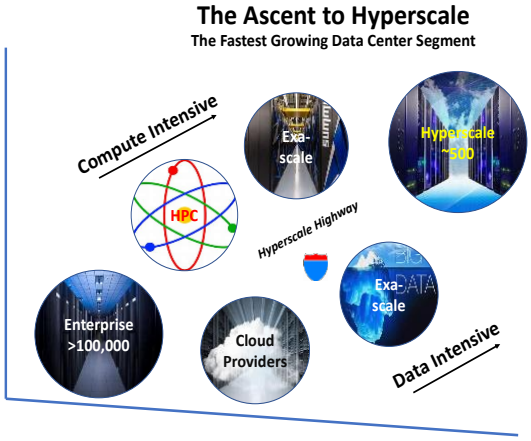
Introduction

HSDCs (Hyperscale Data Centers) have become the fastest growing data center segment generating massive compute, storage, energy and carbon footprint requirements. These are often built by cloud and internet giants, large enterprise data centers, and companies that specialize in designing and building data centers and leasing them to others. Though many HSDCs are nearly twice as energy efficient as traditional enterprise data centers, they still require unprecedented amounts of energy, advanced cooling technologies, from multiple energy sources. Energy hungry server and HDD farms are forcing HSDCs to move as much low-activity, archival and cold data to tape as possible to minimize energy consumption. As a result of these forces, tape has become a pressure relief valve for unabated HSDC expansion.

The Hyperscale Landscape

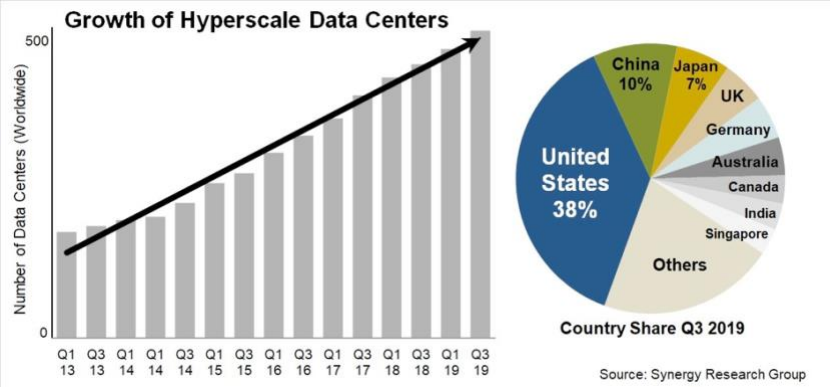
Many data center segments are beginning their ascent to hyperscale status including large enterprises, HPC facilities, the next wave of cloud providers, along with exascale compute and storage facilities. A hyperscale data center is less like a warehouse and more like an enormous fulfillment center or distribution hub. The Citadel is presently the largest data center campus in the world located in Tahoe, Nevada covering 7.2 million square feet or 88.9 soccer fields (360'x225'). Although hyperscale facilities are extremely large, hyperscale is actually more about scalability than size.



Hyperscale Data Center	Massive scaling
Hyper-scale lite (HSL)	The next wave of Hyperscalers
Exascale Computing	One exaFLOPS, or ($1x10^{18}$)
Exascale Storage	One exabyte ($1x10^{18}$) of storage
HPC	Compute Intensive
Cloud	Private, public, hybrid
Enterprise	Large footprint, many apps

Source: Horison, Inc.

