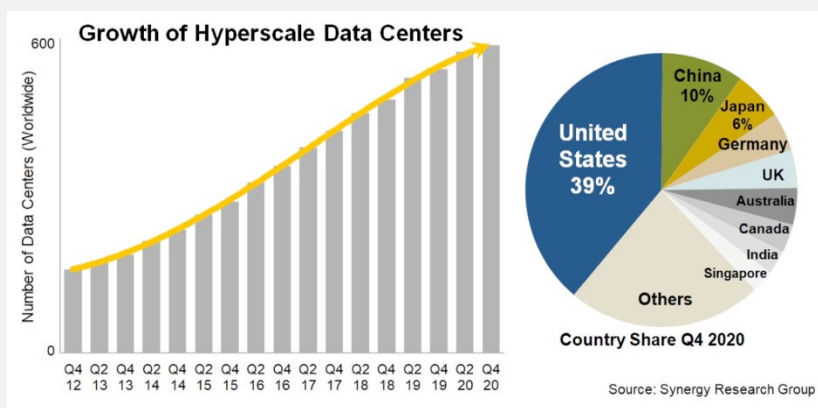
Relentless digital data growth is inevitable as data has become critical to all aspects of human life and it promises to play a much greater role in the future. The amount of data created over the next three years will exceed the total amount of data created over the past 30 years. Most of this data is piling up as archival and cold data waiting to have its potential value unlocked. This accelerated transformation has been increased by the global COVID-19 pandemic, which caused an upsurge in working-from-home, relying heavily on internet-based applications and video communication in an “everything digital” work environment. Nearly 80% of all data is archival or cold with most being stored in the traditional file format, which is facing scaling limits, and the relatively new, fast-growing object format which provides limitless scaling. This enormous growth is the product of generational and new internet-based applications including surveillance, cloud and edge computing, high-performance computing (HPC), the Internet of Things (IoT), social media, entertainment and many more.

Data storage has become far more than a megatrend – it can be viewed as a measure of the progress of humanity, cutting across all walks of life, public and private organizations, and every aspect of the economy. New storage technologies and architectures must be created in response to the growing need for highly scalable capacity at minimal cost for large enterprise, webscale and hyperscale data centers (HSDCs). As a result, today’s enormous archives will command a much more intelligent solution that leverages the advanced capabilities of intelligent data management software and high-availability, scale-out hardware. A compelling digital archive strategy has become a lifeline for enterprise, webscale and HSDCs continued growth. Fortunately, the dawn of a *new class of object storage for cold data* has arrived.

What are the Challenges with Large Enterprise and Hyperscale Cold Storage?

It’s no surprise that most HSDC challenges arise from their sheer size. HSDCs and large enterprises face insurmountable growth of disk farms which are devouring budgets, overcrowding data centers, and creating energy and carbon footprint problems, forcing archive and cold data migration to more cost-effective tape solutions. Cloud archive storage services are relatively inexpensive, but cloud data retrieval and transfer (bandwidth) costs quickly escalate when the amount of data transferred increases. HSDCs,

large enterprises and cloud providers have become the epicenter for modern archiving strategies. Amazon Glacier, Amazon Glacier Deep Archive, and Microsoft Azure are examples of specific hyperscale cloud services for cold data archiving. There are nearly 600 hyperscale data centers in the world – twice as many as five



years ago, according to the latest study by [Synergy Research Group](#). Four of the largest HSDCs have 60 or more data center locations worldwide. In addition, the Cisco Global Cloud Index Report projects HSDCs

will support 65% of all data stored in data centers in 2021. With various industry estimates projecting just over 8 ZB installed capacity by year end 2021, the magnitude of the looming enterprise and hyperscale cold storage challenge becomes obvious.

When Does Data Get Cold?

Archival and cold data is mostly unstructured and is the fastest growing and largest data classification segment. For most data types, data typically becomes archival between 90 - 120 days after the data creation date. The longer data remains static and untouched, the colder it becomes and data retention periods of 100 years or more have become commonplace. Some data can become archival upon creation and then wait for years for reference or further analysis. This adds to the cold data archival pile-up and much of this data is stored on HDDs. Keeping untouched data on spinning HDDs for long periods of time is an expensive, energy consuming storage strategy. The most cost-effective storage solution for archival data is high-capacity robotic tape libraries used in local, hybrid cloud and remote locations which can easily scale to an exabyte or more. Recent energy consumption studies show that the tape carbon emissions can be as much as 87% lower than equivalent capacity HDD systems.

