A Device-Centric Approach to AI Infrastructure Efficiency and Reliability

Technical Guide Commissioned by AMI

February 2025

Author

Rami Radi: Sr. Product Manager and Solution Architect

Table of Contents

Copyright © 2025 AMI | Privileged and All Rights Reserved.

Introduction

Over the past decade, the greatest advancements in data center energy efficiency have occurred in the underlying infrastructure (cooling, UPS, etc..). This segment is crucial because it accounts for roughly 38% of total energy consumption. However, many data center operators have not focused as closely on the IT equipment itself, even though that equipment accounts for nearly two-thirds of the data center's energy consumption (Energy Star, 2024).

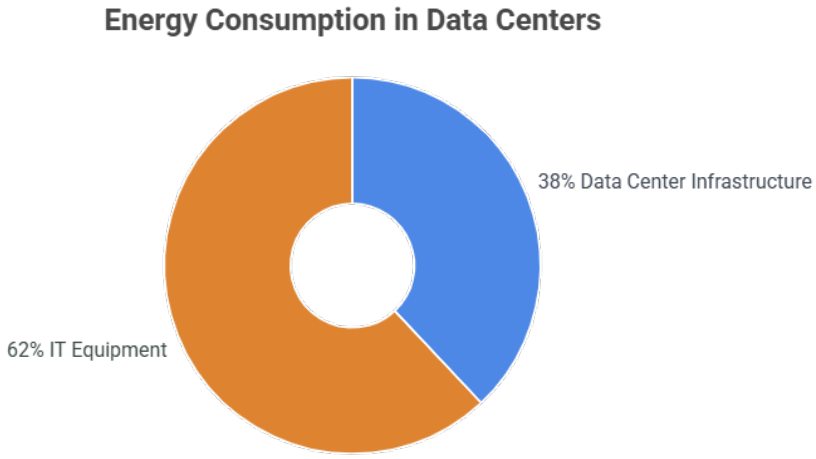

Figure 1: 62% of data center energy consumption is spent on IT Equipment (Energy Star, 2024)

IT equipment energy consumption will continue to grow exponentially as data centers introduce denser, GPU-based AI servers that can consume up to 10X power than traditional servers (Goldman Sachs, 2024). This growth also complicates health monitoring, thermal management, utilization metrics, firmware management, and proactive failure mitigation. These challenges are particularly acute since the dense servers may come from multiple vendors and span diverse generations.

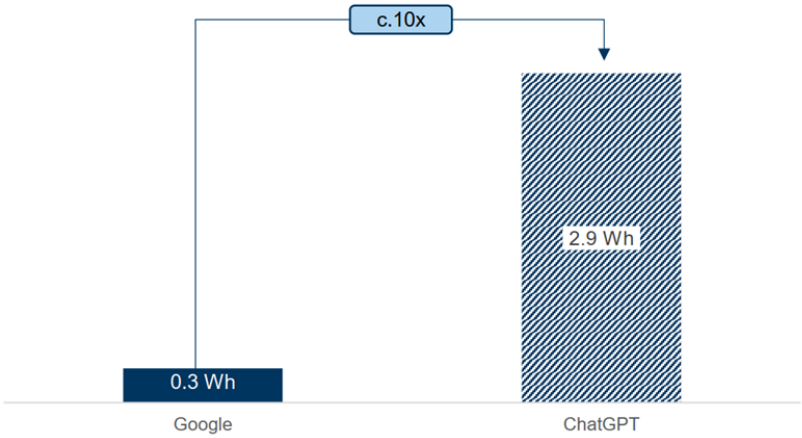

Figure 2: ChatGPT vs. Google power consumption per query/search (Wh) (Goldman Sachs, 2024)

Copyright © 2025 AMI | Privileged and All Rights Reserved.

While Data Center Infrastructure Management (DCIM) tools typically focus on facility-wide metrics, they often lack the granular, per-device insights needed to address localized issues in high-density heterogeneous AI environments. Moreover, single-vendor server management solutions fall short in mixed hardware deployments. Finally, agent-based approaches can introduce both performance overhead and security risks.

