WHITE
PAPER

SPECIFYING THE CORRECT
CABINET FOR AI COMPUTE CLUSTERS:

CONSIDERATIONS FOR DATA CENTER ENGINEERS

#LegrandImprovingLives

legrand®

THE CHALLENGE OF SUPPORTING AI COMPUTE CLUSTERS

The exponential rise of artificial intelligence has revolutionized data center infrastructure. Unlike conventional server environments, AI workloads demand significantly more power, generate intense heat, and require highly specialized networking and power distribution. The shift is not just about scaling up—it's about rethinking how data centers are physically structured to support the weight, density, and cooling requirements of AI hardware.

At the heart of this transformation is a component often taken for granted: the server rack and cabinet. Once considered little more than a metal enclosure, the modern cabinet now plays a critical role in organizing AI deployments. Data center engineers must carefully evaluate specifications to ensure that their cabinets are not only sufficient for today's AI workloads but also capable of supporting the next generation of machine learning models and high-performance computing (HPC) applications. The wrong choice can lead to inefficient cooling, increased operational costs, and long-term scalability issues.

LOAD CAPACITY: THE WEIGHT OF AI INNOVATION

AI servers are built to handle unprecedented computational tasks, and with that power comes a significant increase in physical weight. Traditional data center cabinets are designed for standard 1U and 2U servers, each typically weighing around 17 pounds per rack unit (RU). However, AI-optimized graphics or tensor processing units (GPUs and TPUs) significantly exceed this weight. An NVIDIA H100-based AI server can weigh up to 36 pounds per RU, meaning a fully loaded AI rack may need to support 5,000 pounds of static weight.

Failure to account for this weight can lead to structural issues, equipment damage, and even safety hazards. AI-ready racks are engineered with reinforced steel frames and high-dynamic load ratings, ensuring that they can handle the stress of pre-populated shipping and on-site mobility. When specifying racks for AI, engineers should prioritize designs with robust tubular steel frames and fully welded construction to mitigate the risks associated with extreme density. The tubular steel frames are engineered for strength and rigidity without adding excess weight, distributing the load evenly to prevent flexing or warping under heavy equipment.

SIZE MATTERS: RETHINKING RACK DIMENSIONS FOR AI

In addition to significant increases in weight, the physical footprint of AI cabinets is also changing. Whereas traditional racks are typically 42U in height, 24 inches wide, and 42 inches deep, AI workloads necessitate larger enclosures to accommodate additional hardware and cooling components. AI racks now extend to 48U, 51U, or even 52U in height, with widths reaching 30 to 36 inches and depths expanding to 54 inches.

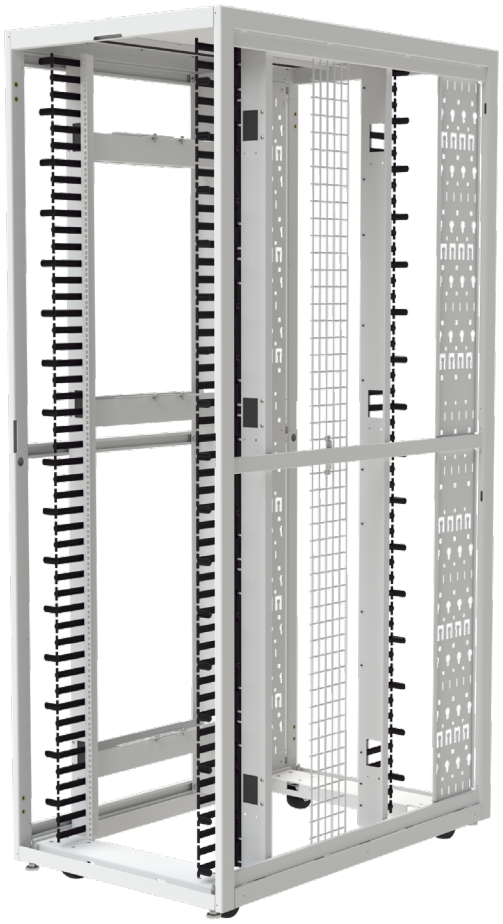
These expanded dimensions serve multiple purposes. Taller racks maximize vertical space, reducing the total number of cabinets needed within a facility. Wider racks allow for more effective cable management and power distribution, ensuring that high-bandwidth networking components are not restricted by tight enclosures. Deeper racks provide additional room for cooling infrastructure, facilitating better airflow and liquid cooling integration.

By adopting AI-ready rack dimensions, data center engineers can future-proof deployments, reducing the need for costly expansions and equipment overhauls in the coming years.

