

WHITE PAPER

UPS FOR AI DATA CENTERS
The challenge for UPS Technology

#LegrandImprovingLives

legrand®

LEGRAND SUPPORTS YOU ON ALL YOUR PROJECTS

Artificial intelligence is reshaping the electrical profile of servers and data centers in ways that go far beyond a simple increase in installed power. Unlike traditional IT environments, where diversified workloads generally create a relatively stable aggregate demand, AI infrastructures—particularly those based on GPU-intensive training and inference—can generate rapid, repetitive and highly synchronized variations in power consumption.

These fluctuations may occur within milliseconds, repeat continuously during computation cycles and create a much more demanding operating environment for the electrical infrastructure that supports them. As a result, the key challenge is no longer only how much power must be delivered, but also how quickly the load changes, how frequently these changes occur and how the entire power chain responds under real operating conditions.

AI servers and clusters can introduce steep ramp rates, large step loads, nonlinear behavior and greater exposure to power quality issues such as voltage disturbances, harmonic distortion, flicker and transient stress. These effects are not confined to the IT load itself: they can propagate through switchboards, transformers, busways, backup systems and the wider electrical distribution architecture of the facility.

In this context, the UPS has become more strategic than ever. It is no longer merely a safeguard against outages, but a critical component for maintaining electrical stability in the presence of highly dynamic and unpredictable loads. Supporting AI-ready environments requires UPS systems capable of managing recurring load steps, fast transients and frequent operating-state changes without compromising output quality, efficiency, or battery life.

For designers, operators, consultants and decision-makers, understanding the impact of AI workloads on the electrical system is now essential to building infrastructures that are not only resilient, but genuinely prepared for the next generation of compute.



Summary

Main Characteristics of AI Data Centers 2

Power Density Jump 2
Not Only Power, but Dynamics 4
Synchronized Cluster Behavior 5
Training versus Inference 5
Power Quality: A Design Variable for AI Infrastructure 7
A New Electrical Infrastructure 8

UPS AI-Ready 10

Impact of AI on UPS Systems 10
A New Generation of UPS: Features and Performance 10
- Stability of the Energy Feed 12
- Grid Protection 13
- Battery Support 14
- High Efficiency Across the UPS Load Profile 15
- The Technical Relevance of SiC-Based UPS Technology 16

UPS as an Element of AI Infrastructure 17

Conclusion 18

Choosing the UPS for AI Applications 19
Keor FLEX UPS: The Legrand UPS Solution 20

MAIN CHARACTERISTICS OF AI DATA CENTERS



Power Density Jump

The first and most visible characteristic of an AI Data Center is the sharp increase in power density compared with traditional enterprise IT environments. Conventional server rooms and legacy data centers have historically been designed around relatively moderate rack densities, typically in the range of 5–15 kW per rack. AI-oriented deployments, by contrast, commonly move toward 30–120kW and even more per rack, with the most advanced liquid-cooled configurations already exceeding this level by a significant margin.

Industry references also highlight flagship AI rack configurations reaching approximately 132 kW per rack, confirming that AI does not represent a gradual increase in power demand, but rather a structural leap in the electrical intensity of the computing environment.

IT SERVER RACKS TYPE	MAIN CHARACTERISTICS
Traditional enterprise IT	Historically lower-density server rooms and mixed IT loads; exact range depends on facility age and design.
Current AI / HPC deployments	Higher rack density driven by GPU servers, high-bandwidth interconnects and liquid-cooling adoption.
Advanced liquid-cooled AI racks	Very high density environments may move well above conventional air-cooled assumptions.
Future extreme rack-density projections	Several hundred kW per rack and beyond are emerging design scenarios, not universal market conditions.

TABLE 1: TYPICAL RACK TYPE

Looking ahead, some projections indicate that rack densities could reach 250–900 kW per rack or more in extreme scenarios. At site level, large AI campuses may reach continuous power absorption in the range of 50–100 MW, with major training jobs potentially exceeding 100 MW. In practical terms, AI can introduce a power density increase from 10 to over 20 times compared with traditional IT environments.



RACK POWER DENSITY SHIFT

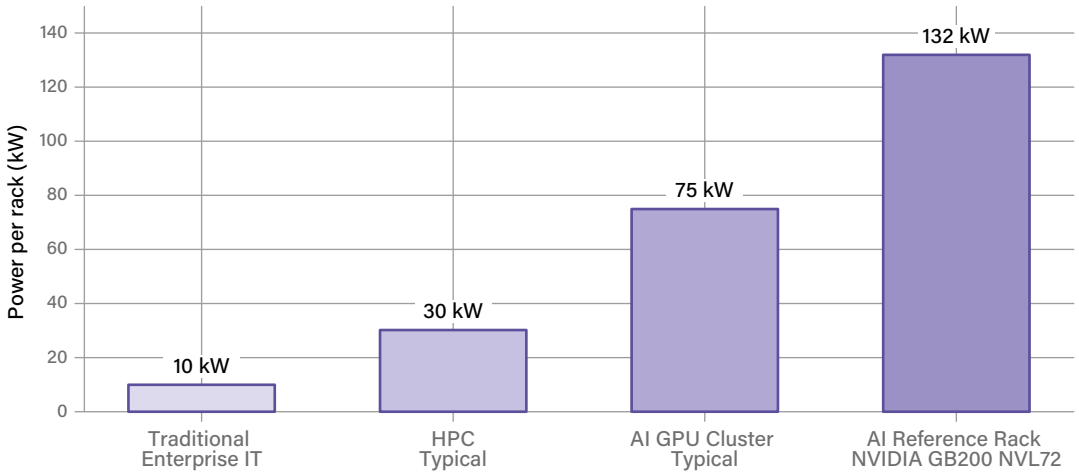


FIGURE 1: RACK POWER DENSITY INCREASE

This power jump is driven primarily by GPU-centric computing architectures, high-bandwidth interconnects, dense memory configurations and the need to sustain massive parallel computation over extended periods. Consequently, electrical design assumptions inherited from traditional IT infrastructures are no longer sufficient.

