HORISON
Information Strategies
Fred Moore, President
Horison.com

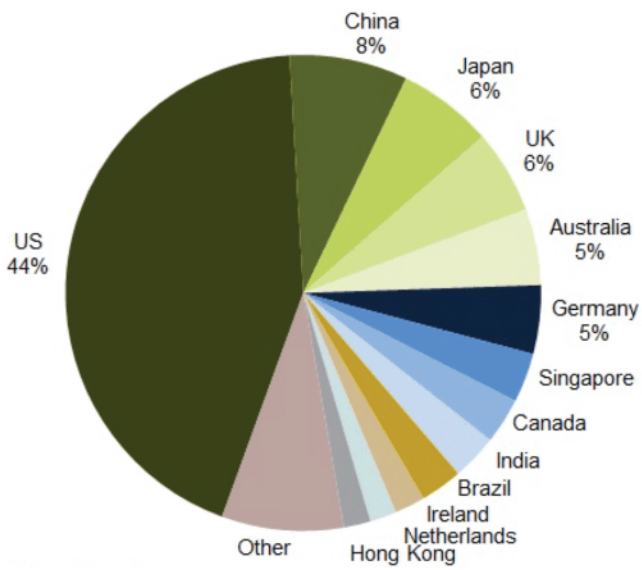
THE ASCENT TO **HYPERSCALE**

Preparing for the Next Wave of Hyperscale Storage Challenges

Started by a few internet and cloud providers in the United States, hyperscale data centers (HSDCs) have now spread across the globe to meet unprecedented data storage requirements. According to the Cisco Global Cloud Index Report, the world's HSDCs are poised to grow from 338 in 2016 to 628 by 2021. That means 290 "hyperscale-lite" datacenters are ascending to become full HSDCs and will begin to experience many of the extreme "startup to scale" challenges of their HSDC predecessors. As the expense and volume of data grows relentlessly every year, the need for more economical and advanced storage solutions to contain this demand is growing in parallel. HSDCs are now ready to catch up and take advantage of the economics of tape at scale.

HYPERSCALE DATA CENTER OPERATORS

Data Center Locations by Country - December 2017



Source:
Synergy Research Group

WHAT ARE HYPERSCALE DATA CENTERS?

The term "hyper" means extreme or excess. While there isn't a single, comprehensive definition for HSDCs, they are significantly larger facilities than a typical enterprise data center. The Synergy Research Group Report indicated there were 390 hyper-scale data centers worldwide at the end of 2017. An overwhelming majority of those facilities, 44% are in the US with China being a distant second with 8%. Currently the world's largest data center facility has 1.1 million square feet. To put this into perspective the standard size for a professional soccer field is 60,000 square feet, the equivalent to about 18.3 soccer fields. Imagine needing binoculars to look out over an endless array of computer equipment in a single facility. Imagine paying the energy bill!

HYPERSCALE refers to a computer architecture that massively scales compute power, memory, a high-speed networking infrastructure, and storage resources typically serving millions of users with relatively few applications. While most enterprises can rely on out-of-the-box infrastructures from vendors, hyperscale companies must personalize nearly every aspect of their environment. A HSDC architecture is typically made up of tens of thousands of small, inexpensive, commodity component servers or nodes, providing massive compute, storage and networking capabilities. HSDCs are implementing Artificial Intelligence (AI), and Machine Learning (ML) to help manage the load, and are exploiting the storage hierarchy including heavy tape usage for backup, archive, active archive and disaster recovery applications.

Finding the lowest TCO

CASE STUDY

This CSP, with hundreds of millions of customers, recognized a need to reduce the expense of storing archival and regulatory internal digital data. The internal data storage, that helps drive operations, often reaches massive capacities. In the case of this CSP, it was recognized that hundreds of petabytes were simply sitting in a cold state on operational HDD infrastructure. As a multi-patent holder for data center innovation, this CSP determined that continuing to innovate in mass data storage was the only way to significantly reduce the expense.

Using internal analytics, this CSP combed the data to determine the touch rate and retrieval requirements of hundreds of "cold" petabytes, finding that 90% of the data was not being touched. This set the stage for determining Service Level Agreements (SLA) and for integrating the best performing, lowest Total Cost of Ownership (TCO) based solution. The answer was tape infrastructure, with open format data.

