

The Secret to Effective Hyperscale Capacity Planning

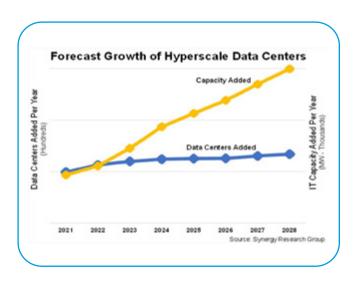
By Todd Schneider Sr Director-Product Management & Hyperscale Business Development



Published: May 2024 chatsworth.com

The social, e-commerce, loT, and AI revolutions are here, demanding a three-fold increase in hyperscale data center capacity over the next four years. But this exponential growth comes with a hidden enemy: crippling complexity in capacity planning.

Hyperscale data centers are constantly evolving to keep up with the ever-changing demands of the digital world while meeting increasingly stringent sustainability initiatives. New equipment, workloads, and use cases—it's all part of the hyperscale landscape. And forget one size fits all—in the hyperscale environment, customization is key. That's where the right suppliers come in. They are the secret to effective capacity planning, scaling hyperscale resources up or down in a flash, maximizing investments, and keeping data centers running at peak efficiency.



A Balancing Act of Power, Efficiency, Uptime, and Agility

Hyperscalers face an unrelenting, insatiable appetite for new technologies that require more data, storage, and processing power. Generative AI models are particularly complex and push the limits. These deep-learning training and inference models require the power of graphical processing units (GPUs). Unlike typical CPU-based servers, GPUs excel at pulling vast amounts of data in parallel from various sources and processing thousands of high-throughput computations simultaneously. But that comes with a catch – up to 10 times more power consumption! While typical servers might need rack power densities of 10 to 15kW, that number jumps to three or even four times as much with GPUs. Scale this to an AI cluster with thousands of interconnected GPUs, and the challenge becomes clear — exponentially higher power requirements, space constraints, and the need for advanced cooling systems. It's a whole new level of data center infrastructure.



Further complicating capacity planning is the need for efficiency and sustainability. As data volumes and process power continue to explode, hyperscalers must expand capacity while simultaneously identifying ways to curtail energy consumption and achieve aggressive sustainability goals. Leading players like Google, Microsoft, Amazon, and Meta have pledged net-zero operations by 2050 or sooner, and regulatory pressure is mounting. The new European Union's (EU) Energy Efficiency Directive (EED) requires data centers with a demand of 100 kilowatts (KW) or more to report their energy performance. Similar initiatives are underway in the U.S., demanding transparency on energy usage, carbon emissions, and other resources for data centers.

While hyperscalers strive to support higher power densities and sustainability, they also remain beholden to unwavering uptime and reliability. They must meticulously plan for capacity expansion while preventing service interruptions. Predictive analytics capabilities are crucial to avoiding capacity deficiencies and proactively addressing issues before they escalate into outages. Maintaining robust security to mitigate cyberattacks and human error remains paramount, requiring compliance with evolving security standards.



As the requirements become more intricate, many standard racks and off-the-shelf components are proving inadequate for hyperscalers' unique needs. New high-density computing applications often demand customized rack solutions that optimize space utilization, improve airflow, support advanced cooling integration, and enhance security and management. While it may seem counterintuitive, customized rack solutions actually speed deployment and improve efficiency by reducing assembly time and labor and lowering transportation costs and waste. However, customization can impede capacity planning from a logistics perspective. Vendors that can't innovate and deliver purpose-built solutions fast enough can delay project timelines. Accelerating equipment capabilities, rapidly evolving workloads, and new use cases exacerbate the challenge, creating dynamic conditions and unforeseen demands that often necessitate last-minute adjustments.

Overcoming these capacity planning challenges to keep fast-paced deployments on track requires a new breed of hyperscale supplier—one with R&D expertise to rapidly design turnkey solutions purpose-built to meet specific needs, manufacturing agility to quickly scale production, and strategic planning and logistics capabilities to adapt to fluctuating forecasts and ensure on-time delivery.

Taming the Beast from Rack to Rollout

The future of Al and other power-hungry applications hinges primarily on what happens at the most fundamental level—the rack. As the building blocks of a hyperscale data center, racks must be optimized for efficiency, reliability, and blazing-fast deployment.

When it comes to effectively expanding capacity, hyperscalers need to make sure their racks provide enough load capacity and horizontal and vertical space to house current and future equipment, cabling, and power needs. They also need the ability to easily integrate racks with top-notch thermal management and advanced cooling technologies that can effectively remove heat in higher rack power densities.

Effective capacity planning also means keeping a close eye on the amount of power consumed by rack equipment compared to total power availability. This information can help identify available rack capacity for future needs, prevent overloads, and track energy consumption and power usage effectiveness (PUE) in line with sustainability initiatives. Monitoring rack-level factors like temperature, humidity, airflow, and leaks can also help assess the effectiveness of cooling systems and quickly detect and address hazards that impact equipment reliability and lead to downtime.