THE NEED FOR INTELLIGENT POWER AND HARDWARE INTEGRATION

As AI data centers scale, power distribution becomes an increasingly complex challenge. Unlike traditional racks, which may house one or two power distribution units (PDUs), AI racks often require three or four PDUs per cabinet to supply the necessary power for high-density configurations, which can demand 30 to 60 kW per cabinet or higher. Standardized power distribution is no longer enough; intelligent PDUs with real-time monitoring and predictive analytics are essential for optimizing power utilization and preventing inefficiencies.

Beyond power, AI racks must seamlessly integrate with a growing array of supporting infrastructure, including multiple types of environmental sensors, redundant networking components, and direct-to-chip liquid cooling solutions. These integrations are critical for both performance and operational efficiency. Engineers should specify racks that not only accommodate today's AI hardware, but also allow for the weight of additional high-density servers for future adaptations, ensuring that the infrastructure remains relevant as processing demands continue to evolve.



T-Series Cabinet - Fully Customizable with optional airflow, cable management accessories and doors, top, side, and bottom panels.

MANAGING THE HEAT: AIRFLOW AND COOLING INNOVATIONS

Heat is the natural result of high-performance AI computing, and as power density increases, so does the challenge of keeping AI hardware within the correct operating temperatures. AI racks can consume 100 kW per cabinet, more than ten times the power draw of a conventional server rack. Traditional air-cooling methods, while effective for standard or even IT loads, are often insufficient to handle the demands of AI compute clusters.

A modern AI-ready rack must integrate with the most advanced cooling strategies. Rear-door heat exchangers, liquid cooling manifolds, and hot/cold aisle containment systems have become essential components of AI infrastructure. Without proper cooling integration, AI hardware runs the risk of thermal throttling, reducing computational efficiency and increasing the likelihood of downtime. Engineers must ensure that rack designs are optimized for airflow, with features such as sealed cable entry points, blanking panels, and precision airflow containment to maximize cooling efficiency while minimizing wasted energy.

CABLE MANAGEMENT: AVOIDING BOTTLENECKS IN AI NETWORKING

One of the most overlooked challenges in AI data center design is the sheer volume of cabling required – often four to five times more than traditional deployments – to support high-bandwidth networking. Unlike traditional architectures, AI clusters rely on dense, high-speed inter-server connectivity—such as NVLink, InfiniBand, or 400G Ethernet—which creates a complex web of east-west traffic that must be carefully planned and managed.

Without proper management, cabling can obstruct airflow, increase serviceability issues, and create potential points of failure. AI racks should incorporate integrated vertical and horizontal cable management solutions, pre-cut routing pathways, and enhanced tie-down options to ensure clean, efficient cabling. By designing with cable management in mind, data center engineers can reduce network congestion and simplify future maintenance, ensuring uninterrupted AI processing.

ACCELERATING DEPLOYMENT WITH PRE-CONFIGURED AI RACKS

The speed at which AI is evolving demands a faster approach to infrastructure deployment. Traditionally, data center teams assemble racks on site, a process that can take days or even weeks. However, the introduction of pre-integrated AI rack solutions is streamlining infrastructure deployment and shortening traditional timelines.

While rack and stack deployments remain common—focusing primarily on loading and wiring IT equipment—pre-configured AI racks go further. The considerable volume of electrical and network wiring involved in AI deployments makes factory pre-configuration necessary to maintain quality and consistency at scale.

Finally, the weight of fully loaded AI racks requires that they be structurally engineered for transportation, ensuring stability and security during shipment and placement. To meet these demands, racks should be compliant with the ISTA 3B standard, which simulates the physical stresses of parcel delivery environments—including vibration, impact, and compression—to ensure their durability in real-world conditions.

By leveraging pre-configured AI racks, data center operators can accelerate time to market for AI applications, reducing labor costs and ensuring infrastructure readiness in record time.

THE FUTURE OF AI INFRASTRUCTURE: RETHINKING RACK DESIGN

As AI continues to advance, the underlying infrastructure must evolve in parallel. The challenges of weight capacity, power distribution, cooling efficiency, hardware integration, and deployment speed all point to the need for purpose-built AI racks that extend beyond the limitations of traditional enclosures.

For data center engineers and facility managers, specifying the right rack is no longer a secondary consideration—it is a critical factor in AI performance and long-term scalability. By adopting AI-optimized racks, organizations can ensure efficiency, reliability, and cost-effectiveness while preparing for the next wave of high-performance computing demands.

ABOUT LEGRAND CABINETS

Legrand is a leading manufacturer of cabinet and containment platforms designed to meet the needs of data center applications from edge and enterprise to hyperscale and colocation. Constructed in the U.S., Legrand cabinets are engineered to solve unique customer challenges. Regardless of the issue—whether it be in containment, power, copper, fiber, or pathways—Legrand has decades of experience delivering turnkey solutions. With both sheet metal and tubular steel construction platforms, Legrand's cabinets guarantee unmatched versatility and stability.

To learn more visit

legrand.us/cabinets

Legrand Cabinets & Containment
800.934.5432
legrand.us