Performance in this case was not to serve SLAs, it was to ensure that data could be streamed to the devices as fast as it was being produced and going to a cold state. Knowing that this data would remain in the infrastructure for years and even decades, the infrastructure team focused on getting to a TCO that demonstrated orders of magnitude of savings below the HDD solutions. The CSP infrastructure team deployed automated tape infrastructure with open format LTFS for storing this data across their data centers, at a calculated TCO savings of nearly 500% over a 10 year period.

Get your free custom TCO report from Brad Johns Consulting:

HYPERSCALE-LITE are large-scale enterprise data centers representing the next wave of hyperscalers. The key to building a hyperscale architecture is to start small to keep upfront investments as low as possible. As demand grows, the HSDC-lite infrastructure should be able to expand easily by adding nodes to the cluster. This is an ideal scaling model for subscriber driven organizations because they can grow the physical data center at the same pace as they add customers. Advanced tiered storage and scale out software are specifically designed make node aggregation and capacity scaling as automated as possible.

Source: Cisco Global Cloud Index, 2016-2021.

HYPERSCALE-LITE ASCENDS TO HYPERSCALE – OVER 600 HSDC'S BY 2021?

Several innovations are driving the ascent from hyperscale-lite to hyperscale: The internet, cloud computing , AI, ML, big data, the growing acceptance of social media, gaming, online shopping and the yet unknown IoT requirements. By 2021, the 628 HSDCs are projected to account for 53 percent of all installed data center servers worldwide. The biggest cloud providers (Amazon, Google, IBM and Microsoft) operate the largest footprints. Each of these four hyperscale cloud companies has at least 45 data center locations, at least three of them per region (North America, Latin America, APAC, and EMEA). The Synergy Report (mentioned earlier) indicated that 24 of the world's major cloud and internet service firms have 16 data center sites on average with tens of thousands of servers.

CHARACTERISTICS OF THE HYPERSCALE DATA CENTER

HSDCs don't publicly share an abundance of information about their infrastructure. For companies who will operate HSDCs , the cost may be the major barrier to entry, but ultimately it isn't the biggest issue - automation is. HSDCs must focus heavily on automating and self-healing environments by using AI and ML whenever possible to overcome inevitable and unexpected failures and delays. Unlike many enterprise data centers, which rely on a large full-time staff across a range of disciplines, HSDCs employ fewer tech experts because they have used technology to automate so much of the overall management process. HSDC characteristics include:

- **Small footprint, dense racks** – HSDCs squeeze servers, SSDs (Solid State Disks) and HDDs (Hard Disk Drives) directly into the rack itself, as opposed to separate SANs or DAS to achieve the smallest possible footprint (heavy use of racks). HSDC racks are typically larger than standard 19" racks.

It is estimated that Google has at least 2 million servers in all its data centers around the world.

- **Automation** – Hyperscale storage tends to be software-defined, and is benefitting from AI delivering a higher degree of automation and self-healing minimizing direct human involvement. AI will support automated data migration between tiers to further optimize storage assets.
- **Users** – The HSDC typically serves millions of users with only a few applications, whereas in a conventional enterprise there are fewer users but many more applications.
- **Virtualization** – The facilities also implement very high degrees of virtualization, with as many operating system images running on each physical server as possible.
- **Tape storage adoption** – Automated tape libraries are on the rise to complement SSDs and HDDs to easily scale capacity, manage and contain out of control data growth, store archival and unstructured data, significantly lower infrastructure and energy costs, and provide hacker-proof cybercrime security via the tape air gap.
- **Fast scaling bulk storage** – HSDCs require fast, easy scaling storage capacity. One petabyte using 15 TB disk drives requires 67 drives and one exabyte requires 66,700 15 TB drives. Tape easily scales capacity by adding media, disk scales by adding drives.
- **Minimal feature set** – Hyperscale storage has a minimal, stripped-down feature set and may even lack redundancy as the goal is to maximize storage space and minimize cost.
- **Energy challenges** – High power consumption and increasing carbon emissions has forced HSDCs to develop new energy sources to reduce and more effectively manage energy expenses.

Maximizing data protection efficiency
CASE STUDY

This CSP has more than 20 major data centers across the globe with exabytes of online data. LTO based digital tape technologies are deployed to ensure all of their data is backed-up and recoverable. The challenges associated with protecting the sheer amount of data extends well beyond the technical, as cost becomes a major variable.