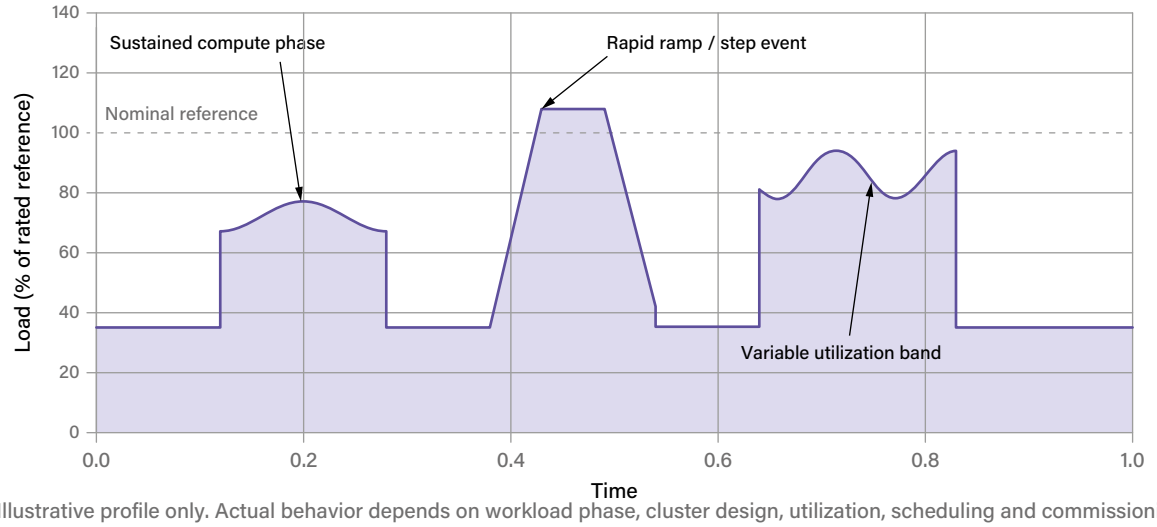
Feeder sizing, busway ratings, transformer loading, breaker coordination, redundancy strategy and cooling integration must all be reassessed against a much higher baseline load. AI data centers are not simply larger versions of conventional facilities; they represent a new order of magnitude in rack-level and room-level power concentration.

Not Only Power, but Dynamics

The defining challenge of AI infrastructure is not only the increase in installed power, but the dynamic behavior of the load itself. Traditional IT estates typically generate a relatively smooth aggregate electrical profile because different applications run asynchronously across large populations of servers.

AI workloads behave differently. Their power demand can vary significantly across different phases of computation, including model execution, communication, synchronization and memory movement. During rapid ramp or step events, demand can rise and fall very quickly, often in short repetitive cycles linked to the internal structure of the workload.

ILLUSTRATIVE AI WORKLOAD POWER BEHAVIOR



Illustrative profile only. Actual behavior depends on workload phase, cluster design, utilization, scheduling and commissioning data

FIGURE 2: AI WORKLOAD PERFORMANCE

These profiles are highly dynamic and impulsive, with fluctuations occurring within milliseconds and repeating continuously during operation.

ILLUSTRATIVE AI WORKLOAD POWER BEHAVIOR

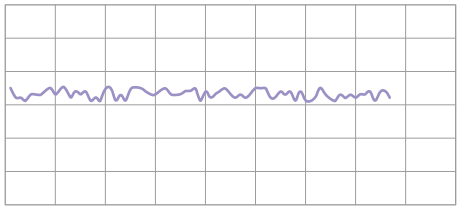


FIGURE 3: TRADITIONAL SERVER POWER ABSORPTION

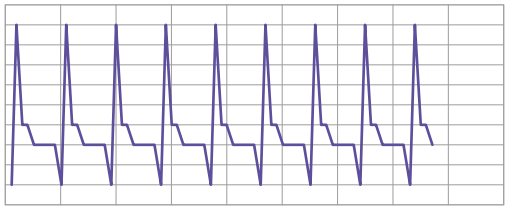


FIGURE 4: AI SERVER POWER ABSORPTION

This fundamentally changes the design question. The issue is no longer only “how many kilowatts must be supplied?”, but also “how quickly can the load change, how often does this happen and how does the electrical infrastructure respond?”

In AI environments, steep ramp rates, recurring step loads and oscillating demand patterns can place significant stress on switchboards, PDUs, UPS rectifiers and inverters, as well as upstream power sources. A system that appears adequately sized from a steady-state perspective may still prove vulnerable if it cannot respond quickly enough to the real operating behavior of the AI load.

Synchronized Cluster Behavior

A further distinguishing characteristic of AI data centers is the synchronized behavior of large GPU clusters. In conventional IT environments, server-level power changes are usually distributed over time and partially cancel each other out at facility level.

In AI training clusters, by contrast, many compute nodes can move through similar processing phases almost simultaneously. This near-synchronous behavior can generate frequent power spikes and drops, creating much larger electrical swings than those typically observed in mixed enterprise workloads.

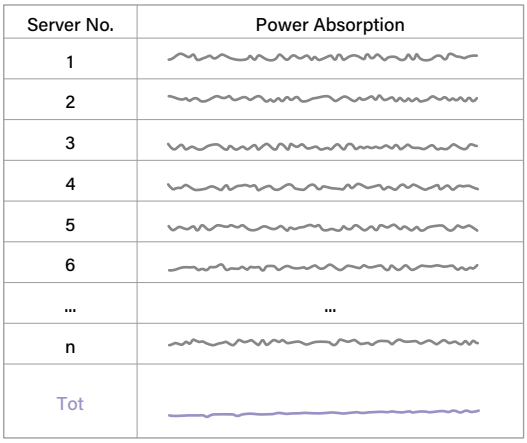


FIGURE 5:TRADITIONAL DATACENTER POWER ABSORPTION

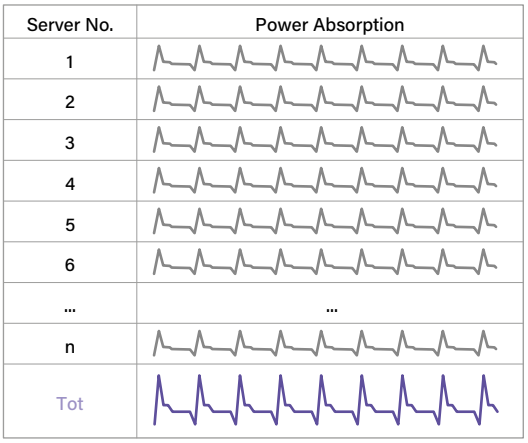


FIGURE 6:AI DATACENTER ABSORPTION WITH AL SERVERS SYNCHRONIZED

This synchronization effect is critical because it amplifies the impact of each compute cycle across the entire power train. What would otherwise be manageable local variation becomes a system-level electrical event.

In more extreme cases, power demand can more than double between idle and peak compute phases These synchronized transitions can stress not only the UPS, but also transformers, generators, busways, power electronics and even cooling-related electrical subsystems if these are responding at the same time.

Training versus Inference

Not all AI workloads behave in the same way. From an electrical perspective, training and inference create different load signatures and therefore different infrastructure implications.

