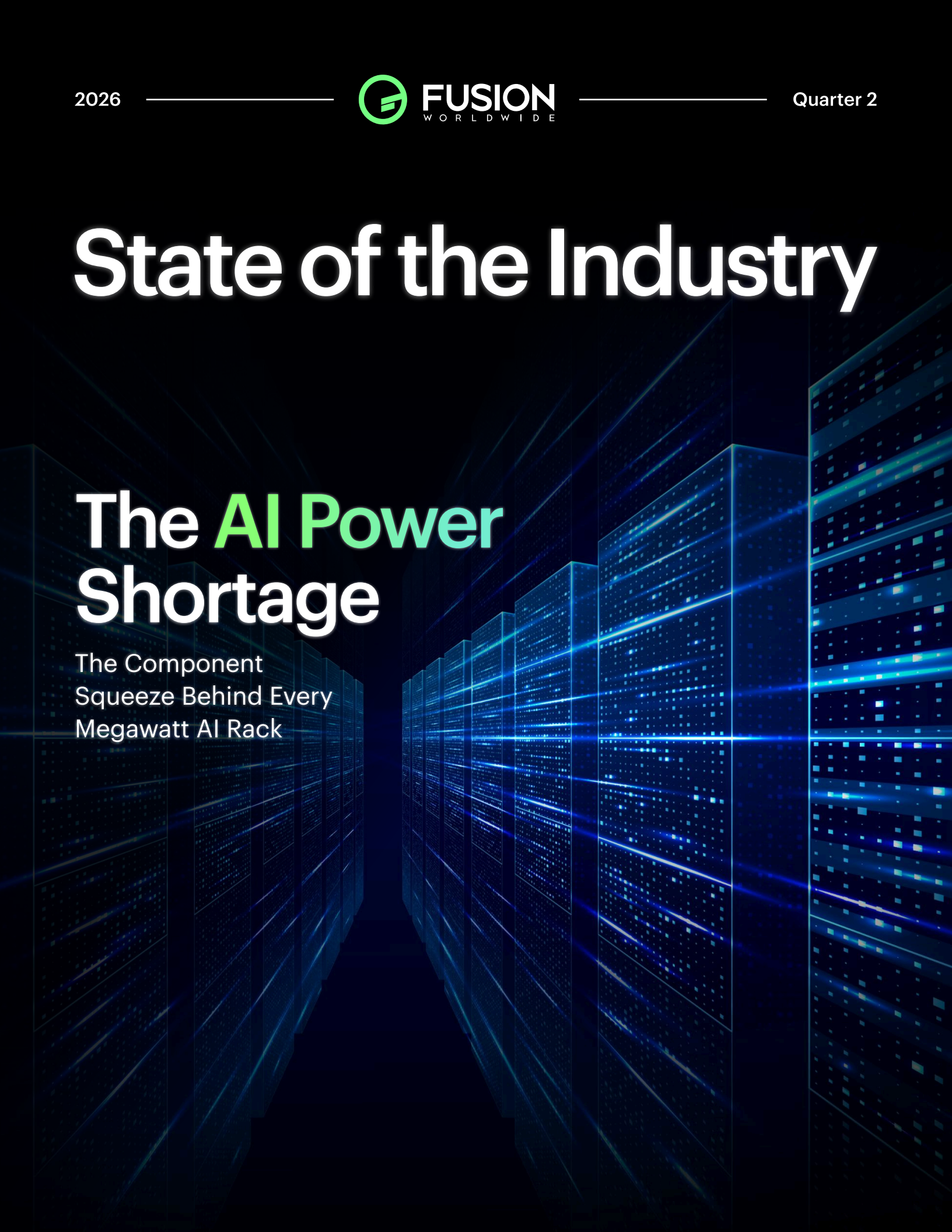
2026 FUSION WORLDWIDE Quarter 2

State of the Industry



The AI Power Shortage

The Component Squeeze Behind Every Megawatt AI Rack

Contents

Power Infrastructure Has Become the Newest Constraint in AI Deployment

For two years, procurement teams building out AI infrastructure focused on securing GPUs, high-bandwidth memory, and networking gear. Every one of those line items can clear on schedule, and the rack still won't come online. Not because of a design flaw. Because a power management IC inside the power distribution unit carries a 40-week lead time nobody had ordered against in time.

