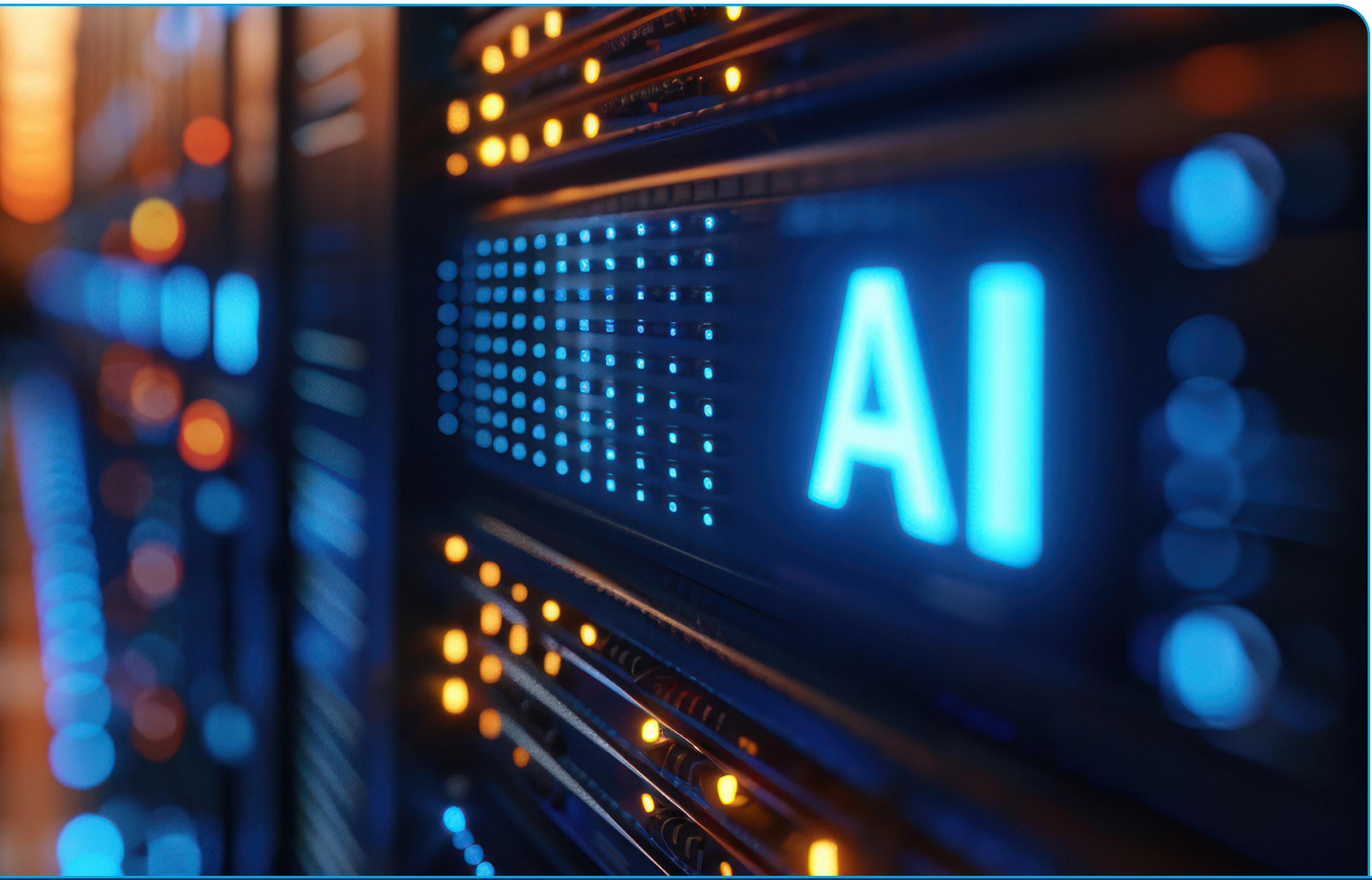
WHITE PAPER

CHATSWORTH PRODUCTS

Powering AI Infrastructure: Rethinking Rack-Level Strategy for High-Density Workloads

By Ashish Moondra
Senior Director of Electronics and Software



Published: October 2025

chatsworth.com

Introduction: The AI Power Challenge

No other single workload has reshaped data center design as abruptly as generative AI. GPU clusters for training and inference have already pushed cabinet power densities from 10–20 kW into levels exceeding 100 kW, with some deployments testing 120 kW or more. Once confined to supercomputers, these rack-level demands are now entering mainstream facilities at a scale never seen before.