Large Enterprises and Hyperscale Data Centers Rely on Object Storage for Cold Data

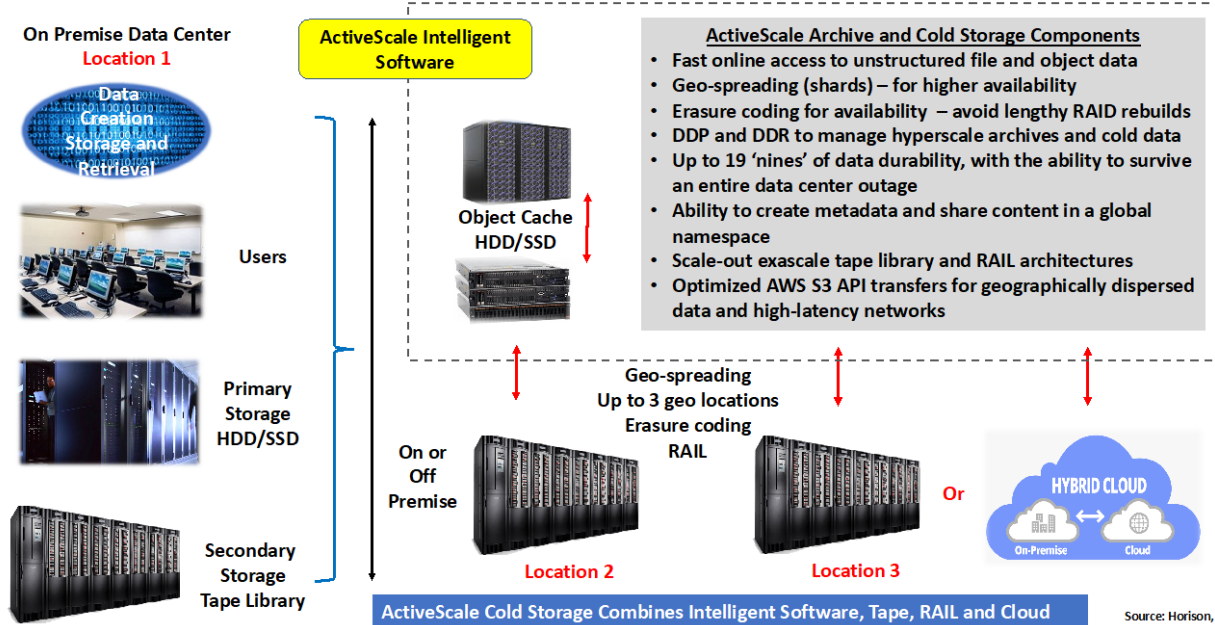
Data growth has made traditional storage formats costly and difficult to manage. Object storage became popular with the Amazon S3 cloud service and has become the de-facto standard format for cloud based archival storage services. Object storage evolved out of the need to store and retrieve any amount of data from any location using the Internet, making it ideal for cloud providers and many HSDCs. For object storage, the data blocks that make up an “object”, together with its metadata, are all packaged together. Traditional file storage can hold millions of files; however, you will eventually hit a ceiling.

Object storage provides the industry’s only limitlessly scalable architecture which can quickly scale to hundreds of petabytes in a single namespace. Object storage has surpassed the growth rate of conventional file and block storage, growing over 35% annually and that trend looks to continue.

ActiveScale Cold Storage Delivers a New Class of Object Storage for Archive and Cold Data

Key large-scale archive and cold storage challenges include easy scale-out capacity, high availability, immutability, metadata assignment, having only one copy of archival data, low TCO, geographical redundancy, and energy efficiency. Fortunately, the capability to effectively address these challenges and scale from petabytes to exabytes of unstructured and object data is now available. Quantum’s S3 enabled [ActiveScale™](#) Cold Storage intelligent software defines a new class of storage that will manage, store, and extract value from large archives of cold data. With ActiveScale, new storage nodes are added to the capacity pool and write activity is optimally distributed across storage resources using [Dynamic Data Placement \(DDP\)](#) heuristics. The DDP algorithm balances data placement and makes capacity readily available to applications. Automating these tasks bring *badly needed relief* to historically time-consuming, labor-intensive storage administrator tasks as the amount of archive and cold data to manage escalates.