As rack power densities push toward the megawatt era, the binding constraint is no longer the chip. It is the power management ICs (PMICs), microcontrollers (MCUs), silicon carbide (SiC) MOSFETs, gallium nitride (GaN) devices, gate drivers, and high-current connectors that get power to it.

The catalyst is physical, not strategic. The average enterprise rack ran 5 to 10 kilowatts before AI, with roughly 8 kW a common design point (International Energy Agency). Nvidia's Hopper-generation DGX H100 pushed that to 30 to 40 kW. Blackwell runs 120 to 132 kW. The incoming Vera Rubin platform is expected to draw 180 to 220 kW, with some architectures projected to exceed 500 kW per rack later this decade (Nvidia).

No power infrastructure built for the last decade of data centers was designed to carry that load. Nvidia's answer is an 800-volt DC architecture that delivers roughly 150% more power through the same copper and eliminates about 200 kilograms of busbar per rack (Nvidia). That is also why the power distribution unit had to be redesigned from a passive metal box into a networked edge-computing platform with embedded controllers, sensors, and DCIM integration. Fusion Worldwide saw this shift play out in person at COMPUTEX 2026: the conversation has moved from the chip itself to everything required to power and cool it.

That redesign is what set off the chain reaction now working its way through the supply chain.

~8kW
Pre-AI Baseline Rack (IEA)

30–40kW
Hopper / DGX H100

120–132kW
Blackwell

180–220kW
Vera Rubin (Incoming)

500+kW
Later this Decade, est.

The Complication

The Constraint Is Not the Chip. It Is What Powers It.

~80%

Top-10 Foundry 8-Inch Utilization, 2025

~90%

Top-10 Foundry 8-Inch Utilization, 2026 (TrendForce)

Every dollar of AI capital expenditure has been chasing the same target: the most advanced silicon money can buy. Three-nanometer logic. The densest HBM stacks. The tightest packaging tolerances available. None of that is where AI infrastructure deployments are getting stuck in 2026.

Intelligent power distribution doesn't run on that silicon at all. It depends on mature-node technology, 28 to 180 nanometer processes, for the 32-bit MCUs, power management ICs (PMICs), gallium nitride (GaN) devices, gate drivers, isolation ICs, and discrete MOSFETs that make it work. Analog and power-management devices gain little from moving to leading-edge nodes; what matters is voltage handling, reliability, and long product life, so mature-node capacity remains the platform of choice even as investment money moves elsewhere.

That is the paradox. While capital pours into advanced packaging, CoWoS capacity, and the HBM supply that's reshaping the memory market, the 8-inch fabs that make PMICs, MCUs, and gate drivers are running tighter than they have in years, and AI-driven demand for power ICs is cited as a key reason foundries can't free up room for other customers (TrendForce).

It is not 2021's blanket shortage. It is narrower, concentrated on the specific mature-node parts that power infrastructure cannot function without, and in practice, harder to see coming.

PMIC, In Plain Terms

A power management IC, sometimes called a power management chip or PMIC semiconductor, regulates, converts, and distributes voltage inside a device. Inside an intelligent PDU, PMICs act as the power management controller for every outlet: they handle per-outlet monitoring, protect against overcurrent, and report back to the facility's DCIM software. No PMIC, no visibility into the rack.

The Short List

Four Components, One Bill of Materials, All Constrained

The risk in a representative intelligent PDU's bill of materials isn't evenly spread. It concentrates in four places, and every one of them is a category procurement has historically overlooked.

Critical | 26-52 Week Lead Time

Silicon Carbide (SiC) MOSFETs

Automotive electrification already claims most SiC wafer capacity for traction inverters. AI infrastructure is now a second, structural demand driver competing for the same lines.