When managing at hyperscale, the best strategies are to employ the most efficient processes in both software and hardware including all of the overhead costs associated with infrastructure and management. First and foremost, this CSP started with a few company-wide principles that guide all projects and their data sets with the first being that all data must be backed-up. Given this principle and the hyper-scale of their data, this becomes a seemingly daunting task. However, if the data is not worth backing-up, the project is not worth doing!

The second principle applied by this CSP is that all back-ups must be tested by restoring samples of the data. This principle ensures that issues with recovery of data are found before the data needs to be recovered. While these guiding principles carry a considerable cost, this burden is far outweighed by the damage of unrecoverable client data.

Data protection systems at this CSP are distributed across multiple sites for fault tolerance and geographic protection with large enterprise LTO tape libraries on each site. Back-ups may be written to any site and Map Reduce is used for process automation. Data is sharded utilizing 20+8 Reed Solomon erasure coding with 40% protection overhead. Data is written to tape using RAIT 4+1 (similar to RAID 4) with data written to 4 tape drives at once. Data on tape is calculated across all tapes to create a 5th parity tape that is tied to the original data set. RAIT and parity represents a 25% overhead for protection.

This CSP utilizes an archive process to retain data for longer periods of time with Archive retention periods set by the client. Archived data is automatically migrated in "N" year to a new generation of media typically within a 10 year period. If a client has set archive retention for 21 years and "N" = 10 years then the archived data would be automatically migrated twice in the span of the data lifecycle before deletion.

The efficiency of this CSPs tape based data protection and archive strategy has resulted in a cost-effective and highly reliable solution. This solution provides greater recoverability while protecting against multiple failures with a total of 75% overhead compared to 100% for a single exact copy of data or 200% for two copies.

HYPERSCALE STORAGE CHALLENGES BECOME EXTREME AT SCALE

If data is the new currency, then storage is the new bank and the sheer size of HSDCs makes storage management a most critical challenge for HSDC architectures. For HSDCs, storage is ideally installed in each node as it is added to the cluster and is then pooled across all the nodes by hypervisor storage software. HSDCs and most IT staffs are under even greater strain as the volume and complexity of managing daily workloads defy the traditional approach of simply adding more expensive disk drives when capacity is maxed out. Too often most of this data is stored on the most expensive storage tiers. Lower cost storage options are available, such as high capacity tape and the cloud which are the optimal solutions for infrequently accessed and cold data. Enabling data centers to get the right data in the right place at the right time is critical and even more so at hyperscale levels.

HDDs have played the major role in capacity scaling and have recently been joined by flash SSDs, but the magnitude of hyperscale storage requirements and costs are shifting more focus to the economics of scale. Tape is playing a fast-growing role in containing HDD/SSD growth for HSDCs as it scales easily and enables ILM (Information Lifecycle Management) to be highly cost-effective by migrating less-active data on disk to tape containing costs in more costly online powered storage. ILM recognizes that the value and characteristics of data itself changes over time and that it must be managed accordingly by establishing policies to migrate and store data on the appropriate storage tier. Traditional storage management techniques have left data centers struggling with as much as 80% of their data being stored on the wrong tier of storage costing organizations millions of dollars per year.

- **Flash SSD challenges** – Flash and solid-state memories now play a major role in accelerating HSDC performance and are much faster, use less power and are easier to manage than HDDs. The Flash Translation Layer (FTL) intelligently manages everything from caching to performance, wear leveling, and to garbage collection, etc. At hyperscale levels, the sheer amount of SSD device management overhead impacts throughput, latency, and cost. Until the next non-volatile memory architecture arrives, the flash overhead will be tolerated as the performance gains are compelling.
- **Disk drive challenges** – Once hyperscale-lite storage reaches hyperscale levels, traditional RAID- and replication recovery architectures can become too expensive and unmanageable. Also, the higher the disk capacity, the longer the rebuild time takes just to restore/recover a failed drive to regain redundancy. A 4 TB disk will take at least 10 hours to rebuild – a 15 TB disk can take several days or even weeks. For most HSDCs, RAID is becoming obsolete or soon will be. With RAID reaching its practical limits, Erasure Coding has emerged as an alternative to RAID in which data is broken into fragments that are stored across different geographical locations using a single namespace.

Leveraging the capacity roadmap

CASE STUDY

A global Cloud Service Provider (CSP) seeked to improve margins on their deep archive storage business. At the same time they wanted to improve their competitive position by offering the lowest possible cost per Gigabyte per month.