WORKLOAD / OPERATING CONDITION	TYPICAL POWER BEHAVIOR	INFRASTRUCTURE IMPLICATION	VALIDATION NEED
Training clusters	High sustained demand with possible cyclic variation	Capacity, thermal design, synchronized load behavior	Measured load profile and transient envelope
Inference platforms	Variable demand with request-driven bursts	Part-load efficiency, fast response, monitoring	Utilization and concurrency model
Phased deployment	UPS may operate below final design load for extended periods	Weighted efficiency and modular growth become critical	Deployment roadmap and expected loading
Mixed AI / HPC environments	Multiple workload patterns coexist	Redundancy and load diversity assumptions need review	Telemetry and commissioning data

TABLE 2: TYPICAL WORKLOAD CONDITIONS

Training versus Inference (continued)

Training workloads generally involve very high, sustained computational demand. They often run across large GPU clusters for extended periods, creating substantial base load at facility level, frequently measured in megawatts. However, this sustained demand may still include repeated oscillations linked to the internal rhythm of the training process.

Inference workloads, on the other hand, may operate at a lower average load but are often characterized by shorter, recurring spikes driven by user requests, real-time processing events, or bursty service demand. From an infrastructure standpoint, inference may be less dominated by long-duration maximum draw, but more exposed to frequent transient behavior and rapid short-term load changes.

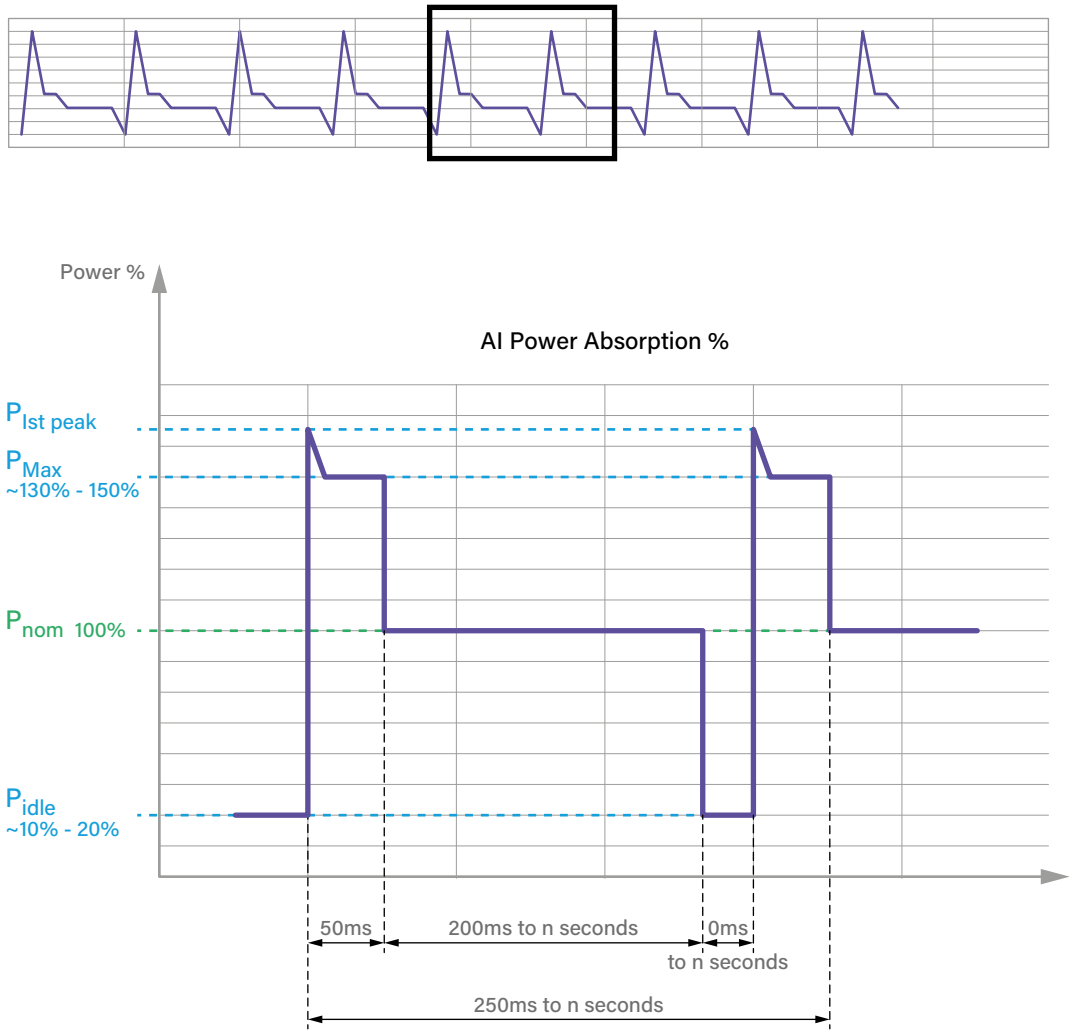


FIGURE 7: INDICATIVE AI SERVER LOAD PROFILE

Modern AI data centers may host both profiles simultaneously. This makes it necessary to understand not only peak demand, but also duty cycle, ramp behavior, temporal diversity and the interaction between different classes of AI applications.

Power Quality: A Design Variable for AI Infrastructure

As AI workloads become denser and more dynamic, power quality becomes a central design concern rather than a secondary operational issue. AI servers, GPU clusters and associated power electronics introduce highly nonlinear demand, which can contribute to voltage disturbances, harmonic distortion, transient events and flicker.