Effective capacity planning goes beyond just what's in the rack – it's also about how quickly those racks get up and running. That means working with suppliers with design and manufacturing capabilities to customize and produce solutions to spec. They should be able to pre-install any required power, cable, and thermal management accessories. Racks should also facilitate fast equipment loading and roll easily into position to meet aggressive project schedules and reduce costs. That means choosing racks with ample mounting and positioning options and those that can perform in a fully loaded configuration with the right packaging, stabilizer, casters, and shock pallet options for quickly and safely rolling into position in hyperscalers' core facilities and their remote edge and colocation data centers.

Purpose-Built Hyperscale Solutions



Hyperscalers can ensure effective capacity at the rack level with CPI's data center solutions backed by direct, one-on-one consultation services, superior R&D, and innovative engineering capabilities. From open racks, cabinets, and cable management to aisle containment, thermal management, power monitoring and security, CPI quickly customizes and delivers turnkey solutions that match specific hyperscale needs.

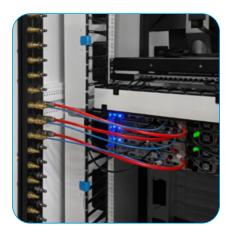
- ZetaFrame® Cabinet System The ideal foundation for hyperscalers, the highly engineered ZetaFrame® Cabinet delivers fast, built-to-order configurability, strength, and scalability. It offers the industry's highest 4,000 lb. dynamic load rating and enhanced high-density cable management while easily integrating power, cable, and thermal management to support next-generation compute. An optional shock pallet is also available to expedite safe delivery of racked equipment.
- Rack Solutions Whether hyperscalers need 19-inch EIA standardized frames or solutions that adhere to the latest Open Rack requirements of the Open Compute Project® (OCP), CPI rack solutions are built to order with industry-leading strength, stability, and durability to meet the demands of mass scale computing.

• Custom Aisle Containment — Backed by CPI's extensive thermal expertise, durable, maintenance-free hot and cold aisle containment solutions for cooling efficiency and economization are built to spec as either free-standing structures or integrated with ZetaFrame cabinets to meet exact hyperscale specifications for rack power densities up to 32kW.

The Elevate™ Adjustable Containment Solution is optimized for the company's ZetaFrame Cabinet System and designed to streamline deployment. Its telescoping design seamlessly adapts to create a duct connecting the ZetaFrame cabinet to the overhead drop ceiling, allowing for quick adjustments to the containment. The Elevate Adjustable Containment Solution is engineered to align with standard ZetaFrame cabinet widths and supports both hot aisle containment and row-level vertical exhaust duct applications. Pre-assembled panels eliminate the need for on-site fabrication of metals or panels and drilling. A sliding service panel facilitates easy access to the cabling infrastructure and other overhead structures, allowing for adjustments and repairs with minimal disruption.

- Thermal Management From snap-in filler panels, air dams and directors to sealing, fans, brush covers, grommets, and vertical exhaust ducts, CPI can outfit cabinets with a variety of thermal management solutions that optimize cooling by preventing bypass airflow, isolating conditioned and exhaust air, and moving hot exhaust air away from equipment.
- Liquid Cooling The ZetaFrame cabinet can be fully customized with all the provisions to support advanced liquid cooling solutions for higher rack power densities, including rear door heat exchangers and the ZutaCore® waterless direct-to-chip liquid cooling solution. ZetaFrame's high load rating and mounting features enable seamless integration with ZutaCore manifolds and heat rejection units, enabling secure transport when used with a shock pallet.
- Advanced Power Management Intelligent high-density eConnect® PDUs meet today's hyperscale higher power
 loads while delivering granular cabinet- and equipment-level power monitoring and management with a web-based
 interface for remote access, automated monitoring, outlet control, and easy configuration. Powerful DCIM can
 analyze information from PDUs in real-time, allowing hyperscales to identify space and power capacity across their
 facilities and monitor power usage across the entire power chain.
- **Environmental Monitoring** Integrated PDU sensors provide remote active monitoring of temperature, humidity, airflow levels, and leaks to quickly detect and prevent hot spots, moisture build-up, and other hazards.
- Security and Protection With remote programming and monitoring through eConnect PDUs, ZetaFrame cabinets
 can be equipped with networked electronic lock solutions to provide 24/7 monitoring of front and rear door access,
 including audit logs to enable compliance with security standards like SOC 2, HIPAA, PCI DSS 3.2.1, SSAE 18, and
 ISO 27001.







It Takes More than a Vendor

While hyperscale data centers can solve rack-level capacity planning challenges with customized solutions that meet their unique needs, planning from a logistics perspective is another challenge—especially in today's unpredictable supply chain plagued by natural disasters, economic uncertainty, international conflict, and the lingering effects of the pandemic. Traditional vendors struggle to keep pace with the demand and customized requirements of hyperscale customers. Research reports that data center construction reached an all-time high in 2023, but extended lead times for key infrastructure components delayed many project delivery timelines.²



Ongoing changes and complexities in equipment, workloads, and use cases combined with a volatile market often require hyperscales to make decisions

relatively close to deployment, resulting in last-minute refinements to solutions and demand forecasting. This takes a vendor with R&D, manufacturing, and logistic capabilities to scale up or down quickly. In addition, hyperscales frequently demand price discounts on large purchases, which can create cost pressures for some vendors.