By embracing a device, architecture, and vendor-agnostic IT-centric solution for data center management, operators can tackle the challenges of modern AI data centers at the component, server, and rack levels.

Power Management at the Device Level

High-performance compute nodes can draw extraordinary power—occasionally exceeding the combined wattage of several standard servers. This makes monitoring and managing power at the rack and device level more critical, and relying on total facility metrics is not enough. Without this visibility, data centers risk inefficient energy distribution, rising operational costs, and challenges in scaling capacity, performance, and sustainability.

According to Uptime Institute's Global Data Center Survey for 2024, nearly half of respondents primarily worked with facilities over 11 years old (Uptime Institute, 2024).

Moreover, the 2024 Report on U.S. Data Center Energy Use produced by Lawrence Berkeley National Laboratory (LBNL) indicates that while AI workloads are growing, conventional servers still represent a significant portion of data center infrastructure. The report also illustrates that while AI servers operate at 80-90% utilization, non-AI servers often run below 60% (LBNL, 2024).

This means that identifying, consolidating, and optimizing underutilized conventional servers remains essential for improving power distribution and resource allocation. This not only reduces energy waste but also improves space usage and overall operational performance.

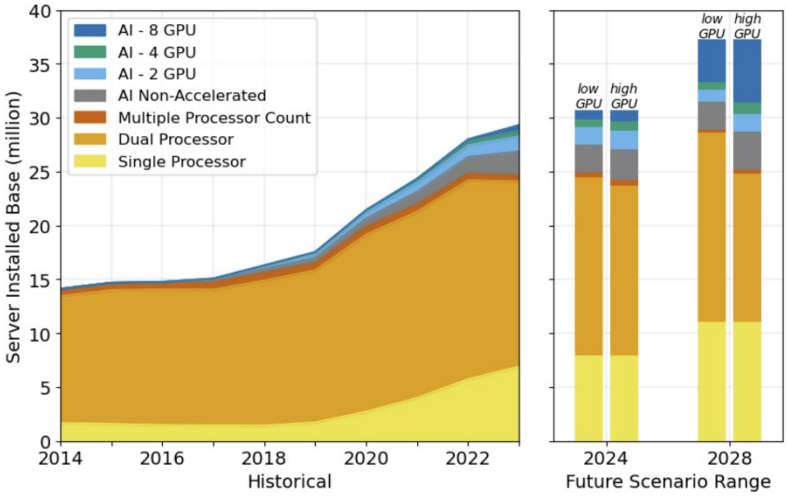

Figure 3. Total server installed base for 2014–2028 with higher bound shipments (left). Adjusted installed base with lower bound GPU shipments (right). (LBNL, 2024).

Copyright © 2025 AMI | Privileged and All Rights Reserved.

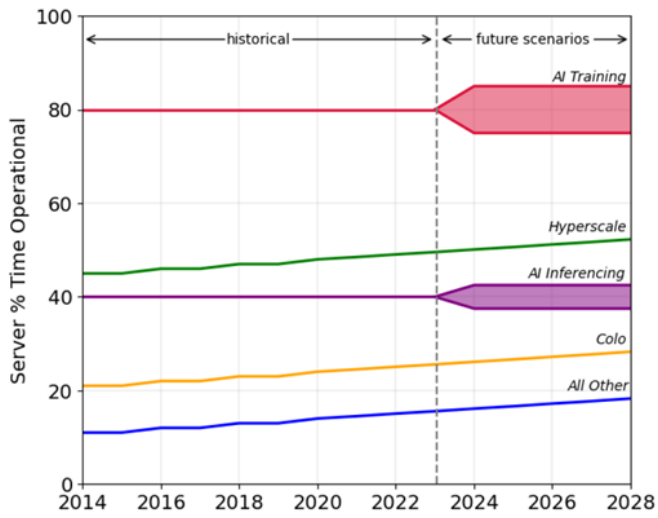

Figure 4: The average operational time of servers for non-AI workloads is often less than 60% (LBNL, 2024).