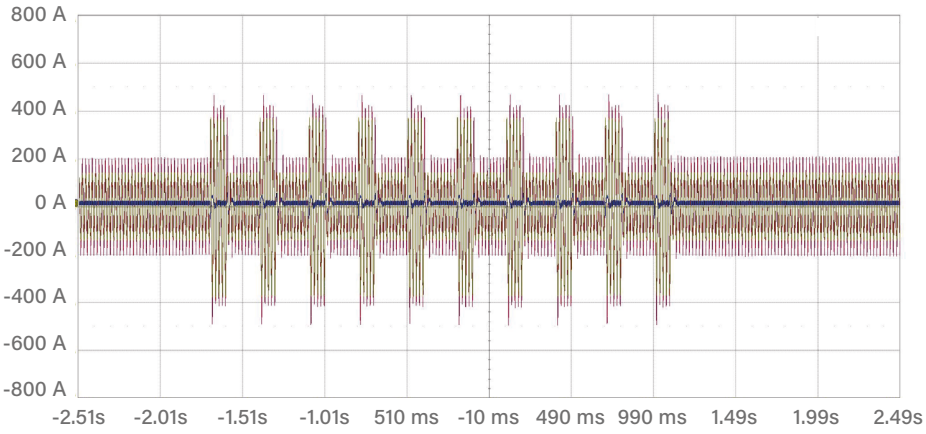


FIGURE 8: AC CURRENT ABSORPTION

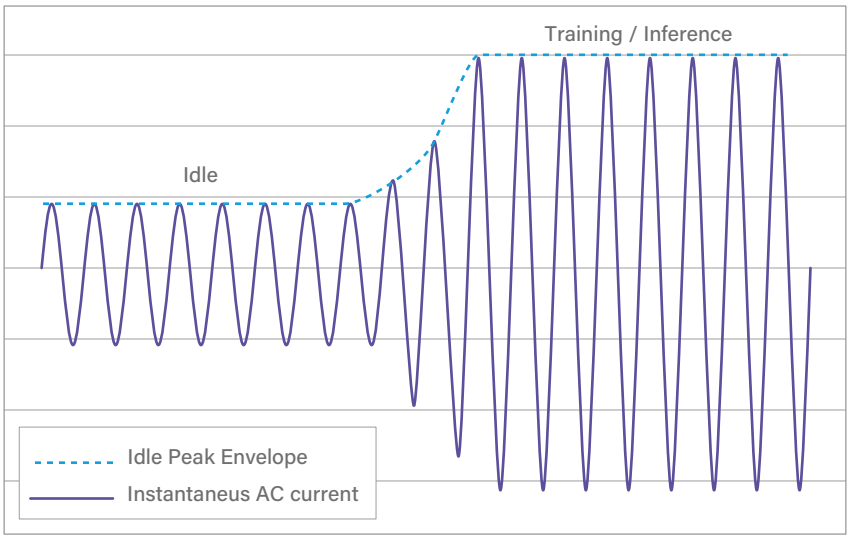


FIGURE 9: AC CURRENT ABSORPTION TRANSIENT FROM IDLE TO TRAINING/INFERENCE

AI-driven facilities operate in highly dynamic load environments where sudden spikes—such as server startup events, inference bursts, or the launch of a training sequence—can cause rapid increases in demand and create measurable stress on the electrical system.

This represents a major shift for operators and designers. Power quality issues are no longer limited to the external grid or utility supply. The internal white space of the AI Data Center itself becomes a source of electrical disturbance.

Fast-changing and repetitive current demand can challenge voltage regulation, increase thermal stress on components, worsen harmonic conditions and reduce overall efficiency. In this context, real-time monitoring of electrical events, harmonics, transient behavior and high-resolution load variation becomes essential to maintain both resilience and performance.

AI therefore introduces not only a capacity challenge, but also a power quality management challenge that must be addressed from the earliest design stages.

A New Electrical Infrastructure

Taken together, these characteristics require a new type of electrical infrastructure for AI data centers. The traditional approach—based primarily on average load, static sizing and classical IT redundancy assumptions—must evolve toward a model capable of managing high density, fast dynamics, synchronized transients and continuously varying operating conditions.

This affects the entire architecture: utility interface, MV/LV transformation, distribution topology, UPS strategy, battery selection, generator integration, busway design, protection coordination and monitoring.

AI-ready infrastructure must therefore be designed not only for capacity, but also for responsiveness, controllability and stability under dynamic stress.

Within this new architecture, the UPS plays a pivotal role. It can no longer be considered a passive backup system reserved for rare outage events. Instead, it becomes a critical stabilizing layer capable of tolerating recurring step loads, maintaining output quality during rapid transitions and interacting intelligently with batteries and upstream sources.

Conventional performance metrics are no longer sufficient to describe real-world behavior in AI facilities. Battery technology choices also become more strategic, since frequent and rapid load variations can accelerate cycling stress if the energy storage system is not properly matched to the application.

For this reason, the electrical infrastructure of the AI Data Center must be conceived as an integrated system in which distribution, protection, power continuity, monitoring and energy storage are all designed around the actual behavior of AI compute.





UPS AI-READY



Impact of AI on UPS Systems

The rise of AI workloads changes the role and operating conditions of the UPS more profoundly than any recent shift in Data Center architecture.

In traditional environments, UPS systems were primarily sized and validated against predictable load evolution, short-duration transients and conventional backup scenarios. In AI-oriented infrastructures, the UPS is increasingly exposed to recurring and highly dynamic step loads generated by GPU clusters whose demand can change rapidly and repeatedly during normal operation.

This creates a new type of electrical stress. The UPS must manage not only peak demand, but also fast load ramps, frequent transitions, nonlinear current draw and repeated cycling conditions that can challenge both power electronics and energy storage.

AI therefore has a direct impact on several UPS design and application criteria: inverter response, overload management, battery interaction, heat dissipation, part-load efficiency, variable-load efficiency and performance during source transitions.

The challenge is not limited to maintaining power during outages. It also includes preserving output stability while the IT load itself is actively oscillating. In other words, the UPS must operate as a dynamic buffer between an increasingly unstable downstream load profile and an upstream power system that may include utility supply, generators, transformers and switching devices with their own response limitations.

Typical UPS selection metrics are no longer sufficient on their own. Nameplate capacity, autonomy time and static overload capability remain important, but they must now be complemented by AI-specific criteria such as tolerance to recurring step loads, transient recovery behavior, battery micro-cycling control and compatibility with rapidly changing input and output conditions.

As AI loads become denser and more synchronized, the UPS must be evaluated as an active infrastructure component operating under a fundamentally new regime.



A New Generation of UPS: Features and Performance

To address these new operating conditions, the market is moving toward a new generation of UPS platforms designed not only for continuity, but also for dynamic load management.

These systems are being developed with stronger control capabilities, improved transient performance, higher overload tolerance, more advanced battery interaction logic and a broader role in maintaining stability across the entire power train.

The shift is both qualitative and quantitative. The next-generation UPS is not merely a larger system with more kVA. It is a platform that combines power conversion, energy buffering, control intelligence and monitoring to absorb the operational consequences of AI load behavior.

In this context, performance must be understood not only in terms of backup duration or nominal efficiency, but also in terms of power-feed stability, upstream grid or generator protection and intelligent battery use to support dynamic operation without accelerating degradation.

▪ Stability of the Energy Feed

One of the primary expectations from a UPS in an AI environment is the ability to maintain a stable and high-quality energy feed to the critical load, even while that load is changing rapidly.

This means preserving output voltage, frequency, waveform quality and transition continuity under conditions that are more demanding than those typically seen in conventional Data Center applications. A key requirement is the ability of a large UPS to maintain regulated output while the downstream load cyclically alternates between high-compute and lower-load phases, including during utility operation and generator transition.



FIGURE 10: STABILITY OF UPS OUTPUT VOLTAGE WITH AI LOAD

For AI applications, feed stability is not simply a matter of “power on” versus “power off”. It is a matter of maintaining electrical quality during constant load movement.

Voltage deviations, brief dips, control-loop instability, or poorly managed transient recovery can all become more problematic when thousands of GPUs operate in a tightly coupled environment.

The new expectation is therefore a UPS capable of decoupling the critical load from instability created both upstream and downstream, ensuring that dynamic AI servers are continuously supplied by an electrically stable source under real operating conditions.