While capacity expansion is expected to occur in existing facilities, hyperscalers still need to build out additional capacity in various locations—from newly-built facilities and remote edge data centers to wholesale colocations and other shared spaces. That means vendors must also have global capabilities and the ability to meet local requirements or regulations.

As leaders in technology, it's no surprise that hyperscalers will quickly turn away from vendors who can't meet their pace and requirements. Even when they can deliver, hyperscalers prefer to work with suppliers who go beyond simply fulfilling purchase orders on time—they want strategic partners with direct account support who they can trust to act as an extension of their own team and provide long-term value.



As a hyperscale veteran fluent in the language of volatility, CPI develops multidimensional strategic relationships with hyperscalers and offers a tactical go-to-market approach that delves deep to support fast, seamless deployments tailored to unique specifications. With superior customer service and the following R&D, engineering, manufacturing, and logistics capabilities, CPI keeps pace with hyperscale demands to help eliminate capacity planning challenges:

- Superior R&D CPI has the R&D capabilities to rapidly engineer design-to-value solutions that meet specific hyperscale use cases and compute power, bandwidth, and storage requirements. Our 10,200 square-foot state-of-the-art R&D and Training Center allows hyperscalers to interact with solutions in one environment and observe real-world thermal and product testing.
- Customization business model CPI has a rich history of delivering purpose-built turnkey solutions; it's in our DNA. With CPI's business model that involves comprehensive collaboration with our customers, a modular platform-based design strategy to eliminate constraints, and adaptable production to accommodate built-to-order solutions, hyperscalers don't have to sacrifice speed and competitive pricing for customization.
- Current capacity analysis CPI thoroughly analyzes current cabinet and
 rack capacity, monitors current power usage against available power, and
 assesses the effectiveness of cooling systems to help hyperscalers identify
 specific infrastructure requirements for expansion and reconfigurations
 within existing spaces.
- Expert market awareness As a trusted advisor for hyperscalers, CPI consistently stays on top of the latest technology advancements, unique needs of various market segments, geographical demand, and regulations to provide scalable, future-proofed solutions that ensure compliance.
- Data-driven insight CPI consistently analyzes sales and spending trends and leverages advanced analytics for data-driven forecasting to anticipate capacity needs and adjust production schedules to align with budget cycles and project timelines, allowing hyperscalers to enjoy fast, seamless deployments.
- Agile global manufacturing CPI's global lean manufacturing, product
 availability, and logistics capabilities, combined with rapid product
 customization and data-driven insights, means we can quickly and reliably
 scale up or down and ensure fast lead times anywhere in the world,
 eliminating supply chain headaches and providing hyperscales with a
 competitive advantage.
- One-on-one support With direct account managers who know the
 volatile hyperscale landscape and engage with technical product and data
 center teams across all functions, CPI provides hyperscalers with more
 than a vendor. They get a true strategic partner with unequaled application
 expertise who understands their specific needs and requirements and is
 dedicated to ensuring long-term value.







Ready to get started? Connect with CPI today and achieve effective capacity planning through a rich relationship that ensures the right purpose-built rack, containment, and power distribution solutions delivered at speed—making hyperscale a reality, not a headache.

References

- ¹ Hyperscale Data Center Capacity to Almost Triple in Next Six Years, Synergy Research Group, October 2023
- ² North America Data Center Trends 2023, CBRE

Contributors



Todd Schneider Director of Product Development, Chatsworth Products

Todd Schneider is the Director of Product Management and Hyperscale Business Development at Chatsworth Products, Inc. Todd has over 25 years experience in the data center space covering a wide variety of functional areas including Product Management, Business Development, Sales Engineering and Product Development. His current roles focus on driving standard new product innovation to CPI's core offerings and delivering customer-centric, customized solutions for Hyperscale Operators. Schneider previously worked at Legrand, Electrorack and Wright Line.

About Chatsworth Products

From the data center to the edge, Chatsworth Products (CPI) is more than just a global manufacturer of products and solutions that protect your ever-growing investment in information and communications technology for IT and industrial automation applications. We're a 100% employee-owned, U.S. based organization that's building a better digital future—one rack, cabinet and PDU at a time.

As a trusted business partner, we are uniquely prepared to respond to your requirements with global availability and rapid product customization, giving you a competitive advantage. With more than 30 years of experience engineering thermal, power and cable management solutions for the data center, enterprise networking and industrial enclosure markets, CPI is well positioned to provide you with unequaled application expertise, customer service and technical support, and a global network of industry-leading distributors.

chatsworth.com techsupport@chatsworth.com 800-834-49699