Resource utilization affects power efficiency in complex ways: a server may draw significant power even when not fully utilized, and tasks requiring 75% CPU utilization may consume nearly as much power as those requiring 100% utilization (NREL, 2025)

This means that power policies can be used to carefully control and reduce power consumption on a server level without impacting performance. When scaled, this can yield significant data center wide savings as shown in the AMI conducted experiment illustrated in Figure 5.

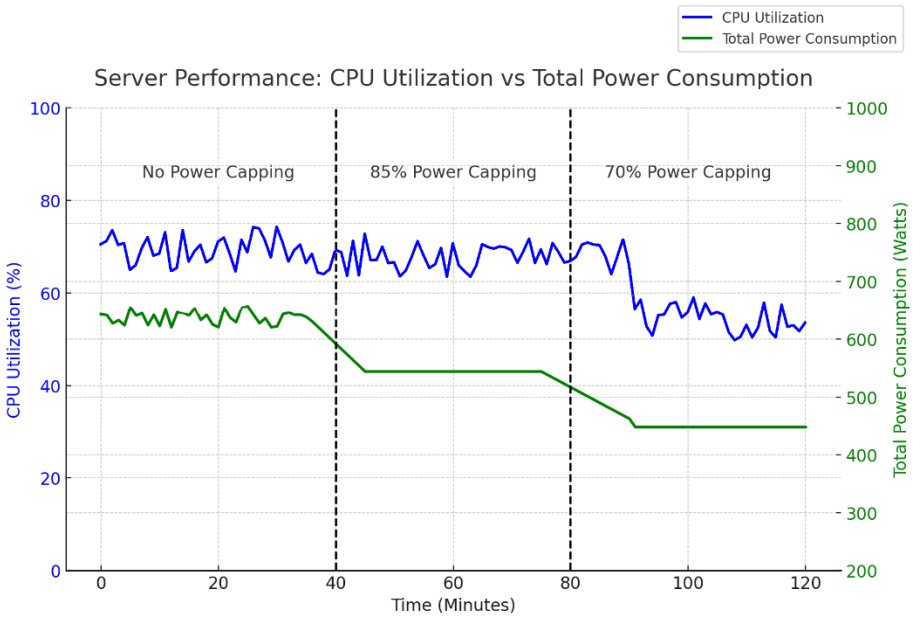

Figure 5: A case when power policies can reduce power by 15% with negligible performance impact for a CPU-bound workload that is not fully utilized.

Copyright © 2025 AMI | Privileged and All Rights Reserved.

Liquid Cooling Oversight

Cooling systems account for 30–40% of total data center energy consumption (EPRI, 2024). As AI racks surpass 40 kW, traditional air-cooling methods become inadequate, making direct-to-chip and immersion cooling essential for managing extreme power densities. However, these advanced cooling solutions introduce new challenges, particularly in monitoring coolant flow rates, pump performance, pressure levels, and leak detection.

Granular visibility into per-device cooling data is critical for detecting issues such as small leaks, declining pump efficiency, or pressure imbalances before they escalate into major failures. Unlike traditional air-cooled systems, where temperature fluctuations can often be managed with redundant airflow solutions, liquid cooling failures can have immediate and severe consequences if not proactively monitored.

Even in legacy data centers, real-time monitoring can guide retrofits, helping operators balance cooling demand without requiring expensive infrastructure overhauls. This is especially important as more AI and HPC infrastructure are integrated into existing data center environments.

The risks of poor liquid cooling oversight are significant. A single cooling system water pump failure can lead to leaks and can result in short circuits, hardware damage, corrosion, and even catastrophic downtime, making real-time health monitoring, analytics, and effective Cooling Distribution Unit (CDU) management essential.

Preventing Catastrophic Component Failures

Reliability is paramount in AI and HPC environments. A single malfunctioning accelerator in a densely populated rack can derail performance-critical workloads. According to a study by Meta, during a 54-day Llama 3, 405-billion-parameter model training run, over half of the unexpected interruptions recorded were caused by issues with GPUs or their onboard HBM3 memory. The study concludes that the GPU annualized failure rates can reach about 9% (Meta, 2024).