ActiveScale Cold Storage – A New Class of Object Storage for Cold Data



ActiveScale presents a standard S3-compliant object interface supporting the entire S3 Glacier ecosystem command set. Enabling a critical aspect of an intelligent archiving strategy, ActiveScale allows customizable metadata to be set at the time you upload the object making it much easier to search and find relevant data essentially unlocking the archives. ActiveScale also employs a small object policy that optimizes the system for read performance by maintaining a full copy of the object on a single drive. For systems under high parallel processing loads, this can allow the system to store more objects simultaneously and lower the aggregate latency of the write process.

Quantum’s New Object Storage Services Bring the Cloud Experience to On-Premises Environments

Besides the typical deployment of hardware on premises, Quantum has also introduced a new, innovative line of petabyte scale object services for active archive and cold data based on the ActiveScale solution set. These new services are purpose-built for large enterprises, government agencies, cloud service providers and research organizations. With two classes of service and no data access fees, these fully managed services should be seriously considered as a highly cost-effective, on-premises solution for both active archiving and cold data storage.

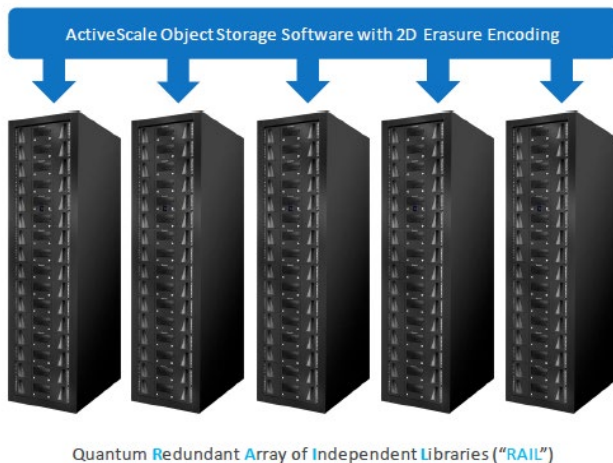
Object Storage Services	For Active Data	For Cold Data
Data Retrieval Time	Sub-milliseconds	Typically under 5 minutes
How data is accessed	S3 Standard Class	S3 Glacier Class
Durability	Up to 19 9's	Up to 19 9's
Minimum Commitments	1 Petabyte Capacity Three-year Term	4 Petabyte Capacity Three-year Term
List Price	\$.015 per GB per Month No Data Access Fees	\$.003 per GB per Month No Data Access Fees

ActiveScale Geo-Spread Configurations Provide Extreme Availability

HSDCs often physically locate servers and storage in geographically dispersed data centers to create a higher availability safeguard known as geographical redundancy. This approach provides business resiliency against catastrophic events, such as fires, floods, hurricanes, or other situations that might bring a data center down. ActiveScale GeoSpread systems using HDDs and/or tape using RAIL (Redundant Arrays of Independent Libraries - introduced by Quantum) can be deployed across three geographically (3-geo) dispersed data centers creating object store data lakes that can tolerate component and site failures. GeoSpread uses error correction with [rateless erasure coding](#) to non-disruptively distribute data across nodes and across geographies. Object data can be accessed across nodes in parallel, increasing aggregate performance for data transfer bandwidth.

Since a 3-geo system must be able to recover objects if one site becomes unavailable, no more than a third of the chunks can be in a single datacenter. With DDP's hierarchical spreading enabled for each level, the system will equally balance the chunks over 18 drives across three data centers. The erasure code storage policy used is referred to 18/8, which results in objects spread across 18 disk drives with a disk safety of 8. Objects can be decoded from any subset of 10 encoded chunks. Critical data and applications can be mirrored to a single, hybrid or multiple cloud service providers for redundancy. If a failure occurs, the on-premises data center can fail over to one of the hybrid-cloud mirrors to maintain availability and then fail back when the incident is resolved.

RAIL and Erasure Coding Define Advanced Hyperscale Archive and Cold Storage Architecture



RAILlike RAIT (Redundant Arrays of Independent Tape), stripes data across different tape cartridges but in different libraries. Using advanced erasure coding, RAIL provides increased geographic resiliency and extreme availability in case of a full data center outage or natural disaster. RAIL enables a durable, high availability, secure and low-cost storage architecture for large archives and cold data for hyperscale data centers.

