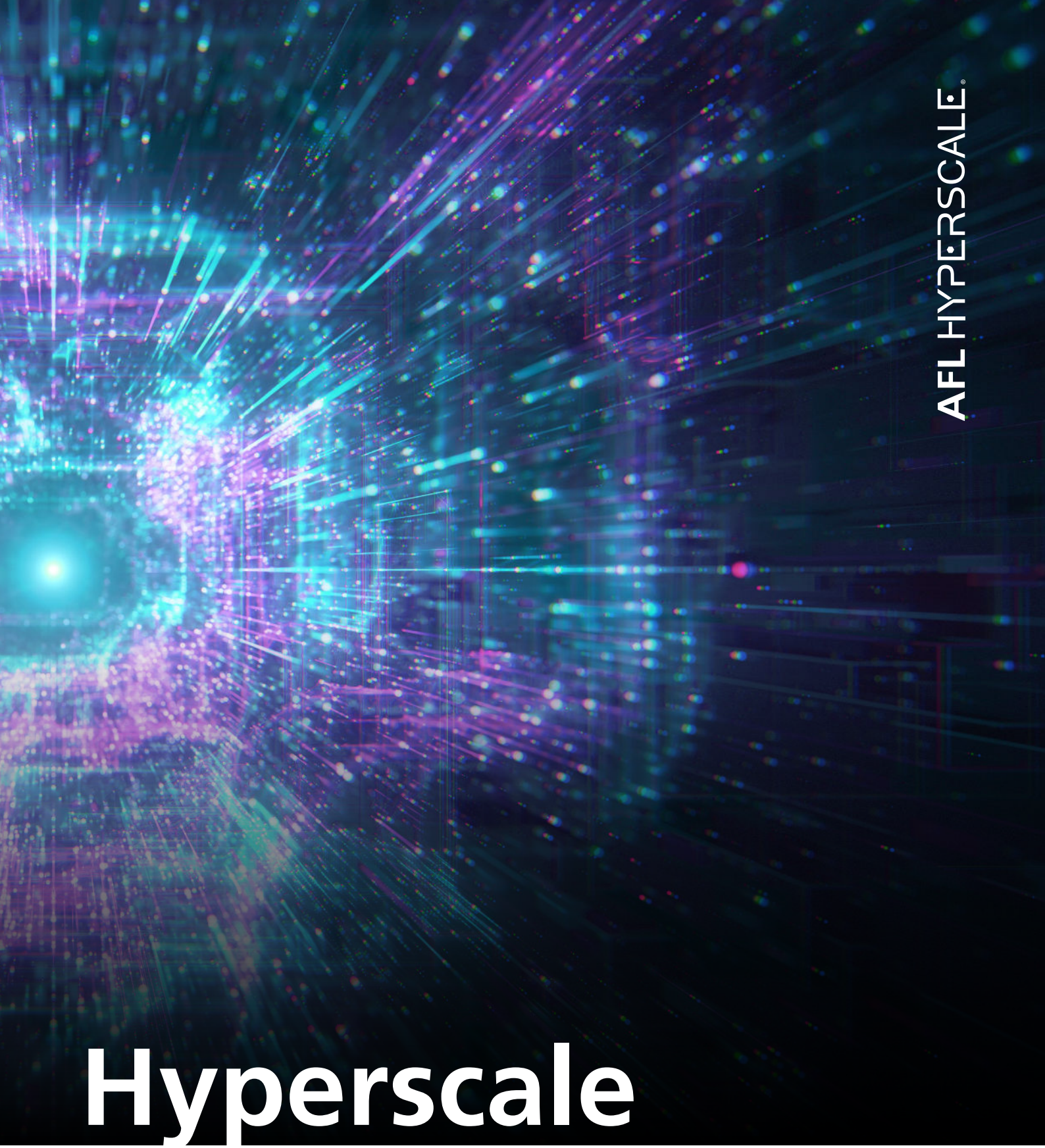


AFL HYPERSCALE®

Hyperscale Rising

Ebook

1 2 3 4

Contents

Introduction

Over the past five years, hyperscale computing has undergone a tremendous transformation that has firmly validated its position as the backbone of our digital world. In this second eBook in the Hyperscale Rising series, we provide readers with an in-depth look at the massive growth and evolution of the hyperscale industry during this relatively short but hugely impactful timeframe.