High | 20-45 Week Lead Time

32-Bit Microcontrollers (MCUs)

Renesas, Microchip, STMicroelectronics, and NXP are issuing accelerated EOL and last-time-buy notices on legacy families that many PDU designs single-source.

High | 20-40 Week Lead Time

Power Management ICs (PMICs)

Demand is climbing across AI accelerators, enterprise SSDs, networking gear, and industrial systems, all competing for the same constrained mature-node analog capacity.

Moderate High | 12-24 Week Lead Time

Passives & Connectors

Automotive-grade inductors and high-current connectors face copper and specialized tooling constraints that rarely make headlines but can delay a rack just as effectively.

STMicroelectronics' own Q2 2026 earnings call described the identical collision, unprompted: industrial and general-purpose MCU demand moving from destocking into replenishment and supply tightness, in the same quarter its silicon carbide business swung back to growth on AI datacenter pull.

Renesas told the same story a quarter earlier, tying rising 32-bit MCU demand and higher wafer input at its Naka fab directly to AI datacenter and client-side AI growth. Two public semiconductor companies, on their own earnings calls, describing the same bill of materials this report is describing.

Automotive, Storage, and Consolidation Are Compounding the Squeeze

The allocation squeeze our floor is watching in PMICs and connectors doesn't stop at power infrastructure. If the mature-node squeeze were only about the PDU, it would be a solvable problem: qualify a few extra suppliers, extend a forecast, move on. It is not only about the PDU.

Automakers have claimed the majority of SiC capacity for traction inverters for years, and PDU manufacturers now compete directly with them for the same wafers. Eaton's own Q1 2026 earnings call put a number on how much of the grid that competition now runs through.

Less visibly, the UPS and battery energy storage (BESS) systems that back up the data center are migrating to the same SiC and 32-bit MCU architecture to hit AI-grade efficiency targets. When a hyperscaler orders a new battery array for an AI cluster, it draws on the same component supply required to build the PDUs for that cluster.

Schneider Electric just put a dollar figure on what that looks like: nearly $2.3 billion in new 2026 deals with Switch and Digital Realty for the power modules, cooling, UPS, and switchgear a hyperscale AI buildout requires.

32 GW

U.S. Data Center Capacity Under Construction (Eaton)

~70%

Of That Capacity is AI-Driven

228 GW

Total U.S. Backlog, ~12 Years at 2025 Build Rates

Consolidation is compounding the squeeze. Schneider Electric, Vertiv, Eaton, Legrand, and ABB collectively hold just over half of the global PDU market (Mordor Intelligence), with Schneider Electric and Vertiv separated by roughly a tenth of a percentage point at the top ("Schneider Electric and Vertiv Dominate").

The leaders are signing direct, long-term supply agreements that lock up capacity ahead of everyone still buying on the open market, and integrating vertically into liquid cooling, battery storage, and software-defined power management, so component availability is becoming as important as price. Fusion Worldwide has tracked the same allocation dynamic play out across the broader discrete component market, where tier-one demand closes the door on tier-two buyers.

50%+

Global PDU Market, Top 5 OEMs

75%+

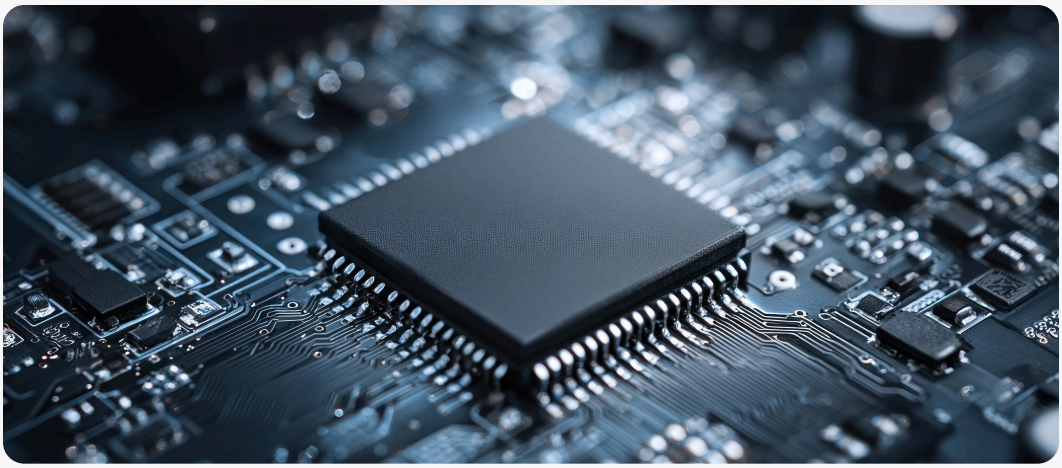
Component Manufacturing in China / Taiwan