Out-of-band, device-level monitoring uncovers early warning signs for each server—like erratic temperature or voltage readings—that might remain invisible to agent-based or in-band tools. And by centrally managing firmware versions across heterogeneous hardware, operators also curb security risks and performance mismatches. Combined, these measures reduce unplanned downtime and the associated maintenance bills.

Copyright © 2025 AMI | Privileged and All Rights Reserved.

Component	Category	Interruption Count	% of Interruptions
Faulty GPU	GPU	148	30.1%
GPU HBM3 Memory	GPU	72	17.2%
Software Bug	Dependency	54	12.9%
Network Switch/Cable	Network	35	8.4%
Host Maintenance	Unplanned Maintenance	32	7.6%
GPU SRAM Memory	GPU	19	4.5%
GPU System Processor	GPU	17	4.1%
NIC	Host	7	1.7%
NCCL Watchdog Timeouts	Unknown	7	1.7%
Silent Data Corruption	GPU	6	1.4%
GPU Thermal Interface + Sensor	GPU	6	1.4%
SSD	Host	3	0.7%
Power Supply	Host	3	0.7%
Server Chassis	Host	2	0.5%
IO Expansion Board	Host	2	0.5%
Dependency	Dependency	2	0.5%
CPU	Host	2	0.5%
System Memory	Host	2	0.5%

Figure 6: A study indicating that during a 54-day period of Llama 3 405B pre-training, about 78% of unexpected interruptions were attributed to confirmed or suspected hardware issues. (Meta, 2024)

Sustainability and Regulatory Compliance

With increasing legislative and market pressures, data centers must improve efficiency and report carbon emissions under regulations like the EU Energy Efficiency Directive (EED). The directive mandates tracking energy usage, Power Usage Effectiveness (PUE), water footprint, and carbon impact, making device-level power and thermal monitoring essential for compliance (EU Energy Efficiency Directive).

Real-time per-node energy monitoring allows operators to optimize cooling strategies, avoiding unnecessary overcooling while maintaining critical workloads. This approach helps reduce greenhouse gas emissions while improving overall energy efficiency.

Beyond compliance, thermal stability directly impacts hardware longevity. Unchecked thermal hotspots accelerate component wear, increase failure rates, and shorten server lifespans, leading to higher replacement costs and downtime. Effective power and cooling analytics help maintain stable operating conditions, improving reliability and sustainability (LDP Associates, 2024).

As the industry moves toward carbon neutrality, tracking PUE, Carbon Usage Effectiveness (CUE), and greenhouse gas emissions is becoming integral to sustainability efforts. Implementing granular energy monitoring and intelligent power distribution aligns with regulatory goals while strengthening competitive positioning.

By integrating precise power and thermal management, data centers can meet regulatory requirements, enhance operational efficiency, and reduce long-term costs—ensuring a sustainable and resilient infrastructure for the future.

Copyright © 2025 AMI | Privileged and All Rights Reserved.

Conclusion

High-density AI deployments need more than just facility-wide dashboards and single-vendor server tools to remain reliable, cost-effective, and compliant with modern efficiency standards. Device and component-level management of power, cooling, and hardware health is indispensable for detecting localized issues—ranging from runaway GPU wattage to faulty coolant pumps—before they escalate into major outages.

By choosing a solution that spans architectures, vendors, and device types, such as AMI® Data Center Manager, operators can strengthen both operational resilience and long-term ROI in an ever-evolving AI landscape.

For example, in an environment with 5,000 servers, adopting AMI® DCM for, can yield $2.2M in annual savings as illustrated here.