The pages ahead will highlight the key drivers for the explosive investments made by hyperscalers in infrastructure such as data centers, servers, and connectivity. We will also explore how major cultural shifts toward streaming media, social networking, and mobile apps have driven unprecedented demand on compute, and how COVID-19 has further accelerated adoption of cloud-based solutions relying on hyperscale providers.

Ultimately, we want to emphasize why understanding the hyperscale landscape is crucial for businesses and technology leaders looking to prepare for the future. The companies that harness hyperscale capabilities effectively today will have a competitive advantage in navigating the digital transformations ahead, across all industries.

Contributors

Alan Keizer
Senior Technology Advisor

Keith Sullivan
Director of Strategic Innovation

Hyperscale Growth and Cultural Impact

In just 5 years, hyperscale computing has transformed from an already large provider of cloud computing and media services into the bedrock of the global digital economy. The pace of infrastructure expansion in that short timeframe is staggering. The hyperscale data center construction boom shows no signs of slowing down either.

By the end of 2021, according to industry tracker Statista, the number of large hyperscale facilities swelled to 700, and with over 300 more data centers in the pipeline, according to Synergy Research, the total operational count will blast past 1,000 by 2024.[1]

To put this into perspective the total area covered by these data centers, if we assume the average size to be 100,000 square feet and size of 1000 data centres to 100 million square feet. That's about 5 times the size of Monaco.

Hyperscalers are the largest users of data center infrastructure, and their spending has a significant impact on the overall data center market. They doubled their annual hardware investments from 2018 from $157 Billion in 2018 to $295 Billion in 2022.

Examples of Big Hyperscaler investments in 2023

Google Announced investing **$1.7 Billion** in expanding operations in Ohio, US alone in Augst 2023.[2]

AWS unveiled plans this year to invest **$35 Billion** in building new data center infrastructure in Virginia to support cloud offerings.[3]

Google announced that it would invest about **$9.5 billion** on data centers and offices in the US in 2023. This was an increase from the previous year's investment of $7 billion.[4]

This explosive growth has coincided with tectonic shifts in the way we live our digital lives. Back in 2018, going down a YouTube or TikTok rabbit hole wasn't even possible. Neither was binge watching the latest Netflix hit series. The apps and sites we now rely on didn't have the capacity or ability to do the things we now take for granted.

That all changed thanks to hyperscale providers recognizing our shift, for example to voracious video consumption and building out their global data center networks accordingly. Video now accounts for nearly 80% of downstream internet traffic, compared to 57% back in 2018.[5]

The COVID-19 pandemic supercharged our embrace of streaming entertainment and social media when lockdowns hit. With nowhere to go, we flooded hyperscale data centers with Zoom calls, Netflix movie nights, and Facebook scrolls.

1. Synergy Research Group's Hyperscale Cloud Market Tracker report
2. Google announces $1.7 Billion investment in 3 Central Ohio data centers
3. Amazon's AWS to invest $35 billion in Virginia
4. Google to invest $9.5 billion in U.S. offices, data centers this year
5. Cisco's Visual Networking Index (VNI), a yearly report forecasting global internet traffic

The Pandemic Impact by the Numbers

Zoom

Zoom's daily meeting participants increased from **10 million** in December 2019 to 300 million in June 2020[6]

Netflix

In the first half of 2020, Netflix added more than **26 million** global subscribers[7]

Social Media

There was a recorded total increase of **300 million** new social media users between April 2019 and April 2020[8]

By the end of 2023, the global datasphere is projected to be 100 ZB, a notable leap from 5 ZB in 2018 and expected to swell further to 175 ZB by 2025. This surge is propelled by cloud computing's rise, expanding mobile device usage, and escalating data generation by businesses and individuals. Remarkably, 80% of this will be unstructured data—like text, images, and videos—growing faster than its structured counterpart.[9]

In essence, over the past five years, hyperscale computing has evolved from a nice-to-have to a must-have. It's now the lifeblood enabling our digital world and economy. The companies providing that connective tissue have invested billions to create a world where we can access apps and services anytime, anywhere.

6. Zoom User Statistics
7. The unassailable rise of Netflix | ACCA Global
8. Hootsuite Digital 2020 Report
9. Forbes – 175 Zettabytes by 2025

The Typical Hyperscale Data Center

Hyperscale data centers are marvels of cutting-edge engineering, containing acres of servers and networking gear powering our connected world.