These figures are not anomalies. According to Uptime Intelligence, AI hardware is doubling rack power demand with every generation. Systems rated at 300 kW per rack—once confined to supercomputers—are slated for mass deployment by 2026[1]. In other words, what feels extreme today is only the starting point. The crushing force of AI is densification paired with generational leaps in power demand, moving faster than most facility cycles can adapt.

What makes this transition different from past inflection points is its singular cause. Virtualization redistributed loads across clusters, cloud abstracted utilization into elastic pools. AI concentrates its disruptive load at the rack. Each new GPU generation drives more power into fewer units, compounding thermal output and pushing AI training workloads to consume six to eight times more energy than conventional compute. At these levels, the cabinet—not the facility—has become the defining unit of design, where power, cooling, and space planning now converge.

This leads to the central paradox: operators may have sufficient upstream megawatts yet still struggle to deliver them reliably to cabinets packed with GPUs. In the AI era, the true metric of readiness is not site bulk capacity, but cabinet deliverability—how many high-density racks a facility can reliably support.

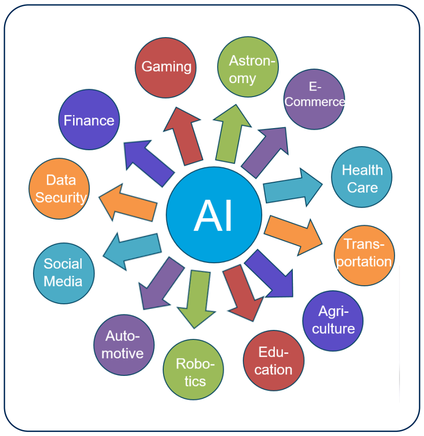
Meeting that standard exposes a new landscape of challenges for operators:

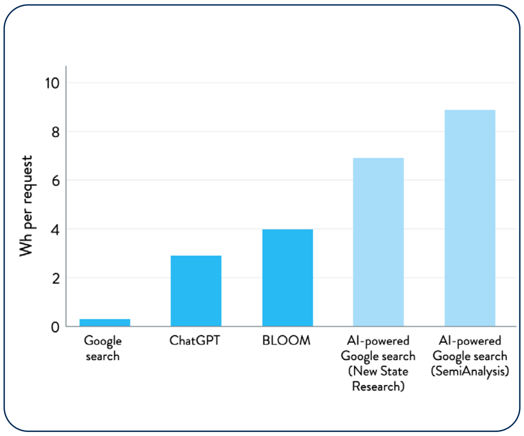
- **Uneven load distribution** within racks, creating localized hot spots and wasted capacity.
- **Physical space consumed by power distribution units (PDUs)**, reducing capacity for airflow corridors or liquid cooling infrastructure.
- **Circuit and breaker limitations,** with many facilities lacking the headroom to support extreme densities.
- **Reactive power management,** where failures are diagnosed after disruption rather than anticipated.
- **A mismatch between infrastructure cycles and hardware evolution,** leaving little room for future growth.

This white paper examines how operators must rethink rack-level power strategy for AI deployments. The focus is on the **white space cabinet as the new unit of design,** where distribution, cooling, and management converge – and how each element must evolve to support AI-class densities.

AI generated queries require 10x the power in comparison to a regular Google search

- **Machine Learning**
- **Natural Language Processing**
- **Graphics Rendering**
- **Robotics**
- **Computer Vision**
- **Algorithms**





From Gray Space to White Space: Where AI Power Pressure Builds

Historically, efficiency gains were pursued in the **gray space:** UPS systems, switchgear, and remote power panels. Operators refined redundancy models, right-sized UPS modules, and optimized distribution paths. The logic was straightforward—loads downstream were relatively predictable and modest. A cabinet drawing 5–10 kW was considered dense. Downstream efficiency was rarely a limiting factor.