~$14.2B

Global PDU Market by 2032

Final assembly is diversifying under a “China Plus One” strategy, with new lines opening in Vietnam, Malaysia, Thailand, India, and Mexico. But more than 75% of power management device manufacturing still sits in mainland China and Taiwan. A PDU assembled in Guadalajara still runs on SiC and MCUs fabricated in Hsinchu.

Moving the last step of assembly does not move the bottleneck upstream. The global data center PDU market is projected to grow from roughly $4.9 billion in 2026 to $14.2 billion by 2032 (MarketsandMarkets), a trajectory that will keep demand for the components behind it climbing well past the current allocation crunch.



What We’re Watching: Power Semiconductor Demand

Across our sourcing network, the real signal isn't rising demand, it's who gets served first. Infineon, Renesas, MPS, Onsemi, and TI are increasingly steering PMIC and power IC output toward AI and hyperscaler accounts, and every other customer is absorbing the lead-time extensions that creates.

Connectors are showing the same pattern. AI isn't just competing for power semiconductor capacity, it's gobbling it, and the customers left waiting are the ones who'll feel this shortage first.

Scenario Outlook

Three Scenarios Through 2028

None of the three paths below put power-component lead times back where they were before 2024. The difference is how much worse it gets before it gets better, and how much runway procurement has to prepare.

Not every analyst is convinced the broader AI semiconductor supercycle holds up under its own weight, and that is a fair debate to have about GPUs and HBM. It has nothing to do with what follows here. This is a tighter, power-specific squeeze, and it is already visible in bill-of-materials lead times today, not in anyone's five-year forecast. Vertiv entered the second quarter of 2026 with a project backlog already above $15 billion on AI-driven data center demand alone, real money on the books before a single one of these scenarios plays out.

Baseline Outlook

65%

Power device deployment lags GPU delivery by 1 to 2 quarters. SiC and MCU lead times hold at 30 to 45 weeks through 2028. Spot premiums run 10 to 15%.