Walking into one of these data center complexes is an awe-inspiring experience. The scale alone is staggering. For example, a large hyperscale data center can span up to a million square feet - equivalent to 18 football fields. More than a hundred thousand servers fill the facility, lined up in precisely arranged rows and columns to optimize airflow for cooling. The collective computing power is immense, capable of executing over 200 PetaFLOPs per second.

Fiber optic cabling forms the nervous system underlying hyperscale data centers. Often, hundreds of miles of cable are woven throughout the building floors, interconnecting servers to three or more layers of high-performance switches. This internal high-speed network enables huge data loads to move seamlessly between servers and out to the internet.

The engineering feat of constructing such a complex continues beyond the building walls. Underground fiber optic networks link data centers across regions to share loads. Above ground, a maze of transmission lines delivers massive power to keep servers running 24/7.

In total, enabling technologies like fiber optics, power management, and cooling systems combine to create the ideal environment for hyperscale data centers to deliver web-scale computing power. The scale of these facilities will only continue growing, mirroring our demand for cloud services and connectivity at the touch of a button.

Just as an example, look inside an Amazon Data Center.The biggest of hyperscale data centers may be a million square feet, and these will have over a hundred thousand servers.

Inside Amazon's US-East-1 Region

"Eight million square feet total in more than 70 buildings." This is the size of 140 American football fields.

Power capacity estimated at 870 MW, enough to meet the average electricity needs of 675,000 U.S. residents.

"Contains about 1 million meters of optical fiber." - equivalent to the distance between Miami and Seattle

In total, the scale of hyperscale data centers will continue growing to meet our insatiable demand for cloud services and connectivity. The engineering capabilities make this immense computing power seem unlimited.

10. Inside Amazon's US-East-1 Region

Hyperscale Spending In 2022

In 2022, hyperscalers channelled an impressive $295 Billion into data center infrastructure, nearly double the $157 spent in 2018:

Servers

A substantial **$187 Billion**, making up **63.2%** of the total investment.

Storage

$52 billion, accounting for **17.8%.**

Networking

$32 Billion, or **10.9%** of the total.

Software

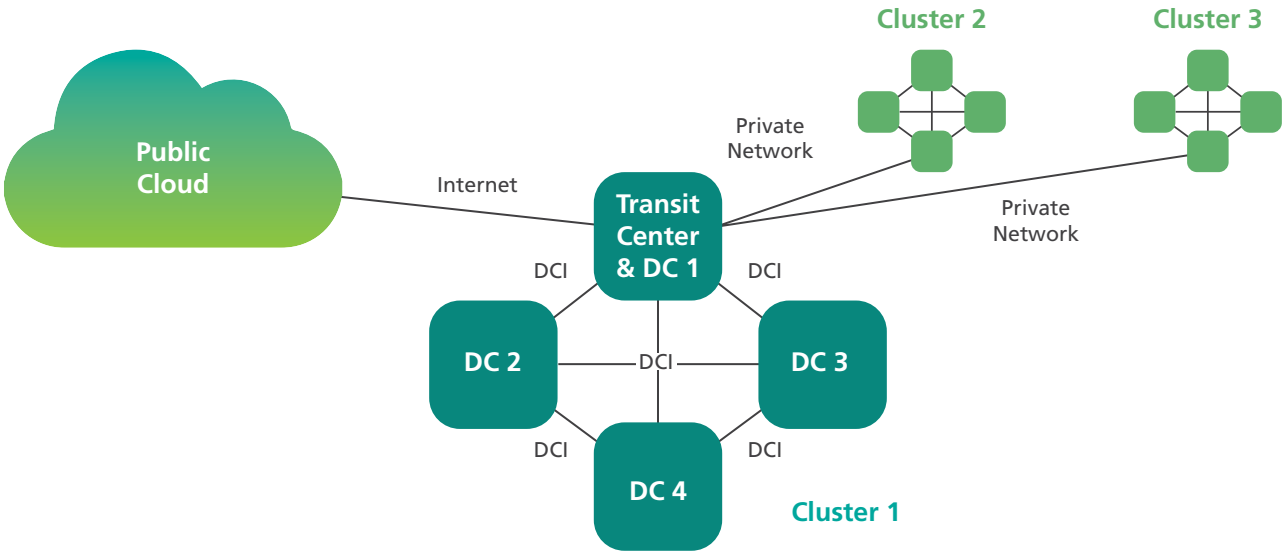
$20 Billion, which is **6.8%**.

