

INTELLIGENCE UPDATE

Electrical considerations with large AI compute



Daniel Bizo

30 Jun 2025

The training of large generative AI models is a special case of high-performance computing (HPC) workloads. This is not simply due to the reliance on GPUs — numerous engineering and scientific research computations already use GPUs as standard. Neither is it about the power density or the liquid cooling of AI hardware, as large HPC systems are already extremely dense and use liquid cooling. Instead, what makes AI compute special is its runtime behavior: when training transformer-based models, large compute clusters can create step load-related power quality issues for power distribution systems in data center facilities. A previous Intelligence report offers an overview of the underlying hardware-software mechanisms (see [Erratic power profiles of AI clusters: the root causes](#)).

The scale of the power fluctuations makes this phenomenon unusual and problematic. The vast number of generic servers found in most data centers collectively produce a relatively steady electrical load — even if individual servers experience sudden changes in power usage, they are discordant. In contrast, the power use of compute nodes in AI training clusters moves in near unison.

Even compared with most other HPC clusters, AI training clusters exhibit larger power swings. This is due to an interplay between transformer-based neural network architectures and compute hardware, which creates frequent spikes and falls (every second or two) in power demand. These fluctuations correspond to the computational steps in the training processes, exacerbated by an aggressive pursuit of peak performance typical in modern silicon.

Powerful fluctuations

The scope of the resulting step changes in power will depend on the size and configuration of the compute cluster, as well as operational factors such as AI server performance and power management settings. Uptime Intelligence estimates that in worst-case scenarios, the difference between the low and high points of power draw during training program execution can exceed 100% on a system level (the load doubles almost instantaneously, within milliseconds) for some configurations.

These extremes occur every few seconds, whenever a batch of weights and biases is loaded on

GPUs and the training begins. This is often accompanied by a massive spike in current, produced by power excursion events as GPUs overshoot their thermal design power rating (TDP) to opportunistically exploit any extra thermal and power delivery budget following a phase of lower transistor activity. In short, power spikes are made possible by intermittent lulls.

This behavior is common in modern compute silicon, including in personal devices and generic servers. Still, it is only with large AI compute clusters that these fluctuations across dozens or hundreds of servers move almost synchronously.

Even in moderately sized clusters with just a few dozen racks, this can result in sudden, millisecond-speed changes in AC power — ranging from several hundred kilowatts to even a few megawatts. If there are no other substantial loads present in the electrical mix to dampen these fluctuations, these step changes may stress capacity components in the power distribution systems. They may also cause power quality issues such as voltage sags and swells, or significant harmonics and sub-synchronous oscillations that distort the sinusoidal waveforms in AC power systems.

Based on several discussions with and disclosures by major electrical equipment manufacturers — including ABB, Eaton, Schneider Electric, Siemens and Vertiv — there is a general consensus that modern power distribution equipment is expected to be able to handle AI power fluctuations, as long as they remain within the rated load.

IT system capacity redefined

The issue of AI step loads appears to center on equipment capacity and the need to avoid frequent overloads. Standard capacity planning practices often start with the nameplate power of installed IT hardware, then derate it to estimate the expected actual power. This adjustment can reduce the total nameplate power by 25% to 50% across all IT loads when accounting for the diversity of workloads — since they do not act in unison — and also for the fact that most software rarely pushes the IT hardware close to its rated power.

In comparison, AI training systems can show extreme behavior. Larger AI compute clusters have the potential to draw what is similar to an inrush current (rapid change of currents, often denoted by *high di/dt*) that exceed the IT system's sustained maximum power rating.

Normally, overloads would not pose a problem for modern power distribution. All electrical components and systems have specified overload ratings to handle transient events (e.g., current surges during the startup of IT hardware or other equipment) and are designed and tested accordingly. However, if power distribution components are sized closely to the rated capacity of the AI compute load, these transient overloads could happen millions of times per year in the worst cases — components are not tested for regularly repeated overloads. Over time, this can lead to electromechanical stress, thermal stress and gradual overheating (heat-up is faster than cool-off) — potentially resulting in component failure.