▪ Grid Protection

A second key function of next-generation UPS systems is the ability to protect the upstream electrical system from the consequences of AI load volatility. This capability comes from specific peak shaving/power smoothing control logic implemented in the UPS and the use of battery systems able to support frequent charge discharge cycles.

In traditional thinking, the UPS primarily protected the load from the grid. In AI data centers, the opposite direction also becomes important: the UPS must help prevent highly dynamic downstream behavior from propagating back toward switchboards, transformers, generators and utility interfaces.

This is especially relevant when synchronized GPU clusters create recurring pulses, steep ramp rates, or harmonic stress that may otherwise destabilize the broader electrical environment.

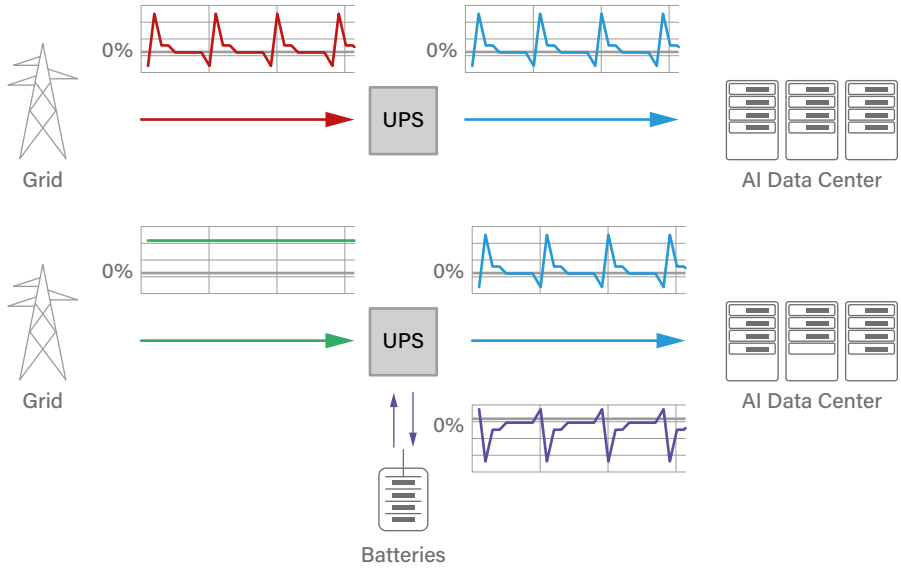


FIGURE 11: THE ABSORPTION PROFILE WITHOUT AND WITH THE USAGE OF BATTERY TO STABILIZE THE ABSORPTION FROM THE GRID

This shielding role is increasingly supported by advanced control strategies such as input power smoothing and controlled battery participation. The objective is to make the load seen by the upstream source more gradual, more predictable and less disruptive than the raw AI demand profile would otherwise be.

This is particularly valuable when the upstream source includes standby generators or constrained grid connections, both of which may be sensitive to fast power excursions.

From a system-design perspective, grid protection becomes a strategic functionality. It is not simply a marketing concept, but a practical way to improve compatibility between AI compute and the wider electrical infrastructure of the site.

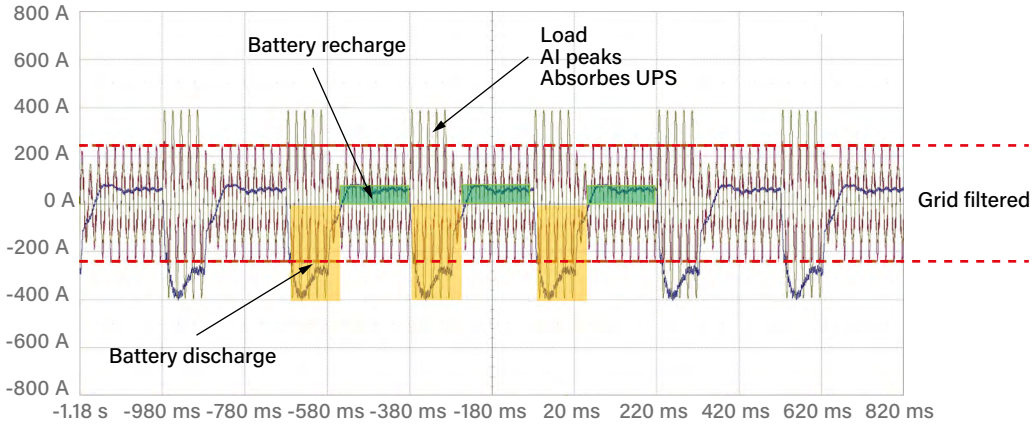


FIGURE 12: THE CURRENT ABSORPTION PROFILE OF AI LOAD VS BATTERY VOLTAGE DURING GRID PROTECTION MODE

A New Generation of UPS: Features and Performance (continued)

▪ Battery Support

Battery behavior becomes significantly more critical in AI-oriented UPS operation than in many traditional applications.

Frequent load fluctuations can lead to repeated charge and discharge activity, short-duration interventions and partial cycling patterns that may not have been dominant in legacy designs. If the UPS uses the battery too aggressively to stabilize AI-driven transients, the result may be accelerated battery wear, additional thermal stress and reduced effective lifetime.

For this reason, modern AI-oriented UPS strategies increasingly include specific control modes designed to limit unnecessary battery usage while still making energy storage available when it is genuinely beneficial to system stability.

Battery chemistry also becomes more relevant. Beyond response capability, the choice of battery technology influences maintenance strategy, lifecycle cost, thermal behavior, charging characteristics and overall resilience under repeated high-rate events.

In AI infrastructure, battery support is therefore not only a runtime issue. It is a design issue at the intersection of UPS control logic, power smoothing strategy, total cost of ownership and long-term operational reliability.

Today, lithium-ion LFP is one of the most common solutions, although other chemistries may offer advantages depending on application requirements, load profile, footprint constraints, safety priorities and lifecycle expectations.

TECHNOLOGY	MAIN PROS	MAIN CONS	BEST USE
LFP	Safe, long life, compact, fast recharge.	Higher upfront cost than VRLA.	Good choice for most new AI data center.
LTO	Excellent for rapid load swings, very fast recharge, very high cycle life.	Expensive, lower energy density.	Good for highly dynamic AI loads and advanced UPS/hybrid setups.
NiZn	High power density, compact, strong safety positioning.	Less mature ecosystem / fewer reference deployments.	Strong alternative when safety and footprint are priorities.
NMC	Very high energy density.	Usually less preferred than LFP on thermal/safety profile.	Good when space is the main constraint.
VRLA	Lowest upfront cost, proven technology.	Shorter life, larger footprint, slower recharge, more temperature sensitive.	Mainly for legacy / CAPEX-driven projects.
Pure Lead AGM	Better than standard VRLA on life and temperature tolerance.	Still bulkier and less future-proof than lithium options.	Mostly a retrofit / transition option.
Sodium-Ion	Promising for high power, rapid recharge, long cycle life.	Still early-stage for mainstream AI Data Center UPS.	Technology to watch / pilot.