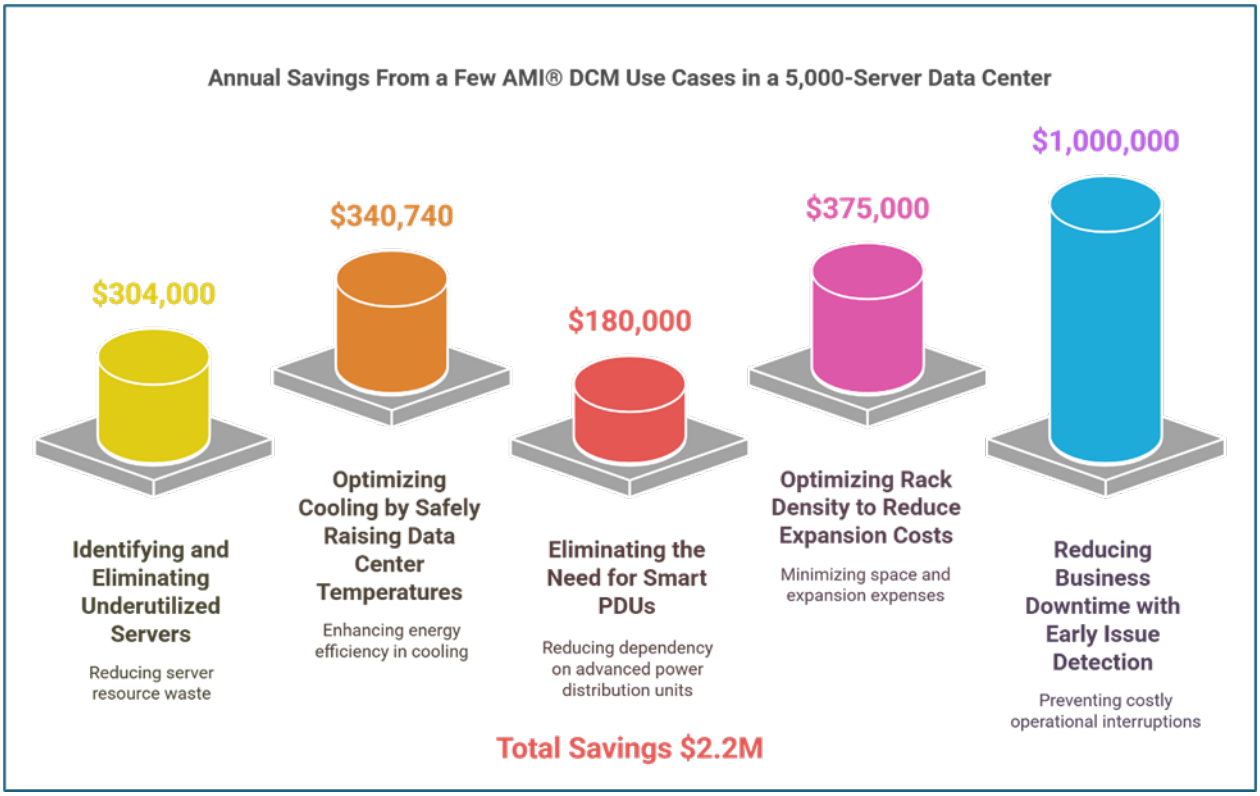

Figure 7: Adopting AMI® DCM can yield $2.2M in annual savings in a Data Center of 5000 servers

For a deeper dive into next-generation AI cluster management, **download the whitepaper** A Holistic Approach to Managing Liquid-Cooled AI Clusters, where **AMI** and **Wiwynn** detail their work on **AMI® DCM v6.0**. For more information about **AMI® DCM**, download the brochure or schedule a demo today.

Copyright © 2025 AMI | Privileged and All Rights Reserved.

References

- Uptime Institute. (2024). *Global Data Center Survey.*
- LBNL. (2024). *2024 United States Data Center Energy Usage Report*
- Energy Star. (2024). *Is Energy Efficiency in Data Centers Still Important?*
- EPRI. (2024). *Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption*
- Meta. (2024). *The Llama 3 Herd of Models*
- Goldman Sachs. (2024). *Generational Growth: AI, data centers and the coming US power demand surge*
- NREL. (2025). *A Beginner's Guide to Power and Energy Measurement and Estimation for Computing and Machine Learning*

About AMI

AMI is Firmware Reimagined for modern computing. As a global leader in Dynamic Firmware for security, orchestration, and manageability solutions, AMI enables the world's compute platforms from on-premises to the cloud to the edge. AMI's industry-leading foundational technology and unwavering customer support have generated lasting partnerships and spurred innovation for some of the most prominent brands in the high-tech industry. For more information, visit www.ami.com.

Copyright © 2025 AMI | Privileged and All Rights Reserved.