This brings the definition of capacity to the forefront of AI compute step loads. Establishing the

repeated peak power of a single GPU-server node is already a non-trivial effort — it requires running a variety of computationally intensive codes and setting up a high-precision power monitor. However, predicting how a specific compute cluster spanning several racks and potentially hundreds or even thousands of GPUs will behave during a training run is difficult to ascertain ahead of deployment.

The expected power profile also depends on server configurations, such as power supply redundancy level, cooling mode and GPU generations. For example, in a typical AI system from the 2022-2024 generation, power fluctuations can reach up to 4 kW per 8-GPU server node, or 16 kW per rack when populated with four nodes, according to Uptime estimates. Even so, the likelihood of exceeding the rack power rating of around 41 kW is relatively low. Any overshoot is likely to be minor, as these systems are mostly air-cooled hardware designed to meet ASHRAE Class A2 specifications — allowed to operate in environments up to 35°C (95°F). In practice, most facilities supply much cooler air, making system fans cycle less intensely.

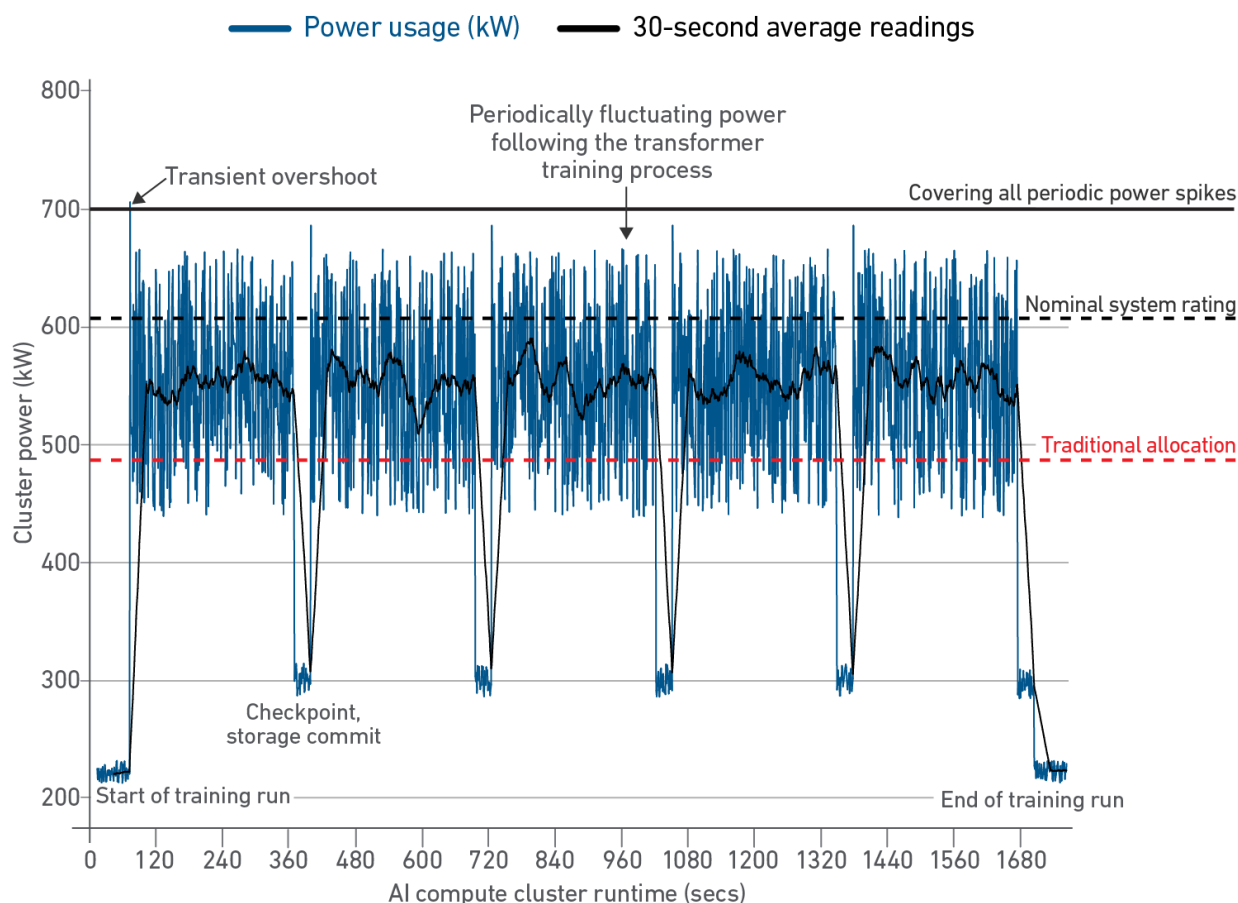
However, with recently launched systems, the issue is further exacerbated as GPUs account for a larger share of the power budget, not only because they use more power (in excess of 1 kW per GPU module) but also because these systems are more likely to use direct liquid cooling (DLC). Liquid cooling reduces system fan power, thereby reducing the stable load of server power. It also has better thermal performance, which helps the silicon to accumulate extra thermal budget for power excursions.