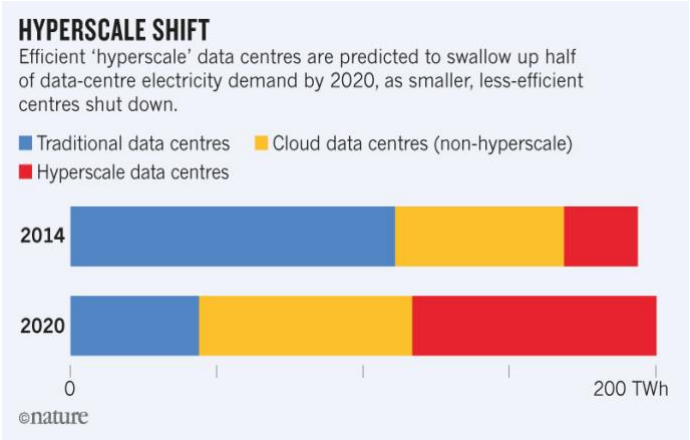
The most recent Synergy Research report states there are 541 HSDCs worldwide, up from 500 in October 2019. There are 176 currently under construction or in planning, while 100 new HSDCs opened in the past two years. The companies with the largest data center footprint are the leading cloud providers including Amazon, Google, IBM and Microsoft each having 60 or more data center locations with at least three in each of the four global regions – North America, APAC, EMEA and Latin America. These companies collectively control over half of the WW CSP (Cloud Service Provider) market.



Hyperscale Energy Demand Accelerates

The explosion of digital content is literally setting the data center market on fire and has made HSDCs the fastest-growing consumers of electricity. 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. Most all tape cartridges spend the vast majority of their life in a slot or rack, electronically disconnected from a computer system consuming *no* energy. The greatest HSDC energy savings opportunities are physically reducing the number of servers and moving TBs and PBs of low-activity and cold data, typically 60% or more of all stored data, from HDD to tape. Fortunately, all of these storage technologies have made significant energy efficiency improvements in recent years helping slow the rapid rate of energy consumption.

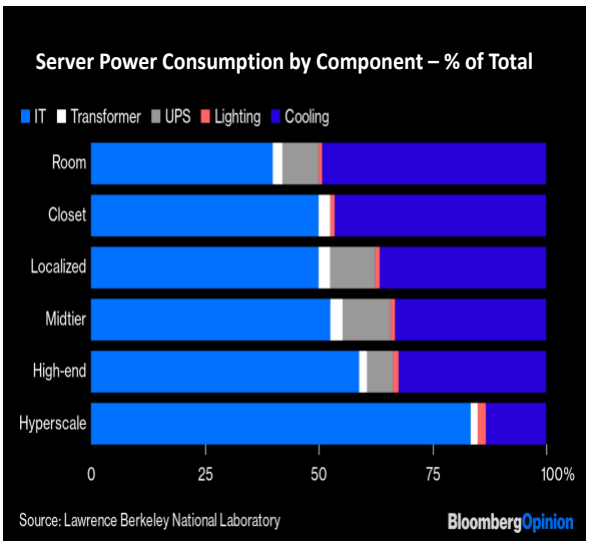
Data centers consume about 2% of the world's electricity currently and is expected to soar up to 8% by 2030. By 2030, the combined information and communications technology industries are projected to consume 20.9% of the entire global electricity supply! A key selection criterion for the geographic location of new HSDCs is often related to the location of fossil fuel and renewable energy facilities. Hyperscale energy demand is providing valuable insights into what the future of computing might look like. As more and more data centers migrate their IT operations data to HSDCs and CSPs, it's possible one day that the vast majority of servers and storage will be reside with them. Maybe then computing would eventually function like an electric utility provider and be the information utility where you just plug in and pay for what you use?



Managing Hyperscale Energy Efficiency

The single largest operational expense in most data centers worldwide, more than powering energy hungry servers, is powering the climate control systems. When a tenant or customer leases space from a HSDC or CSP, that space is often described and contracted in kilowatts rather than in numbers of racks or square footage. This helps ensure that kilowatts are available when a customer needs them. HSDCs have very strong business incentives to conserve energy since the lower amount of energy used, the higher the provider's profit margin. Most HSDCs are equipped with redundant power sources given their complete dependency on an uninterrupted energy supply. *Without electricity, there is no IT industry.*

Power Usage Effectiveness (PUE) is a measure of how efficiently a data center uses energy. PUE is defined as the total energy needed for everything, including lights, signage, surveillance and cooling, divided by the energy used for computing. A PUE of 1.0 would be a perfect score. The energy efficiency progress made by HSDCs has been remarkable. Conventional data centers typically have a PUE of about 2.0 compared to hyperscale at about 1.2. The use of tape will be key to sustaining this level of efficiency.