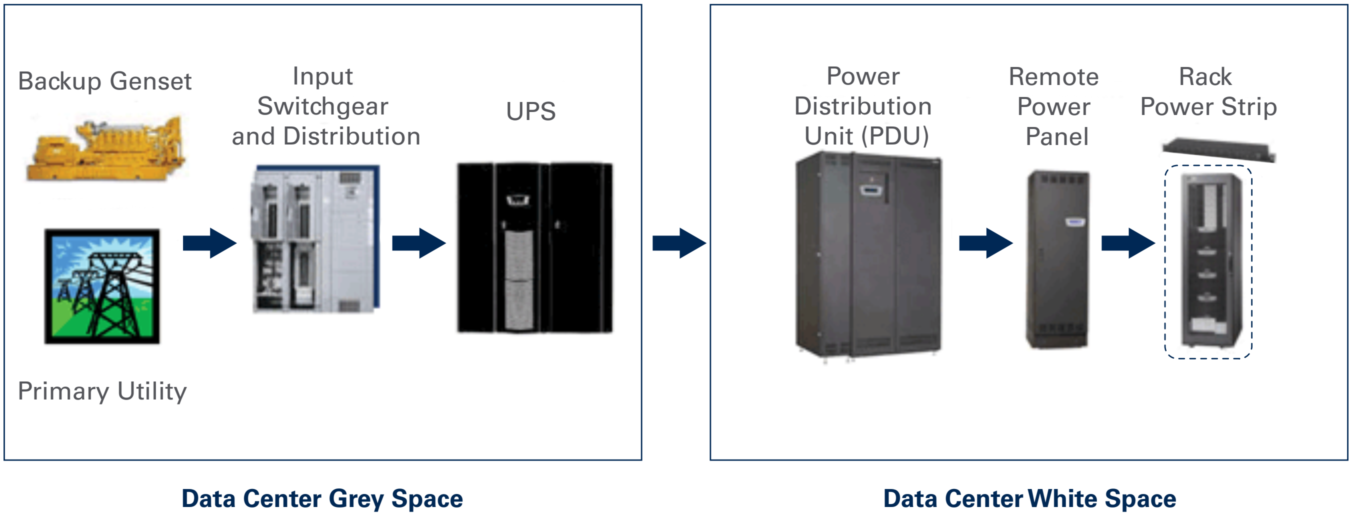
AI breaks that equation. Even if the facility has ample upstream power, delivering it safely and predictably to individual cabinets is a different challenge. As Uptime puts it, each row of future high-performance racks will have equivalent power demand to entire megawatt-scale data centers of the past[1].

For most of the past two decades, cabinets were considered the passive endpoint of the power chain. The prevailing mindset was as long as the upstream system delivered the expected kilowatts, the cabinet would manage. That assumption no longer holds.

At 50–120 kW per rack, small inefficiencies compound into failures. Uneven load distribution across devices can trigger hot spots, tripped breakers, or premature component aging. Even operators with ample upstream supply now see failures rooted in the cabinet itself.

This reframes the cabinet from endpoint to design locus. Power paths, breaker configurations, PDU placement, and airflow clearance now require the same rigor once reserved for UPS system design. Resilience is increasingly determined at the cabinet, not the utility interconnect.

Elements of Data Center Power Distribution



Three operator types illustrate the scale and shape of the challenge:

Hyperscale Operators: Risk of Overbuilding

Hyperscale data center providers can, in theory, afford to redesign around 80–120 kW cabinets. But the real risk is not undercapacity—it's stranded investment. Building electrical paths to the frontier of density may mean underutilized infrastructure if workloads plateau at 60–70 kW. Hyperscalers have the capital to absorb inefficiencies, but even at their scale, miscalculations carry financial and sustainability penalties. The challenge is not only building for peak density but ensuring that utilization matches investment.

Colocation Providers: Asymmetry as the Constraint

Colocation providers face a dual challenge. AI tenants want pods with 80 kW+ racks, but the surrounding hall may still run legacy enterprise loads at 5–15 kW. This asymmetry is not just about electrical balance—it creates operational blind spots. For example, breakers set to protect legacy racks may nuisance-trip under the transient loads of AI, and airflow containment strategies tuned for conventional cabinets can collapse under outliers. The colo risk is reputational: a single AI deployment that destabilizes a mixed environment can undermine tenant confidence across the hall.