Services

$4 billion, representing **1.4%**.

11. Synergy Research Group's Hyperscale Cloud Market Tracker report

Hyperscale as Interconnected Global Infrastructure



Hyperscale computing constitutes much more than discrete data centers – it forms an interconnected global infrastructure network enabling on-demand access to massive capacity. This distributed web-scale architecture allows users to tap into vast computing resources and services from anywhere in the world, abstracted from geographic limitations. The illusion of unlimited capacity is powered behind the scenes by hyperscalers interlinking data centers across regions through dense high-speed fiber optic networks. Advanced data center interconnect (DCI) technology plays a key role in enabling these hyperscale networks. DCI provides the high-bandwidth, low-latency connections between geographically dispersed data centers. This allows workloads and data to be efficiently shared across locations.

For example, a company in Singapore can leverage the full suite of services available on Amazon Web Services. However, the supporting infrastructure is not restricted to data centers only in Southeast Asia. AWS interconnects facilities worldwide so capacity can be efficiently allocated on demand based on varying customer needs. The same applies for an enterprise relying on Microsoft Azure services from an office in Dubai. The regional Azure cloud data center provides a user gateway, which in turn taps into Microsoft's 200 data centers across 60 regions and containing more than 4 million servers.

Enabling this interconnected hyperscale model requires extremely fast and reliable fiber optic data pipelines. Hyperscalers have aggressively built out their high-capacity fiber footprints between data centers across continents. Some have even invested in subsea cables to bolster international connectivity.

The result is a growing web of ultra-fast fiber links between hyperscale facilities worldwide. This allows them to operate seamlessly as a unified global infrastructure network. Companies can now reap the benefits of cloud elasticity and geo-distribution without having to architect their own distributed IT frameworks.

Hyperscale operators have established extensive networks of POP (Point of Presence) and Edge locations to serve as on-ramps. The on-ramps may be internal to co-location data centers or stand-alone edge sites. External users connect locally and the hyperscaler manages communication to designated regional or global cloud computing sites, providing assured bandwidth, security and quality of service.

AI and Hyperscale

The growth of Artificial Intelligence (AI) has been a key driver and beneficiary of hyperscale expansion over the past five years. As AI adoption explodes, access to immense data sets is crucial for training complex models. In response, hyperscale providers have rapidly built out data centers specifically equipped for AI workloads. This symbiotic relationship has fuelled the upward trajectory of both sectors. AI algorithms like deep learning neural networks require processing mind-boggling amounts of data during training. Hyperscale GPU-powered infrastructure has proven ideal for this computationally intensive task. The virtually unlimited capacity available on demand from cloud providers allows companies to develop AI models cost-effectively without massive upfront server investments.

In turn, hyperscale operators continue optimizing their infrastructure for even more powerful AI capabilities. For example, Google has developed proprietary AI chips called Tensor Processing Units tailored to accelerate deep learning applications. Amazon Web Services offers half a dozen machine learning optimized instance types like P5 (H100) and Trn1 (AWS Trainium) and two inference optimized instance types, Inf1 and Inf2. The hyperscale infrastructure supporting these AI leaps depends fundamentally on the fiber optic backbone connecting vast networks of data centers and servers. High-speed, low-latency fiber connectivity allows huge datasets to flow seamlessly during model development and training. As AI processing requirements continue growing exponentially, high-bandwidth fiber will be critical for unleashing the full potential.

As AI proliferates across industries and use cases, demand for hyperscale-powered AI development environments will intensify. From smart assistants to self-driving cars, emerging AI applications depend on hyperscale computing as their engine. This partnership is a defining technology trend of the 21st century. This symbiotic relationship creates a positive feedback loop where progress in one area fuels advancement of the other. The capabilities unlocked at the intersection of AI and hyperscale will shape the next wave of business and society-changing innovations.

Below are just a few examples of that trend:

Meta spends $7.1bn on data centers, servers, and office capex in Q1, still assessing “data center restructuring initiative”[12]

Microsoft Makes Multibillion-Dollar Investment in OpenAI[13]

12. Data Center Dynamics DCD
13. Data Center Knowledge

Hyperscale in 2024 and Beyond

Based on the tremendous growth hyperscale has experienced in a few short years, the pace of expansion shows no signs of slowing down. Analysts predict global hyperscale data center capacity will reach over 4,000 MW and 1000+ Hyperscale Facilities by 2024, more than double the capacity in 2019.