PUE	DCiE	Data Center Efficiency Level
3.0	33%	Very Inefficient
2.5	40%	Inefficient
2.0	50%	Average
1.5	67%	Efficient
1.2	83%	Very Efficient
1.0	100%	Max Efficiency

PUE = Total Facility Power / IT Equipment Power

DCIE = IT Equipment Power / Total Facility Power

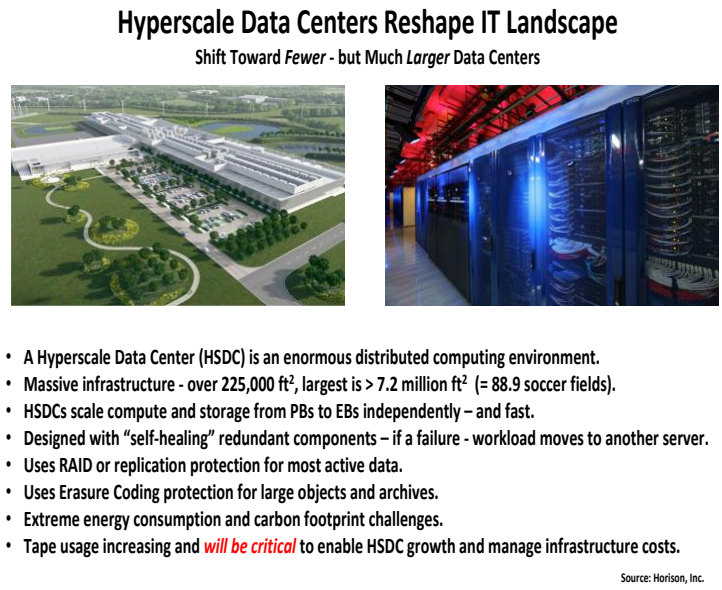
Power (Watts) $P = V \times I$ or $I^2 \times R$
Voltage (Volts or V) $V = I \times R$
Current (Amps or I) $I = V \div R$
Resistance (Ohm) $R = V \div I$
The kilowatt hour (kWh) is a unit of energy equal to 1,000 watt-hours

Data Center Infrastructure Efficiency (DCIE) is the inverse of PUE. DCIE is the percentage value derived by dividing computing equipment power by total facility power. The ideal PUE would be 1.0 and DCIE 100%. Anything that isn't considered a computing device in a data center falls into the category of facility energy consumption.

For years, most data center equipment such as air conditioners, chillers, cooling towers and water pumps were controlled by standard BMS (building management systems) to ensure a normal and safe operation. HSDCs are taking energy management to the next level by implementing Artificial Intelligence (AI), and Machine Learning (ML) to create a more precise PUE prediction model to anticipate and reduce their energy consumption load and to more cost-effectively exploit the storage infrastructure and take advantage of energy efficient tape technology. Relentless data demands, and the unknown impact created by billions of IoT sensors means that everyone including data center operators, equipment manufacturers, data consumers and energy policymakers must intensify efforts to accommodate a potential sharp rise in energy usage later this decade. Every day, the Earth's natural resources are converted into digital data as coal, oil, natural gas, copper, silicon and aluminum are used to build and power the massive computer farms that keep the digital economy moving forward. Careful HSDC management of these resources is paramount to continuing this trend.

The Value of Tape Rises as Hyperscale Data Centers Take Off

HSDCs are implementing solutions to manage massive data growth challenges by leveraging the numerous advantages of tape technology. Advanced LTO tape architectures allow HSDCs to achieve business objectives by providing data protection for critical assets, backup, recovery, archive, DR, quick capacity scaling, the highest reliability, and cybersecurity protection via the air gap. These benefits are expected to increase for tape in the future as HSDCs aggressively push to produce sustainable energy in such a way that it meets the needs of the present without compromising the ability of future generations to meet their needs.