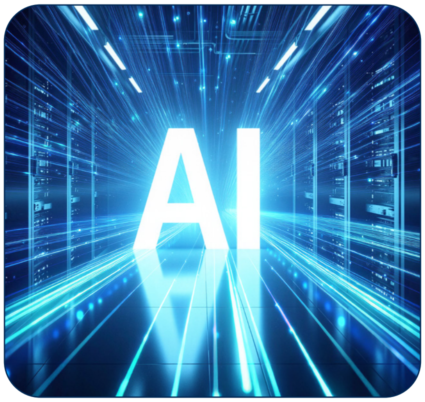
Enterprises: Compressed Timelines, Zero Cushion

Enterprises often face the steepest barriers. Most are not re-architecting entire facilities; they are carving out high-density pods inside structures never designed for them. The mismatch is temporal: AI refresh cycles are measured in months, while corporate real estate cycles are measured in years. This gap forces compromises—such as retrofitting within undersized rooms, leasing colo space, or deferring projects until infrastructure catches up. Each option carries a penalty: sunk cost, higher operating expense, or strategic delay.

Redefining Capacity

The lesson is not that gray space is irrelevant —it remains critical. Rather, white space has become the critical frontier. In an AI-driven world, a 50 MW facility that can only sustain 10% of racks above 50 kW is less competitive than a 40 MW facility that can sustain 40% of racks at those levels. In an AI era, capacity is not only about supply, but where and how that supply can be used without compromise.

Emerging Trends in Power Delivery for AI Data Centers



The last major rethink of electrical topologies in data centers came with the shift from single-phase to three-phase distribution in the 1990s and early 2000s. Since then, most change has been incremental. AI workloads break that pattern.

Operators are being forced to revisit foundational assumptions about how power is delivered, balanced, and protected. As a result, several key trends are emerging that are reshaping the way operators approach power delivery in AI-driven environments:

Bringing Three-Phase Power Down to the Cabinet Level

The most visible trend is the adoption of 240/415V three-phase power at the cabinet level. The rationale goes beyond efficiency: balancing loads effectively across an entire data center begins with balance inside each cabinet. By managing distribution at the lowest common denominator, operators can propagate stability upstream and avoid the phase imbalances that silently degrade resilience.

Three-phase distribution also reduces copper use and simplifies wiring, but the larger point is systemic balance. Done correctly, cabinet-level balancing ensures that megawatts delivered to the site are utilized evenly and predictably across white space.

The barrier, however, can be the retrofit cost. Many enterprise and colocation sites are still wired for 208V, and migrating to 240/415V requires new breakers, panelboards, and sometimes wholesale rewiring—investments few operators will make unless AI demand forces their hand. This explains the uneven adoption curve: hyperscalers lead, colos experiment, and enterprises hesitate.

Moving Toward Higher Distribution Voltages (480V and beyond)

As rack densities climb toward 100 kW and beyond, even 240/415V begins to strain. The arithmetic illustrates why: a 100A 208V deployment supports roughly 28.8 kW, while the same amperage at 415V supports 57.5 kW. Higher voltages unlock more usable power without increasing amperage—and since amperage drives distribution losses, the efficiency case is compelling.

Typical Circuit	Typical Plug Type	Max. Capacity (kW)
3-ph 30A, 208V	L21-30	8.6
3-ph 30/32A, 415V	IEC 60309 30A 3P+N+G / L22-30	17.3
3-ph 60A, 208V	IEC 60309 60A 3P+G	17.3
3-ph 60A, 415V	IEC 60309 60A 3P+N+G	34.5
3-ph 100A, 208V	IEC 60309 100A 3P+G	28.8
3-ph 100A, 415V	IEC 60309 100A 3P+N+G	57.5

Power capacity by typical input circuits and plug types.

The next step is 480V delivered directly to the cabinet. In North America, nearly every facility already receives 480V at the service entrance; bypassing step-down transformers reduces conversion losses and compresses footprint.

Delivering higher voltages to the cabinet also redefines expectations for the PDU. Standard units designed for 208V are quickly outmatched; **high-power PDUs capable of handling 240/415V and even 480V become essential to make these architectures viable in practice.** Without them, upstream efficiency gains are lost at the last step of distribution.

The obstacle is compatibility. Most IT gear power supplies were designed with 250V limits, and adoption of 480V downstream is constrained until server power supplies evolve. This creates the classic “chicken-and-egg” adoption dilemma: facilities are reluctant to build for 480V until IT gear supports it, while vendors wait for demand signals before enabling their hardware.