IT hardware specifications and information shared with Uptime by power equipment vendors indicate that in the worst cases, load swings can reach 150%, with a potential for overshoots exceeding 10% above the system's power specification. In the case of the rack-scale systems based on Nvidia's GB200 NVL72 architecture, sudden power climbs from around 60 kW and 70 kW to more than 150 kW per rack can occur.

This compares to a maximum power specification of 132 kW, which means that, under worst-case assumptions, repeated overloads can amount to as much as 20% in instantaneous power, Uptime estimates. This warrants extra care regarding circuit sizing (including breakers, tap-off units and placements, busways and other conductors) to avoid overheating and related reliability issues.

Figure 1 illustrates the power pattern of a GPU-based compute cluster running a transformer-based model training workload. Based on hardware specifications and real-world power data disclosed to Uptime Intelligence, we algorithmically mimicked the behavior of a compute cluster comprising four Nvidia GB200 NVL72 racks and four non-compute racks. It demonstrates the expected power fluctuations during these training clusters and underscores the need to rethink capacity planning compared with traditional, generic IT loads. Even though the average power stays below the power rating of the cluster, peak fluctuations can exceed it. While this estimates a relatively small cluster with 288 GPUs, a larger cluster would exhibit similar behavior at the megawatt scale.

Figure 1 Power profile of a GPU-based training cluster (algorithmic not real-world data)



UPTIME INSTITUTE 2025

uptime
INTELLIGENCE

In electrical terms, no multi-rack workload is perfectly synchronous, while the presence of other loads will help smooth out the edges of fluctuations further. When including non-compute ancillary loads in the cluster — such as storage systems, networks and CDUs (which also require UPS power) — a lower safety margin above the nominal rating (e.g., 10% to 15%) appears sufficient to cover any regular peaks over the nominal system power specifications, even with the latest AI hardware.

Current mitigation options

There are several factors that data center operators may want to consider when deploying compute clusters dedicated to training large, transformer-based AI models. Currently, data center operators have a limited toolkit to fully handle large power fluctuations in a power distribution system, particularly when it comes to not passing them on to the source in their full extent. However, in collaboration with the IT infrastructure team/tenant, it should be possible to minimize fluctuations:

- **Mix with diverse IT loads, share generators.** The best first option is to integrate AI training compute with other, diverse IT loads in a shared power infrastructure. This helps to diminish the effects of power fluctuations, particularly on generator sets. For dedicated AI training data center infrastructure installations, this may not be an option for power distribution. However, sharing engine generators will go a long way to dampen the effects of AI power fluctuations. Among power equipment, engine generator sets will be the most stressed if exposed

to the full extent of the fluctuations seen in a large, dedicated AI training infrastructure. Even if correctly sized for the peak load, generators may struggle with large and fast fluctuations — for example, the total facility load stepping from 45% to 50% of design capacity to 80% to 85% within a second, then dropping back to 45% to 50% after two seconds, on repeat. Such fluctuation cycles may be close to what the engines can handle, at the expense of reduced expected life or outright failure.