This CSP, with over 15 data centers globally has consistently seen services related to archive type retention data growing at over 15% annually. Recognizing that this part of the business model is dependent on the lowest cost infrastructure, this CSP looked for new ways to reduce overall archive storage costs.

This CSP initially partnered with industry storage makers to reduce the cost of storage through traditional unstructured data methods. The internal development project resulted in a dramatic improvement in density and accessibility in the marketable product. However, the results from operational testing were less than desirable, falling significantly short in several critical categories: cost per GB of the storage media, overall access performance for large data sets, and hardware durability. This precipitated new thinking.

This CSP engaged with the leading tape industry professionals to expand on the well documented tape roadmap for capacity. After a deep dive in to methodologies, this CSP implemented a single site archive based on LTO tape technology. With a focus on usability, data center integration and cost per GB, this CSP tested the product capability and developed erasure based writes to tape. The calculated economy of scale provided by the tape infrastructure enabled this CSP to lower the public pricing of their solution by just over 28% in a single year, while maintaining the gross margins in the significantly competitive “cold data” IaaS market.

- **Equipment acquisition and TCO** – Heavily favors tape for lowest cost/TB and TCO of any storage solution. The HDD TCO at scale is typically 5-8x higher than tape for equivalent storage.
- **Soaring energy costs** – HSDCs have decommissioned older power-hungry server hardware and embraced more efficient technologies like server virtualization and tiered storage. The fastest way to lower storage energy costs is to move low activity data from disk to tape as tape has the lowest energy cost of any storage solution.
- **Security, cybercrime protection** – Tape provides hacker-proof cybercrime security easily achieved via air gap.
- **Hardware upgrade cycle frequency** – Disk lasts 4-5 years before replacement. Tape drives typically last 7-10 years. Modern tape media life is rated at 30 years or more.

THE VALUE OF TAPE RISES RAPIDLY AS HYPERSCALE DATA CENTERS GROW

Today HSDCs are leveraging the many advantages of tape technology solutions to manage massive data growth and long-term retention challenges. Keep in mind most digital data doesn't need to be immediately accessible and can optimally and indefinitely reside on tape subsystems. Some data requires secure, long-term storage solutions for regulatory reasons or due to the potential value that the data can provide through content analysis at a later date. Advanced tape architectures allow HSDCs to achieve business objectives by providing data protection for critical assets, backup, recovery, archive, easy capacity scaling, the lowest TCO, highest reliability, the fastest throughput, and cybersecurity protection via the air gap. These benefits are expected to increase for tape in the future.

Fighting the cybercrime epidemic has become a major problem for most data centers and HSDCs are no exception. Tape can play a key role in its prevention and provides WORM (Write-Once-Read-Many) and encryption capabilities providing a secure storage medium for compliance, legal and any valuable files. Tape, as an "Air Gap" solution, has gained momentum providing an electronically disconnected copy of data that prevents cybercrime disasters from attacking data stored on tape. Disk systems remaining online 7x24 are the primary target as they are always vulnerable to a cybercrime attack.

HSDCs are taking advantage of tiered storage by integrating high-performance SSDs, HDD arrays and automated tape libraries. Even though HSDCs are struggling with the exploding growth of disk farms which are devouring IT budgets and overcrowding data centers, many continue to maintain expensive disks often half full of data which often has little or no activity for several years. Obviously, few data centers can afford to sustain this degree of inefficiency. The greatest benefits of tiered storage are achieved when tape is used as its scalability, lower price and lower TCO plays an increasing role as the size of the storage environment increases. For the hyperscale world "adding disk is tactical – adding tape is strategic."

Areal density refers to how many bits of information can be stored on a given surface area of a magnetic disk drive or tape media. On April 9, 2015 Fujifilm in conjunction with IBM

