



## There is No Al Without Power

An interview with Adam White from Infineon on Powering the Next Generation of Al Data Centers



The rapid advancements in artificial intelligence (AI) applications have drastically increased the power demands within data centers, thereby amongst others imposing significant challenges on power grids. To explore how industry leaders are responding to these growing pressures, we sat down for an exclusive interview with Adam White, Division President of Power & Sensor Systems at Infineon Technologies AG.

With an emphasis on the semiconductor technologies that transform power management, this discussion will dive into the industry's comprehensive strategies for improving energy efficiency, reducing total cost of ownership (TCO), and strengthening system robustness in AI data centers. Adam White provides a detailed look at how Infineon is tackling the present and upcoming industry challenges with cutting-edge semiconductor solutions as part of the company's ongoing innovative approach to sustaining the growth of AI technologies.

A key focus will be on Infineon's 'power first' and 'grid to core' principles, which prioritize optimizing power delivery at every stage, from the energy grid to the core of AI data centers, ensuring more efficient and reliable energy flow.



Wevolver Team: Can you provide an overview of your role at Infineon and your journey in the power and sensor systems division?

**Adam White:** "I've been with Infineon for several years, currently serving as the Division President of the Power & Sensor Systems division. My journey began in engineering, where I was deeply involved in research and development, and over the years, I transitioned into roles that broadened my scope to include marketing and sales. Before my current position, I was the Executive Vice President and Chief Marketing Officer of this division, which gave me a holistic view of our market strategies.



Connect with <u>Adam White on LinkedIn</u> to follow his insights and join the conversation on innovative and energy-efficient semiconductor solutions.

My role primarily involves steering the division towards innovative, energy efficient solutions that address the growing needs of the power and sensor systems markets and thereby driving the digitalization and decarbonization of our planet. Given the necessary current global emphasis on energy efficiency, for example in Al applications, our work has never been more critical. We aim to create systems that lead the market in terms of technology and set standards for power efficiency and quality."

Wevolver Team: Given Al's growing importance, can you explain the role that semiconductors play in supporting Al technologies?

**Adam White:** "Al's role in modern technology is transformative, and it fundamentally relies on the capabilities provided by semiconductors. Semiconductors are at the heart of Al; they help to power, collect, process and manage the vast amounts of data that Al systems require to function. This includes everything from basic computations to complex machine learning tasks that enable Al to 'learn' from data.

Our focus at Infineon is to consider the entire ecosystem necessary to support AI functionalities, particularly the power management from grid to core. Power semiconductors ensure that AI data centers operate efficiently, managing the energy flow from the power grid all the way to the processors, the core, which is critical as the energy demands of these systems increase rapidly.



Simply put, there is no Al without power.

Moreover, as AI applications extend their reach from cloud data centers to edge devices, the role of semiconductors diversifies and expands. Edge machine learning (ML) is the process of running machine learning algorithms on computing devices at the periphery of a network (e.g. consumer devices, home / major appliances, factories) to make decisions and predictions as close as possible to the originating source of data rather than in a central environment such as in a cloud computing facility. And by that leveraging advantages of Edge ML like low latency, lower cost and improved security and data privacy for a more seamless, energy-efficient and secure user experience. With Infineon's sensors (XENSIV™ portfolio), microcontrollers like our PSOC™ Edge family and our ML software offering, we cover the entire edge ML workflow. This holistic approach, from power management in data centers, collecting information needed for AI training to data processing, highlights the crucial role of semiconductors in enabling and advancing AI technologies.

Wevolver Team: Infineon emphasizes a 'grid to core' approach in powering Al. Can you explain what this means?

**Adam White:** "Certainly! The 'grid to core' approach is a comprehensive strategy that Infineon has developed to address the power needs of AI from the very beginning of the energy supply chain—starting at the grid—right down to the core components, the processors (e.g. GPUs, TPUs). This strategy is about managing power conversion through various stages efficiently and effectively to meet the demanding requirements of AI data centers.

When we talk about power entering from the grid, it usually comes in at about 220V AC, and within the data center, it undergoes several conversion steps before it powers the AI cores. These steps include converting 220V AC to a lower voltage DC, such as 48V, then stepping down further to 12V or even 6V, and finally to the precise voltage needed by the processors (e.g. GPUs, TPUs), often around 1V DC. Optimizing each stage of the process is vital to minimizing energy loss and maximizing efficiency. By collaborating closely with industry leaders, we're driving the development of next-generation AI power solutions that will redefine the future of artificial intelligence. Infineon's unique position allows us to provide solutions at each step. We don't just supply components; we understand and innovate across the entire power conversion chain, ensuring that each step as well as the complete flow is as efficient as possible. By optimizing how power is handled at each stage, we can significantly reduce the overall power consumption and enhance the sustainability of AI operations. This approach helps save energy, reduces the TCO, increases the application robustness and ensures that the systems can handle the increasingly complex computations required by advanced AI algorithms without compromise."