277V-Ready Servers: The Enabler for 480V

The emergence of servers with 277V-compatible power supplies changes the equation. By tapping one phase and neutral of a 480V system, operators can now deliver power directly to the rack without additional step-down gear. This reduces losses, simplifies architecture, and aligns with the natural distribution voltage already present in most U.S. facilities. Adoption is still early, led by hyperscalers, but broader OEM support could make 277V readiness a standard feature of the next refresh cycle.

Adoption is growing but not yet universal. Early adopters are hyperscalers that can influence vendor roadmaps. Enterprises and colos remain dependent on what OEMs make broadly available. Over the next refresh cycle, 277V capability may become standard—but until then, operators must manage mixed fleets, complicating distribution planning.

Direct Current (DC) Architectures

DC distribution is also resurfacing as a serious, if longer-term, option. Telecom operators have long relied on -48V DC, and some hyperscalers are piloting 384V DC and even 800V DC at the cabinet. The case is straightforward: eliminating AC–DC conversion stages improve efficiency and can reduce points of failure.

Yet ecosystem inertia remains powerful. IT gear, breakers, and connectors are all optimized for AC; standards for high-voltage DC are immature; and safety concerns are real. For now, DC remains in pilot mode, but the pressures of AI may push adoption faster than past cycles. When cabinet loads approach the scale of small industrial plants, the potential efficiency gains of DC become harder to ignore.

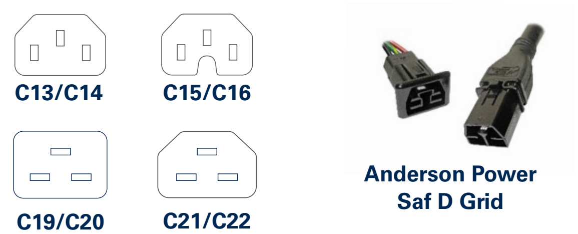
Connector Evolution at the IT Equipment Power Supply Level

Even the smallest interface points are adapting. Traditional C14 and C20 inlets are giving way to higher-temperature C16 and C21 variants, built to withstand the elevated rear-rack conditions of AI deployments. In parallel, 30A-rated SAF-D-GRID connectors—designed in the familiar C13 form factor—are emerging on AI hardware. The amperage ratings are unchanged in some cases, but the thermal headroom and current-handling capacity are not.

These shifts underscore a broader truth: supporting AI workloads requires rethinking not just topologies and voltages but every link in the distribution chain, down to the connectors themselves.

Power Supply Connector Trends - AI Equipment

- Higher kW capacity power suppliers use 16A or higher connectors
- C20 connectors being replaced with C21 / C14 with C16 – higher temperature versions
- Some AI power supplies coming with 30A rated Saf-D-Grid connectors



Rethinking Redundancy

Finally, redundancy is also under review. The 2N model, once the gold standard, is increasingly unsustainable at AI densities. At 10–20 kW, doubling every component was acceptable. At 100 kW, it is the equivalent of powering a neighborhood for a single rack.

Operators are recalibrating toward N+1 and N+2 approaches, where all power paths are active simultaneously. This reduces stranded capacity but raises the complexity of deployment in traditional data centers that have all been designed around 2N architectures.

At a Glance: Emerging Power Delivery Trends for AI

Trend	Why It Matters for AI Deployments
Cabinet-Level Optimization	Balancing loads at the cabinet prevents phase imbalance and ensures stable utilization across the facility.
240/415V Three-Phase Distribution	Reduces copper use, simplifies wiring, and enables load balancing at the cabinet level.
480V to the Cabinet	Doubles usable power for same amperage, reduces losses, and eliminates extra transformers.
277V-Ready Servers	Allow direct 480V distribution, bypassing step-down gear and aligning with facility voltages.
DC Architectures (48V / 384V)	Eliminates AC-DC conversions for efficiency and reliability; potential fit at ultra-high densities.
Connector Evolution	Higher-temp C16/C21 and 30A safety connectors withstand heat and higher current loads.
Rethinking Redundancy	N+1 and N+2 models reduce stranded capacity at scale, but require more complex coordination.

Critical PDU Design Features for AI Workloads

With these distribution trends in mind, attention turns to the cabinet's last link in the chain: the PDU. At AI densities, PDUs are no longer passive accessories but active points of engineering risk. Their design affects airflow, restart behavior, breaker coordination, and even human error rates.

This section examines which PDU features have become decisive at scale and how they shape long-term reliability.