Creating multiple copies of archival and cold data with RAID for HSDCs and large-scale data centers is giving way to more cost-effective erasure coding and RAIL architectures. HDD rebuild times using traditional RAID have become excessive as HDD capacities increase sometimes taking more than a week to rebuild a failed HDD while increasing exposure to data loss. ActiveScale supports advanced erasure coding (HDD-based) and now RAIL (tape-based) configurations where data is dispersed across three

locations for extreme availability with ActiveScale’s [Dynamic Data Placement](#) technology (DDP). Even with a full data center outage, the 3-site configuration delivers continuous data availability and sustained operations. ActiveScale has implemented [Dynamic Data Repair \(DDR\)](#) which performs data integrity audits and automated repair functions to address potential media degradation and sends alerts to the systems management layer. Agents are installed on nodes for out of band management and do continuous background integrity checking and healing. System monitoring tasks include environmental parameters (temperature, fan) as well as disk health statistics to provide additional levels of availability.

Benefits of ActiveScale Geo-spreading Using Tape and RAIL Architectures

Immutable Objects	WORM and Encryption for Object Storage Archives Using Tape.
Security	Tape Air Gap Provides Defense Against Cybercrime, Ransomware Attacks.
Portability	LTO Media Easily Portable in Case of Disaster, HDDs Difficult to Physically Move.
Energy, CO2 Sustainability	Tape Uses Much Less Energy and Has Much Lower Carbon Emissions Than HDDs Less Energy Consumed Results in Lower Carbon Emissions.
Tape and Cloud Ecosystem	Native Cloud Applications Can Write to and Read from Tape Without Specialized or Proprietary Software.
Archive Durability, High Availability	LTO Reliability Bit Error Rate (BER 1×10^{19}) Has Surpassed HDDs (BER 1×10^{16}). Media Life >30 Years for all Modern Tape. Advanced Erasure Coding with RAIL
Geo-spreading	RAIL, DDP, and DDR Provide Advanced Hyperscale Availability Services.

The ActiveScale Cold Storage system can scale objects from petabytes to exabytes while providing the availability, durability, and performance needed for those hyperscale use cases. Modern tape architectures using RAIL provide the most cost-effective scaling, the lowest TCO and carbon footprint, the highest reliability, the fastest throughput, and cybersecurity protection via the air gap. The combined value of these benefits is expected to increase for tape as the preferred storage technology going forward as large-scale cold and archival data requirements show no sign of slowing.

Conclusion

The arrival of Quantum’s S3-enabled [ActiveScale](#) Cold Storage intelligent software signals a new class of object storage that can manage, store, and scale from petabytes to exabytes of cold data. Large enterprises, webscale and hyperscale companies experience the same cold storage challenges and should seriously consider ActiveScale as a highly cost-effective solution providing unlimited capacity. Large-scale data centers represent the fastest growing data center segment today and are pushing carbon footprint, energy demand, and sustainability concerns to center stage as storage demands continue to soar. With no slowdown in sight, re-engineering archive and cold storage strategies to more cost-effectively contain extreme data growth and unlock untapped data value is a critical goal. The ability to connect, analyze, protect, and act on all your data, wherever it is located, will be the key to success. The optimal storage architecture must simplify complexity and help organizations tap the value of their cold data without forcing budgets to scale at the same pace as data growth. The new class of object storage provided by Quantum’s ActiveScale for Cold Storage is ready to meet this challenge.

About the Sponsor

Quantum.

Quantum technology, software, and services provide the solutions that today's organizations need to make video and other unstructured data smarter – so their data works for them and not the other way around. With over 40 years of innovation, Quantum's end-to-end platform is uniquely equipped to orchestrate, protect, and enrich data across its lifecycle, providing enhanced intelligence and actionable insights. Leading organizations in cloud services, entertainment, government, research, education, transportation, and enterprise IT trust Quantum to bring their data to life, because data makes life better, safer, and smarter. Quantum is listed on Nasdaq (QMCO) and the Russell 2000® Index. For more information visit www.quantum.com.

About the Author



[Horison Information Strategies](#) is a data storage industry analyst and consulting firm specializing in executive briefings, industry seminars, market strategy development, whitepapers and research reports encompassing current and future storage technologies. Horison identifies disruptive and emerging data storage trends and growth opportunities for end-users, storage industry providers, and startup ventures.

© Horison Information Strategies, Boulder, CO. All rights reserved.

WP00268A-v01 Oct 2021