TABLE 3: BATTERY TECHNOLOGY TYPE

▪ High Efficiency Across the UPS Load Profile

Efficiency is a critical parameter for UPS systems in AI data centers, but it can no longer be assessed at a single operating point.

Because AI workloads are highly dynamic, shifting rapidly between idle conditions, inference peaks, communication phases and sustained training cycles, the relevant metric is the ability of the UPS to maintain high and stable efficiency across a wide load range.

A narrow peak-efficiency value is no longer sufficient if performance degrades during frequent load transitions, as this leads to higher conversion losses in real operation.

From both a technical and commercial perspective, efficiency stability directly affects operating expenditure. Electrical losses increase energy consumption and are dissipated as heat, adding to cooling demand. In high-density AI facilities, where both power and thermal loads are already elevated, a flatter efficiency curve provides a dual benefit: lower direct electrical losses and reduced cooling energy, with a positive impact on total cost of ownership.

For this reason, peak efficiency alone is not an adequate indicator for AI-ready UPS platforms. Greater importance should be given to weighted efficiency, measured according to recognized standards such as IEC 62040-3 and the European Code of Conduct for UPS, which evaluate performance across multiple load points.

This approach provides a more realistic basis for estimating annual losses, operating costs and lifecycle value under actual AI load profiles.

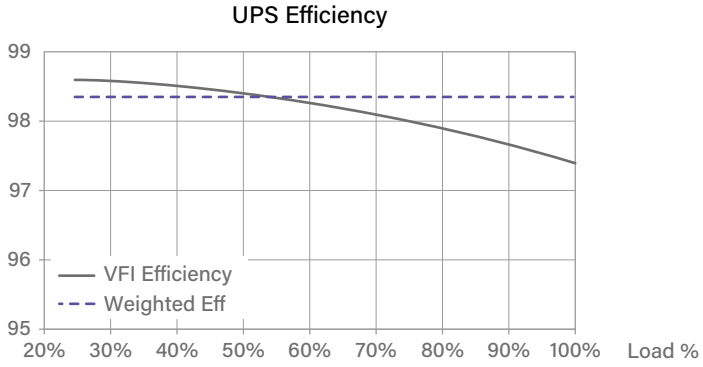


FIGURE 13: EFFICIENCY PROFILE WITH WEIGHTED EFFICIENCY OF A LAST GENERATION UPS SIC BASED TECHNOLOGY.

Accordingly, UPS selection criteria for AI applications should include weighted efficiency values and full efficiency curves across the expected operating range, ideally supported by performance data under representative dynamic AI load conditions.

In this context, next-generation UPS systems can exceed 98% efficiency while maintaining very high efficiency across the load range, enabled by SiC-based power electronics, advanced control algorithms and high-performance internal power distribution architectures.

EVALUATION POINT	WHY IT MATTERS
Efficiency at 25/50/75/100% load	AI and phased deployments may not operate at one steady full-load point.
Weighted efficiency	Better reflects expected operating profile than a single peak point.
Losses in kW	Directly converts efficiency into heat and energy cost.
Cooling impact	UPS heat rejection adds cooling energy requirement.
10-year TCO sensitivity	Makes efficiency commercially visible for decision-makers.

TABLE 4: KEY POINTS TO CONSIDER FOR ENERGETIC EVALUATIONS

▪ The Technical Relevance of SiC-Based UPS Technology

Silicon Carbide (SiC) is an advanced wide-bandgap semiconductor technology that is increasingly adopted in high-performance UPS systems. When integrated into a well-engineered UPS architecture, SiC can reduce power-conversion losses, improve thermal performance and support higher power density compared with conventional silicon-based technologies.

For AI data centers, where power demand is high and operating conditions can vary significantly, UPS efficiency is a critical factor. Every kilowatt lost during power conversion becomes heat that must be removed by the cooling infrastructure. By maintaining high efficiency across a broad load range, SiC-based UPS solutions can reduce energy losses, lower cooling requirements and contribute to a more attractive total cost of ownership (TCO).

SIC DESIGN LEVER	TECHNICAL RELEVANCE	BUSINESS IMPLICATION
Lower conversion losses	Reduces energy dissipated in the UPS power stage.	Lower OpEx and lower cooling burden.
Higher switching capability	Can support more compact and responsive power-conversion design when properly engineered.	Better fit for high-density technical rooms.
Improved thermal behavior	Helps reduce heat stress inside the UPS architecture.	Supports reliability and service-life objectives.
Part-load efficiency potential	Supports strong efficiency where phased or variable deployments may operate.	Improves lifecycle TCO modeling.
Architecture integration	Benefits depend on controls, busbar design, serviceability and validation.	Avoids technology hype; focuses on measurable value.

TABLE 5: SIC KEY FEATURES WHICH BRINGS TECHNICAL AND ECONOMICAL ADVANTAGES

The value of SiC extends beyond peak efficiency. Its strategic advantage lies in the combined benefits of improved energy performance, reduced thermal stress, compact system design and support for high-density deployments. Lower internal losses can help create a more favorable operating environment for power electronics while enabling greater installation flexibility in space-constrained electrical rooms.

However, SiC technology alone does not guarantee superior UPS performance. The overall effectiveness of a UPS depends on the complete system architecture, including control algorithms, thermal management, protection systems, firmware, serviceability and validated performance under real operating conditions. Therefore, the key consideration is not simply the use of SiC devices, but the measurable operational benefits delivered by the complete UPS solution.