High-Power PDUs: Enabling the Next Voltage Step

The most immediate implication of higher distribution voltages is the need for PDUs that can accept and manage them directly. Conventional PDUs, designed for 208V inputs, quickly become a limiting factor in AI deployments where 240/415V three-phase or 480V distribution is introduced at the cabinet. High-power PDUs close that gap. By taking higher voltages directly, they eliminate unnecessary step-down conversions, reduce losses, and align the last link in the chain with the architectures operators are beginning to adopt upstream.

The role is not simply about compatibility. At 80–120 kW per rack, the PDU becomes the gatekeeper of usable capacity: if it cannot handle the higher voltages safely and predictably, the efficiency gains of redesigning upstream power paths collapse at the cabinet. High-power PDUs therefore serve as both bridge and safeguard, ensuring that facility-level shifts translate into reliable, cabinet-level deliverability.

CPI eConnect® High-Power PDUs engineered to accept 415V and 480V inputs directly, and 30A breaker options designed for selective coordination at cabinet level.

Compact design keeps airflow clear and cable management simple.

Form Factor and Airflow: Hidden Cost of Bulk

In AI racks, every obstruction in the rear zone has a measurable effect on cooling efficiency. A PDU that intrudes into cable channels or blocks airflow paths can force higher fan speeds, interfere with liquid manifolds, or compromise cooling efficiency. What seems like a mechanical detail translates into reduced electrical headroom.

One strategy for minimizing form factor is the use of **3-phase Wye configurations such as 240/415V,** which allow for **single-pole breakers that consume less footprint.** Coupled with **30A branch circuit breakers instead of 20A,** operators reduce the total number of breakers required, freeing valuable rear-cabinet space for airflow and cable management.

Relay Technology: Recovery Predictability

The argument for bistable relays is often framed as energy savings. The deeper issue is predictability of restart. AI jobs can run for weeks; a power event followed by uncertain outlet state risks inconsistent recovery across hundreds of servers. Even if only a fraction fails to restart cleanly, the entire training run may need to be repeated. Relay choice, therefore, is less about watts saved and more about eliminating restart ambiguity—a risk that grows exponentially with rack density.

Outlet Flexibility: Managing Generational Churn

AI hardware lifecycles are collapsing; GPU generations now turn over in less than two years. Fixed outlet configurations lock operators into today's requirements, forcing mid-cycle rewiring or overbuild. Flexible receptacle designs reduce this friction, but they also raise new management questions: how to enforce safe allocation, prevent mismatched connections, and track outlet use over time. The trade-off is operational: more flexibility demands stronger governance, or else cabinets drift into undocumented complexity that complicates troubleshooting.

CPI's eConnect PDUs with QuadLock Outlets accept multiple plug types, supporting mixed AI hardware while reducing the risk of overbuild.

Breaker Sizing and Coordination: The Weak Link Problem

20A branch circuits on a PDU, once standard for enterprise racks, are no longer viable at 60–100 kW loads. **30A breakers are emerging as the de facto benchmark**, allowing higher device counts per branch and aligning with the draw of GPU-dense servers.

This shift gives operators more usable headroom within each circuit, reducing the need to spread devices across multiple branches. 30A branch circuit breakers also minimize the number of breakers needed to support finite capacities - keeping the overall footprint of the PDU low, allowing room for proper airflow and deployment of liquid cooling infrastructure.

A single mis-set or poorly coordinated breaker can black out half a rack. These failures are rarely exposed during commissioning, when loads are steady and predictable. They tend to emerge only in production, under the transient surges and uneven draw that characterize GPU training.

The real benchmark is not simply the adoption of 30A breakers but their integration into architectures where coordination is explicitly engineered and verified. Assuming protection schemes will scale from legacy designs is a critical error; at AI densities, the breaker becomes a frontline determinant of availability rather than a background safeguard.

Balanced Phase Outlets

With AI deployments where equipment power supply draws are high, it becomes very important that the loads within the cabinet get balanced at the phase and branch circuit level to allow optimal utilization. Balanced phase outlets (as the name suggests) allows loads to be deployed easily within the cabinet while keeping equipment power cord lengths at a minimum.

Error Prevention and Serviceability: Human Factors at Scale

Accidentally unplugging a single device can halt training jobs that have consumed thousands of GPU-hours. Color-coded outlets provide clear visual guidance for phase and load separation, reducing the chance of mis-wiring during installation or service. Locking connectors ensure cables remain seated under vibration, high cable tension, or human disturbance, preventing the kinds of subtle disconnects that are difficult to diagnose but catastrophic under load.