Constrained Scenario

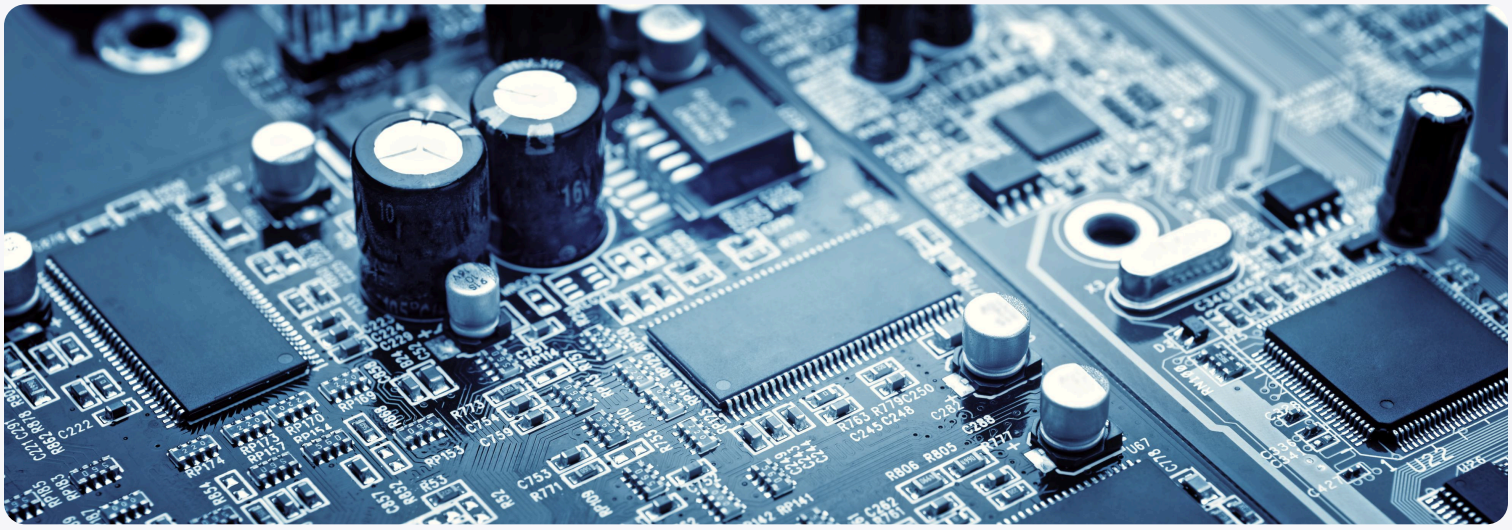
20%

Severe shortages delay AI cluster energization by 6+ months. Lead times exceed 52 weeks under strict allocation.

Capacity Relief

15%

New mature-node capacity arrives in late 2027. Lead times stabilize at 20-26 weeks, still above pre-2024 norms.



What Procurement Leaders Are Doing

The organizations that get this right will be the ones still shipping racks on schedule in 2027. The ones that don't will find a three-dollar microcontroller standing between them and a multi-million-dollar deployment. The difference comes down to four moves. In every one of those scenarios, the deployments that stay on schedule are the ones that got **Out in Front™** of it.

Extend the horizon.

Move forecasting and PO placement out to 52 weeks for power devices, SiC MOSFETs, and 32-bit MCUs. A six-month forecast no longer secures capacity.

What We're Seeing: Buyers are already trying to convert 12-month POs into 18-month blanket orders mid-cycle, and getting told the line closed months ago.

Qualify a second source before you need one.

Single-sourcing any part on the PDU bill of materials, PMICs, inductors, or connectors, is now an unacceptable risk.

What We're Seeing: Every panicked call we get right now traces back to a buyer with exactly one approved vendor on the PMIC line.

Buffer strategic inventory.

The carrying cost of safety stock on wide bandgap semiconductors is small next to the cost of a stalled AI deployment.

What We're Seeing: The accounts holding six months of SiC and MCU safety stock today are not the ones calling us asking for miracles.

Build for agility, not just performance.

Work with engineering to qualify drop-in replacements for high-risk legacy MCUs, and treat an independent distribution partner as part of the sourcing strategy, not a fallback plan.

What We're Seeing: Whoever's still treating an independent distributor as a plan B, instead of a standing part of the sourcing strategy, is explaining a stalled rack to their own customer.

State of the Industry

The Electron Is the New Transistor

The bottleneck has migrated downstream, from the GPU into the power and facility infrastructure that supports it. The organizations that stay on schedule are the ones with visibility three layers down the bill of materials, months before it matters.

Worried there's a PMIC, a MOSFET, or a connector buried in your BOM with a 40-week lead time nobody's watching? Fusion Worldwide helps organizations catch that risk before a three-dollar microcontroller stalls a multi-million-dollar deployment.

Contact our team to evaluate your BOM, or visit fusionww.com.

Americas

Global Headquarters
Boston
+1 617 502 4100
boston@fusionww.com

EMEA

Regional Headquarters
Amsterdam
+31 20 667 6020
amsterdamoffice@fusionww.com

APAC

Regional Headquarters
Singapore
+65 6311 5250
singaporeoffice@fusionww.com

FUSION WORLDWIDE