Several trends will shape the ongoing evolution of hyperscale computing over the next five years and beyond:

01

The AI Tsunami will grow even stronger.

Hyperscalers, other cloud and co-location operators will scramble to secure an AI position. AI will draw large venture capital funding and command a large portion of hyperscale capital budgets.

02

Continued Growth in Core Hubs

Key metro areas will remain central hubs, with Northern Virginia, Tokyo, London, Singapore, Silicon Valley, Chicago, Frankfurt, and Beijing seeing major sustained investment.

03

Emergence of New Markets

Previously underserved markets like Africa, South America, and Eastern Europe will see more hyperscale buildout to provide broader access.

04

Edge Computing Adoption

Placing micro data centers closer to end users will reduce latency and congestion for content delivery networks, IoT, and real-time services.

05

Innovation Spawning New Demand

Emerging technologies like the metaverse, autonomous vehicles, IoT networks, and quantum computing will necessitate exponential capacity growth.

06

Sustainability Initiatives

Energy efficiency, water conservation and renewable power will become bigger priorities amid environmental criticisms.

There will be greater reliance on high-speed fiber links between data centers to enable seamless geo-distribution. The future direction will clearly be continued expansion, but with greater attention to geographic reach, minimizing latency, optimizing efficiency, and boosting sustainability. By staying at the forefront of these trends, hyperscale providers can retain their dominance in powering tomorrow's innovations.

A Key Partner for the Future of Hyperscale Success

The rise of hyperscale providers like Amazon, Microsoft and Google is underpinned by vast fiber optic networks. This critical infrastructure enables their continued growth and innovation through:

- Tremendous data flows between hyperscale data centers riding on underlying fiber backbones. This provides the capacity and speed they require.
- Investment in metro and long-haul dark fiber that delivers resilience through route diversity and ability to scale up.
- Fiber as the ideal medium for data transit given its superior bandwidth, reliability, and security.
- Control over fiber networks providing the infrastructure to seamlessly meet future demands.
- Flexibility to integrate new technologies and workloads enabled by fiber network agility.
- Enabling the dramatic connectivity demands inside the AI optimized data center. AI computing clusters need 5 to 10 times as many connections as conventional compute and storage facilities.

Fiber is the indispensable strategic foundation enabling hyperscalers to thrive now and dominate as demands escalate. Comprehensive fiber assets future-proof their ascent.

As hyperscale computing continues rapidly evolving in the coming years, the infrastructure underpinning these global networks will be more crucial than ever. Companies specializing in data center connectivity solutions like AFL Hyperscale will play an indispensable role in this future.

With its extensive expertise across fiber optic cabling, network design, and creating tailored solutions for individual data centers, AFL Hyperscale has established itself as an industry leader.

Its involvement in developing standards and technologies further cements its position at the forefront of advancing the hyperscale landscape.

Looking ahead, AFL Hyperscale's global footprint, local support, and focus on innovation positions them as an ideal partner in tackling the opportunities and challenges of tomorrow's hyperscale requirements. As this e-book has explored, interconnectivity is imperative to realize the full benefits of distributed hyperscale architecture. AFL Hyperscale's specialty in high-performance data center connectivity solutions directly addresses that need.

AFL HYPERSCALE®

Data Center Cabling and Connectivity Experts

AFL Hyperscale creates purpose-built fiber optic connectivity solutions for data centers.

As the first cabling and connectivity provider focused purely on hyperscale, colocation, and enterprise data centers, we intimately understand the unique infrastructure, performance, and scaling challenges facing these facilities.

With deep expertise in data center interconnection, white space, and emerging technologies like AI, we pioneer innovative fiber optic solutions engineered specifically for the demands of high-density data centers.

Leveraging 40+ years of optical networking expertise, our global team provides purpose-built connectivity solutions that enable rapid deployment, maximum uptime, and reduced costs for data centers worldwide. Backed by decades of experience, we offer invaluable guidance on optimizing network design to deliver future-proof solutions anywhere they are needed.

AFL Hyperscale. The World, Connected.

www.aflhyperscale.com