Technology improvements over the past decade have vaulted LTO tape into the top reliability spot with a BER (Bit Error Rate) of $1x10^{20}$, four orders of magnitude more reliable than HDDs at $1x10^{16}$. Media life for LTO is rated at 30 years or more making it well suited for long-term archival storage applications. Tape capacity and data rates have soared reaching reached 18 TBs native capacity (45 TB compressed) on the latest LTO-9 drives delivering faster data rates than any HDDs. Robotic tape library capacities have surpassed one exabyte ($1x10^{18}$) becoming the first exascale storage system.

HSDCs are also realizing the benefits of tiered storage by integrating high-performance SSDs, HDD arrays and automated tape libraries to cost-effectively balance data allocation. Even though HSDCs constantly battle the exploding growth of disk farms which are devouring IT budgets and overcrowding data centers, many continue to maintain expensive HDDs often over 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 and the larger the data center the greater this inefficiency becomes. LTO's easy scalability, lower price and the lowest TCO play a greater role as the size of the storage environment increases. Several studies indicate the HDD TCO ranges from 5-7x higher than tape as the greatest benefits of tiered storage are achieved when tape is used. *For the HSDCs, adding HDDs is tactical – adding tape is strategic.*

Fighting the cybercrime epidemic is a major problem for all modern data centers and HSDCs are no exception. Tape provides data security with WORM (Write-Once-Read-Many) and encryption capabilities providing a secure, immutable storage medium for compliance, legal and any valuable data. Tape as an "Air Gap" solution, has gained momentum providing an electronically disconnected copy of data that prevents cybercrime disasters from directly attacking data while stored on tape. HDD arrays remaining online and spinning 7x24 are the primary cybercrime target as they are always vulnerable to attack.

LTO Tape Highlights

Maintaining a hyperscale data center is no small feat. As HSDCs strive to reduce energy consumption and carbon emissions, moving inactive data to tape storage provides many compelling benefits. In addition to energy efficiency, the LTO program has been fueled by a decade of strong technological development and continues to play a major role in traditional backup, active archive and disaster recovery applications in addition to effectively addressing many new large-scale storage requirements. As a result, the benefits tape provides is steadily expanding in many new areas.

- tape has the lowest carbon footprint and the lowest TCO (by 1/5th to 1/7th) compared to HDDs
- tape is less expensive ($/TB) to acquire than HDDs
- the 10-year LTO roadmap for tape technology is well defined with few foreseeable limits
- tape easily scales capacity by adding more cartridges - HDD scales by adding more drives
- tape storage has a higher ROI and a much longer replacement timeframe than HDDs
- tape libraries have intelligent, faster, more efficient robot movement and exabyte capacities
- LTFS provides a standard open file system with media partitions for faster "HDD-like" access
- tape supports file, block and object storage formats, which are increasingly popular with large-scale archival and cloud deployments
- the tape air gap provides protection from direct cybercrime attack

Conclusion

Reaching hyperscale status won't just creep up on you - it will run over you. The ascent to hyperscale is reshaping the entire IT industry and may become the future of computing. These large data centers can use enough electricity to power a whole city. A data center in Virginia was the first to reach 1 gigawatts of capacity which could power 700,000 homes and this degree of impact is forcing data centers to build in areas where electricity is cheap and generated from multiple sustainable sources. With the amount of energy consumption being a critical consideration, HSDCs are becoming the epicenter for advanced data archiving strategies. As the HSDC grows, sustaining larger amounts of archival data on HDDs will quickly become prohibitive. The rich LTO technology improvements of the past 10 years indicate that tape will continue to be the most cost-effective storage solution for the unprecedented HSDC challenges ahead. For data centers currently managing petabytes or more of data, the best time to evaluate new storage architectures may already be in your rear-view mirror. For data centers planning for their future, *the ascension to hyperscale and beyond will soon make tape storage solutions mandatory for sheer economic survival.*

Sponsored by the LTO Consortium

Linear Tape Open (LTO) Ultrium is a high capacity tape storage solution developed and continually enhanced by Hewlett Packard Enterprise, IBM and Quantum and promoted by the LTO Program. Linear Tape-Open LTO, the LTO logo, Ultrium and the Ultrium logo are registered trademarks of Hewlett Packard Enterprise, IBM and Quantum in the US and other countries.

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.