

## The Real Challenges of AI (and Why Most Data Centers Aren't Ready)

The conversation around AI is everywhere. Most of it focuses on models, new applications, and what AI might do next. Very little of it talks about what it actually takes to run these workloads at scale.

That's the problem.

AI isn't just another application category. It is a fundamentally different class of computationally intense workloads. When organizations start infusing AI into their existing applications and processes, they quickly discover that it changes how infrastructure must be designed, deployed, and operated.

Power requirements rise sharply. Cooling becomes a primary constraint. Systems grow significantly more complex — and more expensive. At the same time, expectations for performance and time-to-solution only increase.

This is not a minor upgrade. It is a structural shift.

The organizations best prepared for this shift aren't necessarily the ones with the biggest budgets or the newest hardware. They're the ones that have already been operating under similar constraints for years.

High-performance computing environments have long dealt with these exact challenges: running large, complex workloads at scale, optimizing every watt and every cycle, and constantly balancing performance against cost and power limits.

Now those same pressures are appearing across a much wider range of data centers. Enterprise applications are becoming AI-infused. New workloads are emerging that demand HPC-class infrastructure. Many environments are being pushed well beyond what they were originally designed to handle.

The technology itself is rarely the limiting factor. How it is deployed and operated usually is.

## Where the Pressure Shows Up

Power is the first and most visible constraint. Traditional racks that once ran comfortably at 15–20 kW are now being asked to support 40 kW, 60 kW, or more — with rapid fluctuations as AI workloads spin up and down.

Cooling follows close behind. Air-cooled systems that worked reliably for years are hitting hard limits, forcing organizations to seriously evaluate liquid cooling and other alternatives.

Cost is the third pressure point. These systems are not only more powerful, they're significantly more expensive to acquire and run. Underutilized or poorly matched infrastructure is no longer just inefficient — it's financially unsustainable.

## Efficiency Becomes the Differentiator

The real key going forward is not simply adding more hardware. It's using what you already have far more effectively.

That means gaining clear visibility into where power and compute are actually being used (and wasted), matching workloads to the right systems, reducing idle resources, and continuously measuring and adjusting operations.

High-performance computing environments have been doing this for decades. The rest of the industry is now being forced to learn the same lessons — quickly.

This site exists to help bridge that gap.

We examine these issues through a practical **Infrastructure Efficiency Framework** organized around six core areas:

- **Data Center Assessment** — Understanding what you actually have and where inefficiencies hide.
- **Cooling Technology** — Managing the massive new heat loads AI systems generate.
- **Power Infrastructure** — Delivering reliable power efficiently at high densities.

- **Systems OEMs & Integrators** — Choosing the right partners for system-level design.
- **System Components** — Selecting accelerators and other critical technologies wisely.
- **Operational Efficiency** — Maximizing useful work from the resources you have.

Each area plays a vital role in helping data centers adapt to the demands of AI and other high-intensity workloads.

The website itself isn't in line with today's fashions. You won't find flowing pictures of landscapes or abstract art to scroll past. It's designed to be straightforward, with a high meat-to-fluff ratio. Pretty? Not so much. Useful? That's the goal.

Let us know if it is — or isn't.