At 80–120 kW per rack, these features are not convenience options; they are safeguards against some of the most common and costly points of failure in the white space. In AI deployments, where the PDU itself represents a single point of success or failure, neglecting outlet-level error prevention turns routine service into a systemic risk.

CPI PDUs feature color-coded, built-in locking outlets to safeguard against disconnects under cable tension or service activity.

Managing AI Power Proactively at Scale

Proactive management at AI scale requires a cultural shift as much as a technical one. The operational risk is no longer limited to tripped breakers or downtime; it extends to financial losses, tenant trust, and security exposure. Competitiveness will hinge on systems designed to anticipate stress, contain faults, and prevent disruptions from cascading.

Here are the capabilities operators need to design for—so that failures become forecastable, isolated, and non-catastrophic:

Outlet-Level Visibility: From Optional to Non-Negotiable

At AI densities, visibility cannot stop at the rack or branch circuit. The cabinet becomes the true point of variability, and without **outlet-level data**, operators are effectively blind to the dynamics that destabilize high-density deployments.

Uneven phase loads, creeping imbalances, or a single overdrawn outlet can tip a cabinet into fault under AI workloads. These conditions rarely announce themselves at higher aggregation points; they surface only when measured at the outlet. For MTDC operators, the stakes are higher still: AI tenants often demand direct access to this telemetry, making outlet-level monitoring not just an operational safeguard but a market requirement.

The shift is clear—outlet-level visibility has become the baseline for safe load management, proactive diagnostics, and tenant trust in AI-ready environments.

Capability	Advantages
Basic Power Distribution	• Easily distribute power to IT loads within a cabinet / rack
Local Metering	• Balance loads across phases and branch circuits at initial install
Remote Metering - Input	• Manage overall cabinet level power consumption • Balance loads across all input phases • Chargeback based on cabinet level power consumption
Remote Metering – Branch Circuit	• Balance loads across phases and branch circuits at all times • Notification of availability issues before problems occur • Chargebacks based on overall PDU power consumption
Remote Metering – Outlet Level	• Know power consumption of IT equipment • Justify purchase of new equipment • Help determine unused servers • Chargebacks / regulatory credits based on actual IT equipment power consumption
Remote Power control – Outlet Level	• Reboot power to hung up equipment remotely • Scheduled shut down and start up for energy savings (IT Labs) • Provisioning of outlets • Turn off outlets if a leak is detected with liquid-cooled deployments
Integrated Environmental Monitoring / Access Control	• Use PDU for overall cabinet level management

From Overprovisioning to Precision: The Case for DCIM

Without granular monitoring and predictive analytics, operators must reserve capacity for peak scenarios, leaving large portions of infrastructure stranded. **Data Center Infrastructure Management (DCIM)** platforms reduce this waste by identifying real utilization patterns, automating alerts, and guiding load redistribution in real time.

While DCIM systems have historically been viewed as "nice-to-have" overhead, at AI densities, they become non-negotiable to maintaining efficiency at scale.

Integrated Environmental and Security Data

AI racks erase the divide between thermal, electrical, and security concerns. Exhaust zone temperature rates are high, liquid cooling introduces new risks, and PDUs now operate at the edge of their thermal envelope. Cabinet-level sensing of temperature, humidity, and leaks—analyzed alongside breaker and load data—is the only reliable way to detect early warning signs before failures propagate.

At the same time, the value of AI hardware makes cabinets prime security targets. Smart locks, role-based access, and tamper monitoring are converging with power oversight, transforming the cabinet into a governed unit of infrastructure. By integrating environmental telemetry with physical and digital security, operators gain the visibility and control needed to protect workloads worth millions per rack.

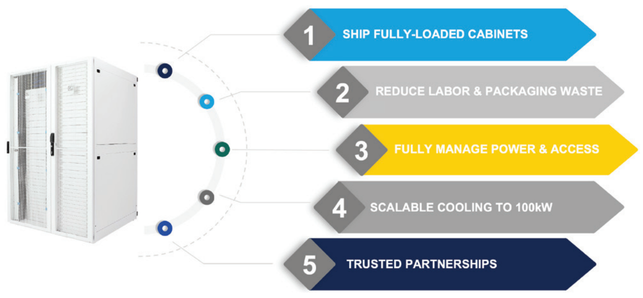
Integrated solutions like CPI's ZetaFrame® Cabinet System delivering power, environmental sensors, and access control in one solution.