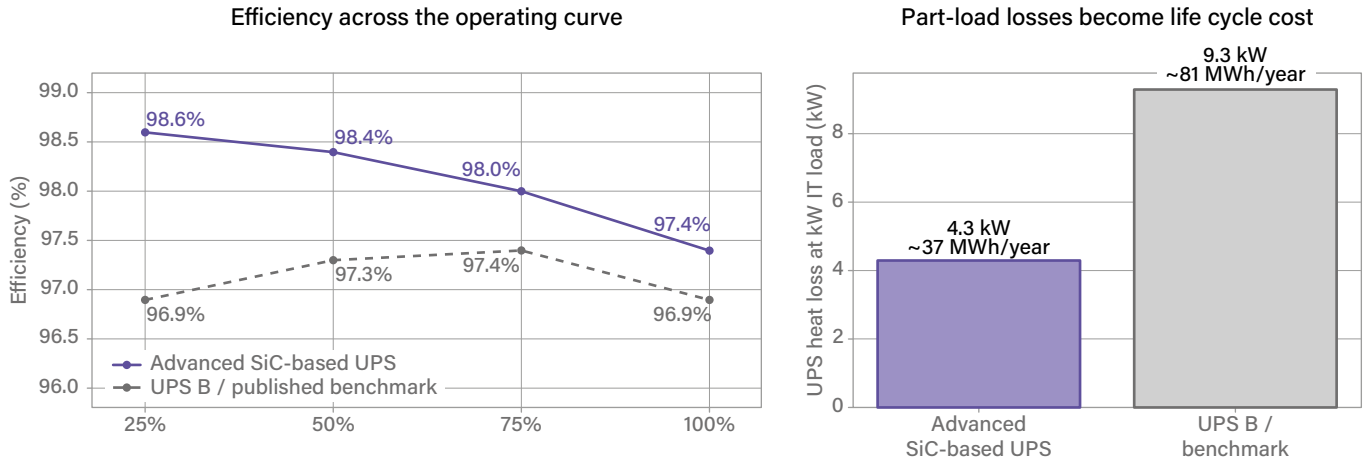


FIGURE 15: EXAMPLE OF COMPARISON OF EFFICIENCY AND LOSSES BETWEEN SIC UPS AND TRADITIONAL UPS

When evaluating a SiC-based modular UPS for AI applications, decision-makers should consider validated metrics such as full-load and partial-load efficiency, weighted efficiency, thermal-loss performance, dynamic-load response, maintenance strategy and modular scalability. For example, advanced SiC-based UPS platforms can deliver efficiency levels of up to 98.6% in double-conversion mode, scalability up to 1.2 MW per frame and up to 4.8 MW in parallel configurations, subject to the specific product rating, operating mode and approved system configuration.

UPS as an Element of AI Infrastructure

These developments lead to a broader conclusion: in an AI data center, the UPS can no longer be viewed as an isolated backup component. It must be considered a structural element of the AI infrastructure itself.

As AI workloads become denser, faster and more synchronized, the UPS assumes a multifunctional role. It protects critical loads from outages, stabilizes voltage and frequency under dynamic conditions, reduces the propagation of load volatility toward the upstream network, coordinates with batteries and energy storage systems and contributes to the overall controllability of the site.

In many ways, the UPS becomes a core node in the electrical architecture of AI-ready facilities.

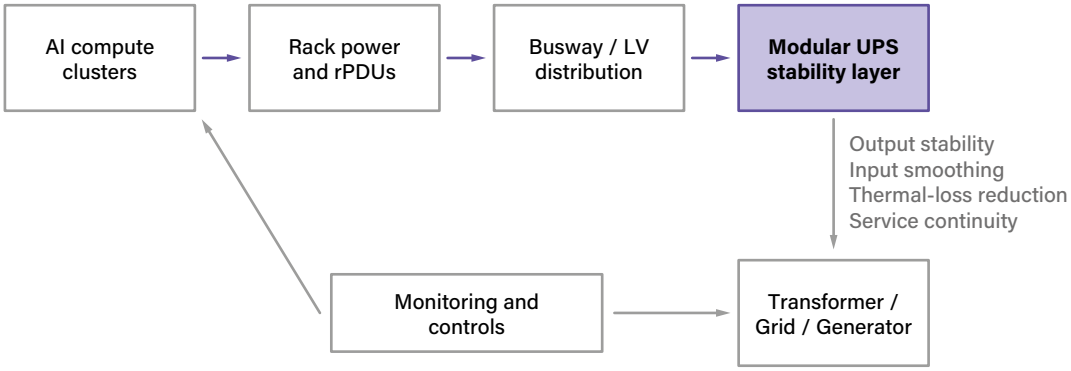


FIGURE 16: EFFECT OF AI CLUSTER TO THE UPSTREAM ELECTRIC INFRASTRUCTURE

This has implications not only for engineering, but also for procurement, commissioning and lifecycle management. The UPS must be specified with AI-relevant validation criteria, monitored with greater granularity and integrated into the broader site strategy for power quality, resiliency, cooling interaction and operational visibility.

A "good UPS" in the AI era is no longer defined only by autonomy time and efficiency under static conditions. It is defined by how effectively it enables the Data Center to host next-generation compute without compromising electrical stability or business continuity.

In this sense, the UPS is evolving from a defensive device into an enabling technology for scalable AI infrastructure.

VALIDATION CRITERION	WHY IT MATTERS IN AI ENVIRONMENTS	EVIDENCE
Output voltage deviation during load steps	GPU server PSUs and downstream power electronics need stable supply during rapid transitions.	Factory or site acceptance test with defined load-step profile
Transient recovery time	Slow recovery can create repeated stress during cyclic AI operation.	Oscilloscope or power-quality recording under representative load
Input current behavior / smoothing	Limits disturbance propagated toward transformers, generators and grid interface.	Input current profile with and without dynamic downstream load
THDi under variable load	Nonlinear behavior can increase upstream power-quality issues.	Harmonic measurement across several load points
Thermal-loss profile	UPS losses become heat and affect cooling energy.	Losses in kW at 25/50/75/100% load
Battery participation frequency	Frequent micro-cycling may influence service life and maintenance strategy.	UPS log data and control-mode explanation

TABLE 6: UPS DESIGN CRITERIA UNDER AI INFRASTRUCTURE PRESSURE

CONCLUSION



Artificial intelligence is redefining the electrical assumptions that have shaped Data Center design for decades.

The challenge is no longer limited to higher rack density or larger installed capacity. AI introduces a new class of electrical load: more concentrated, more dynamic, more synchronized and more demanding in terms of transient response, power quality and infrastructure stability.

These characteristics affect the entire facility, from electrical distribution and protection to cooling and backup power. They require a new design approach based not only on capacity, but also on responsiveness, controllability and resilience under continuously changing operating conditions.

Within this new landscape, the UPS becomes one of the most strategic components of the AI-ready data center. Its role extends far beyond backup autonomy. It must stabilize the energy feed, support recurring step loads, help shield the upstream network from fast load variations and manage battery interaction intelligently under highly dynamic operating conditions.

The UPS is therefore no longer only a protective device for rare outage events. It is an enabling platform for the safe, stable and efficient operation of next-generation compute environments.

Choosing the UPS for AI Applications

Selecting a UPS for AI applications requires a broader, application-driven approach than in conventional Data Center projects.

Nominal power alone is not sufficient. The UPS must be evaluated for its ability to manage rapid and repetitive load variations, maintain output stability during fast transitions, operate reliably with generators and prevent downstream load volatility from affecting the upstream electrical infrastructure.

Efficiency should also be assessed across the actual operating profile. A flatter efficiency curve can reduce both conversion losses and cooling demand under highly dynamic AI workloads.

Battery technology and control strategy are equally critical. Frequent short-duration fluctuations may increase battery activity, making chemistry, cycling capability, thermal behavior and control algorithms key selection criteria. In this context, lithium-ion solutions can offer advantages over VRLA in responsiveness and repeated cycling capability, although the final choice should always be validated against the specific application and UPS architecture.