HYPERSCALE-LITE HIGHLIGHTS

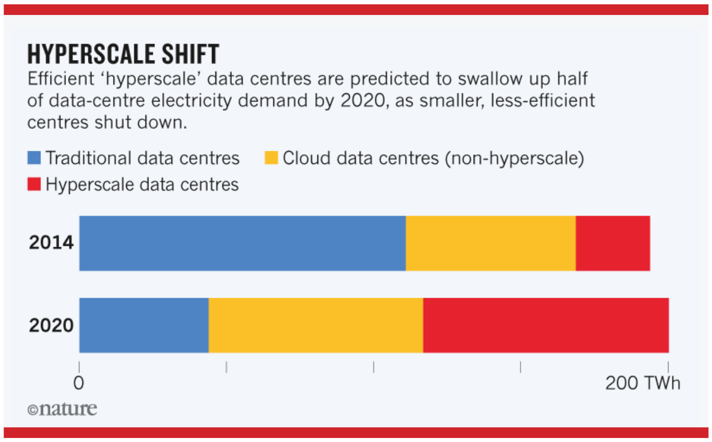
- Hyperscale-lite represents large-scale data centers becoming the next wave of hyperscalers.
- The shift to Hyperscale-lite is fueled by the migration of many smaller data centers to fewer, but much larger.
- Electricity is the HSDC lifeblood – "Without electricity, there is no IT industry".
- As the HSDC grows, tape will become mandatory to contain petascale and exascale storage environments as sustaining these capacities on disk will become prohibitive.
- "Remember that tape scales by adding more media and disk scales by adding more drives."
- "Adding disk is tactical – adding tape is strategic."
- "The ascent to hyperscale will soon make tape mandatory for sheer economic survival."

demonstrated (not announced) a new record in areal density of 123 Gb/in^2 on linear magnetic particulate tape. More recently Sony and IBM demonstrated 201 Gb/in^2 with potential for a 330 TB native cartridge and Fujifilm's Strontium Ferrite next-generation magnetic particle promises more than 400 TB (33.7 times more storage capacity than LTO-8) on a cartridge with an areal density of approximately 224 Gb/in^2. These areal densities coupled with the scalability of tape promise to deliver significant competitive advantages over HDDs for the foreseeable future.

In addition, the OCP (Open Compute Project) has an Archival Storage sub-project that will focus on unique archival solutions including tape. The solutions and contributions for improving archival storage solutions can range from entire system designs down to testing methodology used to characterize a media type. Projects will include solutions that are optimized for hyperscale deployments with the unique workloads and datacenter designs.

KEY POINTS TO REMEMBER ABOUT MODERN TAPE:

- tape is less expensive to acquire ($/TB) and to operate (TCO) than disk
- tape is more reliable than disk by at least three orders of magnitude
- the media life for modern tape is 30 years or more for all new media
- tape drive data rates have reached 400 MB/sec, HDDs transfer data at 160-220 MB/sec
- tape libraries are delivering intelligent, faster, and more efficient robotic movement
- the INSIC roadmap and the 10-year LTO roadmap is well defined with few foreseeable limits



HYPERSCALE ENERGY ISSUES MOUNT

HSDC are energy consumers. High-density, multi-core data center servers typically use between 500 and 1,200 watts while HDDs use about 6-15 watts per hour, approximately three times more than SSDs. A typical desktop computer uses between 65 and 250 watts per hour. Reducing the number of servers and moving low-activity data from disk to tape present the greatest HSDC energy savings opportunities.

CONCLUSION

The HSDC represents the fastest growing data center segment today. Today's hyper-scale companies are re-engineering storage strategies to manage extreme data growth and long-term data retention requirements and are now ready to catch up and take advantage the economics of tape at scale.

The many rich tape technology improvements of the past 10 years suggest that tape will continue to be the most cost-effective storage solution for the unprecedented HSDC challenges ahead. The time to evaluate and prepare for transitioning to a new storage architecture is best done well before the requirement has arrived. For hyperscale-lite data centers currently managing petabytes of data, the best time to evaluate new architectures may already be in your rear-view mirror. Reaching hyperscale status won't just creep up on you - it will run over you. As a result, "the ascent to hyperscale and beyond will soon make tape mandatory for sheer economic survival."

ABOUT THE SPONSORS

FUJIFILM Recording Media U.S.A., Inc. delivers breakthrough data storage products based on a history of thin-film engineering and magnetic particle science such as Fujifilm's NANOCUBIC and Barium Ferrite technology. Our mission is to enable organizations to effectively manage the world's exponential data growth with innovative products and solutions, recognizing the social responsibility to preserve digital content for future generations.

International Business Machines Corporation is an American multinational information technology company headquartered in Armonk, New York, with operations in over 170 countries. IBM offers a full range of tape storage products including drives, autoloaders, libraries, virtual tape systems, IBM Spectrum Archive software and Hybrid solutions.

Horison Information Strategies is a data storage industry analyst and consulting firm specializing in executive briefings, 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.