Best PDU Deployment Practices to Avoid Delays and Downtime

For AI projects, deployment risks are not only about speed but unpredictability. Schedules tied to hardware availability or competitive launches leave little room for rework once equipment arrives on-site. The most effective way to protect timelines is to reduce variability upstream.

Factory pre-installation of PDUs, sensors, and supporting components transforms the cabinet from a frame into a deployment-ready system. Integration and testing at the factory ensure interoperability is verified once—rather than rediscovered across dozens or hundreds of racks in the field.

Key Benefits of an Integrated Cabinet System



The advantage extends beyond power integration:

- **Cooling alignment –** When rear-door heat exchangers, liquid manifolds, or sensor arrays ship pre-integrated, operators avoid field-level friction between mechanical and electrical trades. Synchronization happens at the factory, not during commissioning.
- **Sustainability benefits –** Pre-integration reduces packaging waste and truck rolls, aligning deployment with corporate sustainability commitments.
- **System-level validation –** Load conditions and controls are tested together in advance, ensuring power and cooling interoperate as designed before reaching the white space.

The payoff is twofold: faster installation and consistent reliability. Cabinets that ship with power and cooling systems already installed protect operators from last-minute delays, keep commissioning aligned with program deadlines, and provide predictable performance at scale.

Conclusion: Building a Future-Ready AI Power Ecosystem

AI is not a temporary spike in demand but a structural shift in computing. It is exposing the limits of conventional design and making the cabinet the true unit of resilience.

Without a cabinet-level strategy, operators risk stranded capacity, thermal failures, costly deployment delays, or overbuilding infrastructure that delivers little efficiency. With thoughtful planning, the opposite is possible: resilient, efficient, future-ready ecosystems.

Three priorities stand out:

- **Engineer for AI densities –** PDUs, breakers, and thermal resilience must scale to 30A circuits and beyond.
- **Integrate as a system –** Power, cooling, and structural design must converge at the cabinet.
- **Design for refresh cycles –** High dynamic load ratings and flexible outlets preserve reliability as hardware generations accelerate.

The lesson extends beyond AI. Quantum, neuromorphic, and other post-AI architectures will test the same boundaries. Decisions made today will determine whether future workloads can be energized on schedule—or whether another round of retrofits and compromises will be required.

Infrastructure That Meets AI at the Cabinet

Meeting the demands of AI at scale requires infrastructure designed from the cabinet outward. CPI's portfolio is built with these requirements in mind:

- **High-Power PDUs –** Capable of accepting 240/415V or 480V inputs directly, CPI's eConnect® PDUs align with higher-voltage architectures now entering mainstream use. With 30A breakers, QuadLock outlets, and tolerance for sustained 65°C environments, they provide a foundation for reliable cabinet-level power.
- **Cabinet Systems Built for AI Densities –** CPI's ZetaFrame® Cabinet System supports high dynamic load ratings (4,000 lbs. or more) and integrates seamlessly with high-power PDUs and advanced cooling, ensuring electrical and mechanical stability across refresh cycles.
- **Deployment-Ready Integration –** Factory pre-integration of PDUs, sensors, and containment options reduces on-site variability, compresses deployment timelines, and protects AI programs from costly delays.
- **Visibility and Control –** Intelligent monitoring capabilities, integrated with CPI PDUs and environmental sensors, give operators the outlet-level data and predictive alerts necessary to manage high-density racks at scale.

References

[1] *Uptime Intelligence: AI to trigger radical overhaul of data center electrification*
https://intelligence.uptimeinstitute.com/resource/ai-trigger-radical-overhaul-data-center-electrification

Contributors

Ashish Moondra, Senior Director of Electronics and Software

Ashish has more than 25 years of experience developing, managing and selling rack power distribution, uninterruptible power supply (UPS), energy storage and Data Center Infrastructure Management (DCIM) solutions. Ashish has previously worked with American Power Conversion, Emerson Network Power and Active Power, and has been an expert speaker at various data center forums including teaching a BICSI Masterclass on power protection, distribution and management.

chatsworth.com
techsupport@chatsworth.com
800-834-4969