Procurement should therefore require evidence under representative AI load conditions, including efficiency behavior, overload capability, thermal performance, dynamic-load response and utility-to-generator operation.

More broadly, UPS selection for AI environments should follow a system-level design logic, considering interaction with switchboards, transformers, generators, cooling systems, monitoring platforms and energy storage.

Given the mission-critical nature of AI data centers, UPS systems must be designed to deliver the highest levels of reliability, fault tolerance and operational continuity, while minimizing the impact of maintenance and service activities. A modular architecture with front-access serviceability can support redundancy and scalability, while contributing to high MTBF and low MTTR. This design approach enables faster and more efficient maintenance, reduces operational risk and ultimately increases overall system availability.

IN BRIEF

- Size for dynamics, not only for kW. Verify tolerance to recurring step loads and fast transitions, not only steady-state capacity.
- Check stability under real AI load profiles. Request test evidence under cyclic, synchronized and rapidly varying operating conditions.
- Assess efficiency across the full operating range. Do not focus only on peak efficiency; evaluate efficiency according to standards such as IEC 62040-3 and the European Code of Conduct for UPS to support more realistic OPEX and TCO comparisons.
- Evaluate battery strategy carefully. Consider chemistry, micro-cycling behavior, lifecycle expectations and the control algorithms used to support dynamic loads. In most cases, the UPS manufacturer should validate the battery solution and recommend the most suitable models for its UPS architecture.
- Consider upstream compatibility. Verify interaction with generators, transformers and constrained grids, especially where AI clusters create sharp ramps and synchronized load variations.
- Modular UPS Architecture, with redundancy and scalability and low MTTR
- Require monitoring-ready architecture. Select solutions that support detailed energy, power quality, module-level diagnostics and operational visibility during commissioning and throughout the lifecycle of the facility.

EVALUATION AREA	QUESTION TO ASK	PREFERRED EVIDENCE	RISK IF IGNORED
Load profile	How fast and how often can the load change?	Project telemetry or validated load-step test	UPS appears sized but behaves poorly under dynamics
Efficiency model	What weighting method reflects expected operation?	Full efficiency curve and weighted calculation	OPEX and heat losses underestimated
Thermal behavior	How much UPS loss becomes room heat?	Loss model at expected load points	Cooling burden and electrical-room temperature underestimated
Redundancy and modularity	How does each module behave under N, N+1 or parallel states?	Architecture diagram and service procedure	Growth creates redesign or downtime exposure
Serviceability	Can modules be accessed and replaced safely?	Maintenance method statement and access review	Downtime risk and operational friction increase
Upstream source compatibility	How does the UPS behave with generators and source transitions?	Generator test data and commissioning plan	Instability during transitions or ramp events
Monitoring	What data is visible during commissioning and operation?	Module-level diagnostics, power quality and energy data	Dynamic issues remain invisible until incidents occur

TABLE 7: DECISION MATRIX FOR CONSULTANTS AND OPERATORS

Keor FLEX UPS: The Legrand UPS Solution

Keor FLEX UPS by Legrand has been specifically engineered to address the demanding power protection requirements of next-generation AI data centers. Built on advanced Silicon Carbide (SiC) technology, Keor FLEX delivers up to 98.6% efficiency in Online Double Conversion (VFI) mode and maintaining high efficiency across a wide operating range—from 25% to 100% load. This results in a certified weighted efficiency of 98.35%, with all performance figures independently validated by accredited third-party laboratories.

For AI-driven environments, where power demand can fluctuate rapidly and unpredictably, this level of efficiency translates directly into reduced energy consumption, lower cooling requirements and optimized thermal management. The outcome is a measurable reduction in total operating costs while supporting sustainable Data Center operations.



FIGURE 18: KEOR FLEX UPS



FIGURE 19: KEOR FLEX POWER MODULE 100KW

The system's modular architecture, based on 100 kW power modules, provides outstanding scalability and operational flexibility. Organizations can configure the UPS according to current capacity requirements while preserving the ability to expand seamlessly up to 1.2 MW in a single system and up to 4.8MW with units in parallel. In addition, Keor FLEX supports N+X redundancy configurations, ensuring high availability and resilience for mission-critical applications.

Serviceability has been designed as a core feature. With full front access and hot-swappable power modules, maintenance activities can be performed rapidly and safely, minimizing downtime and maximizing system availability. This approach significantly simplifies both routine servicing and future system upgrades.

To further enhance operational reliability, Keor FLEX incorporates advanced predictive diagnostics and monitoring capabilities. Continuous system analytics help identify early signs of component ageing, optimize operating parameters and support proactive maintenance strategies throughout the UPS lifecycle. This predictive approach contributes to greater reliability while reducing the risk of unexpected outages.

AI data centers place unique demands on power infrastructure due to the highly dynamic and impulsive nature of GPU and accelerator-based workloads. Thanks to its robust power electronics and advanced control architecture, Keor FLEX can handle these rapid load variations while maintaining the highest level of power quality, compliant with IEC/EN 62040-3 Class VFI SS 111 requirements. The result is stable, uninterrupted power delivery that safeguards sensitive AI computing resources and maximizes operational performance.

Beyond protecting downstream equipment, Keor FLEX also helps mitigate the impact of AI workloads on the upstream electrical infrastructure. Through a dedicated operating mode leveraging lithium-ion battery technology, the UPS can absorb and compensate for rapid load transients by supplying short-duration power peaks through controlled battery charge and discharge cycles. This approach smooths demand fluctuations seen by the utility grid, while preserving battery health and ensuring full backup energy availability in the event of a mains failure.

A key differentiator of the Keor FLEX solution is the ability to validate performance under realistic AI operating conditions. Through Legrand's dedicated factory testing laboratories, equipped with proprietary AI load simulators, customers can witness system performance under customized test scenarios tailored to their specific application requirements. This provides a rigorous and transparent demonstration of efficiency, load handling capability, power quality and resilience before deployment.



FIGURE 20: AI LOAD ABSORPTION SET UP



FIGURE 21: AI LOAD SIMULATOR EQUIPMENT

Keor FLEX combines ultra-high efficiency, AI-ready load management, modular scalability, predictive intelligence, serviceability and advanced grid-support capabilities in a single platform. These characteristics make it suitable to protect modern AI data centers, delivering the performance, resilience and operational efficiency required to support the next generation of high-density computing environments.







legrandgroup.com

youtube.com/user/legrand

linkedin.com/company/legrand

x.com/Legrand

World Headquarters
and International Department
87045 Limoges Cedex - France
Tel: +33(0)5 55 06 87 87

legrand®

EXB26101 ©Legrand - December 2026