- **Select UPS configurations to minimize power quality issues, overload.** Even if a smaller frame can handle the fluctuations, according to the vendors, larger systems will carry more capacitance to help absorb the worst of the fluctuations, maintaining voltage and frequency within performance specifications. An additional measure is to use a higher capacity redundancy configuration, for example, by opting for N+2. This allows for UPS maintenance while avoiding any repeated overloads on the operational UPS systems, some of which might hit the battery energy storage system.
- **Use server performance/power management tools.** Power and performance management of hardware remain largely underused, despite their ability to not only improve IT power efficiency but also contribute to the overall performance of the data center infrastructure. Even though AI compute clusters feature some exotic interconnect subsystems, they are essentially standard servers using standard hardware and software. This means there are a variety of levers to manage the peaks in their power and performance levels, such as power capping, turning off boost clocks, limiting performance states, or even setting lower temperature limits. To address the low end of fluctuations, switching off server energy-saving modes — such as silicon sleep states (known as C-states in CPU parlance) — can help raise the IT hardware's power floor. A more advanced technique involves limiting the rate of power change (including on the way down). This feature, called "power smoothing", is available through Nvidia's System Management Interface on the latest generation of Blackwell GPUs.

Electrical equipment manufacturers are investigating the merits of additional rapid discharge/recharge energy storage and updated controls to UPS units with the aim of shielding the power source from fluctuations. These approaches include super capacitors, advanced battery chemistries or even flywheels that can tolerate frequent, short duration but high-powered discharge and recharge cycles. Next-generation AI compute systems may also include more capacitance and energy storage to limit fluctuations on the data center power system. Ultimately, it is often best to address an issue at its root (in this case the IT hardware and software) rather than treat the symptoms, although these may lie outside the control of data center facilities teams.

The Uptime Intelligence View

Most of the time, data center operators need not be overly concerned with the power profile of the IT hardware or the specifics of workloads running on them — rack density estimates were typically overblown to begin with, and overall capacity utilization tends to stay well below 100% in any case. Even so, safety margins, which are expensive, could be thin. However, training large transformer models is different. The specialized compute hardware can be extremely dense, creates large power swings, and is capable of producing frequent power surges that are close to or even above its hardware power rating. This will force data center operators to

reconsider their approach to both capacity planning and safety margins across their infrastructure.

The following Uptime Institute experts were consulted for this report:

Chris Brown, Chief Technical Officer, Uptime Institute

ABOUT THE AUTHOR



Daniel Bizo

Over the past 15 years, Daniel has covered the business and technology of enterprise IT and infrastructure in various roles, including industry analyst and advisor. His research includes sustainability, operations, and energy efficiency within the data center, on topics like emerging battery technologies, thermal operation guidelines, and processor chip technology.

[**dbizo@uptimeinstitute.com**](mailto:dbizo@uptimeinstitute.com)

About Uptime Institute

Uptime Institute is the Global Digital Infrastructure Authority. Its Tier Standard is the IT industry's most trusted and adopted global standard for the proper design, construction, and operation of data centers – the backbone of the digital economy. For over 25 years, the company has served as the standard for data center reliability, sustainability, and efficiency, providing customers assurance that their digital infrastructure can perform at a level that is consistent with their business needs across a wide array of operating conditions.

With its data center Tier Standard & Certifications, Management & Operations reviews, broad range of related risk and performance assessments, and accredited educational curriculum completed by over 10,000 data center professionals, Uptime Institute has helped thousands of companies, in over 100 countries to optimize critical IT assets while managing costs, resources, and efficiency.