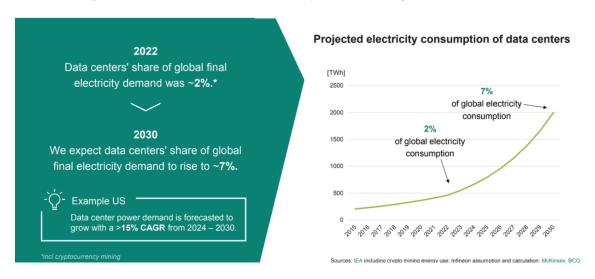
Wevolver Team: Why do we need to have a closer look at the energy efficiency of Al data centers – could you explain what is happening with the rise of ever-smarter generative Als and accelerating compute?



**Adam White:** "Focusing on energy efficiency in Al data centers is crucial for several reasons, particularly as we witness the rapid advancement and increased adoption of generative Al technologies. These sophisticated Al models require an unprecedented level of computational power, which in turn leads to substantial energy consumption. Here's why energy efficiency is becoming a critical concern:

- 1. **Compute Increase:** The computational power required to train state-of-the-art Al models has been doubling approximately every 3.4 months since 2012. This exponential growth in compute requirements directly translates into increased power demands, adding to the overall drain on the grid.
- 2. **Impact on Power Consumption:** With more complex computations, the power needed grows significantly. This isn't just about the electricity used by the processors themselves but also the infrastructure required to support them, including cooling systems and power delivery networks.
- 3. **Global Electricity Consumption:** Data centers account for about 2% of the world's total electricity use, but this is expected to rise to 7% by 2030. This increase is substantial and could become challenging to serve.

Electricity - Substantial increase in compute driven by Al



Source: Infineon Technologies

4. Localized Impact in Data Center Hubs: The impact is even more pronounced in regions dense with data centers. For example, the United States is projected to see Al data centers consume up to 16% of the total electricity supply in 2030. In Ireland, this number could be as high as 32% already in 2026.

Given these dynamics, we must focus on enhancing energy efficiency. This means for example developing more efficient AI algorithms or optimizing data center infrastructure by



implementing innovative and energy efficient power management solutions that significantly reduce power delivery network losses. Addressing these challenges is essential for environmental sustainability and ensuring the economic viability of scaling up AI technologies."

Wevolver Team: You mentioned the 'drain on the grid' caused by Al data centers. Could you elaborate on what this means?

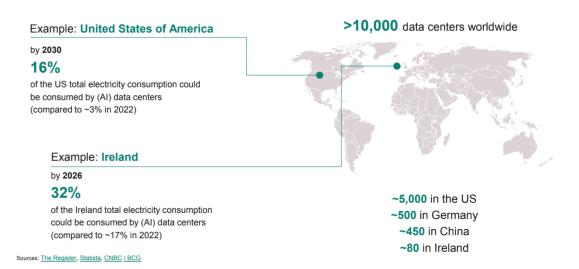
**Adam White:** "The 'drain on the grid' refers to the significant and growing demands that Al data centers place on our electrical infrastructure. This demand is particularly intense in regions known as data center hubs, where the concentration of large-scale data operations leads to a disproportionate consumption of the local power supply. These hubs often face unique challenges, as the high density of data centers can stress the existing power infrastructure, necessitating upgrades or expansions that are costly and time-consuming.

## <Embed: We power Al podcast - why is Al so power hungry | Infineon>

Recognizing the scale of this issue, several countries and regions have implemented legislation to curb the power consumption of data centers. These laws encourage or mandate the use of more energy-efficient technologies and practices to help mitigate the overall impact on the grid.

## Drain on the grid – Data center hubs are especially prone to experience a substantial pressure on their grid





Source: Infineon Technologies

At Infineon, we believe that a key part of the solution is to adopt a 'power first' approach. Traditionally, power considerations in AI have often come after computational needs — essentially, an afterthought. But increasing compute per Watt is only one side of the coin, efficient power consumption being the other. By integrating thoughts of increased energy



efficiency and reduced power consumption right from the design and creation phase of new AI server systems, we can significantly curtail the overall power requirements.

For example, every 1 million next-generation GPUs currently consume, on average, about 0.7 GW of power, which is equivalent to the output of two mid-size coal-fired power plants. With projections indicating that we might expect around 7 million GPU shipments by 2025, the potential cumulative demand could necessitate the equivalent of building approximately 13 new coal-fired power plants just to support this growth. Clearly, this is unsustainable, and by thinking 'power first,' by prioritizing increasing energy efficiency we aim to drastically reduce these numbers and ensure a more sustainable integration of AI technologies into our societal fabric."

Wevolver Team: Also, environmentally speaking, this could pose some challenges, right?

**Adam White:** "Indeed, the environmental challenges posed by expanding AI data centers are significant. As we scale up AI capabilities, we must be acutely aware of the following dimensions:

- Electricity and CO2 Emissions: The massive energy consumption required by AI
  data centers strains our power grids and contributes significantly to CO2 emissions.
  As these facilities consume more power, often sourced from non-renewable energy,
  the carbon footprint of digital operations continues to grow.
- 2. **Water Use:** All data centers require extensive cooling mechanisms to prevent overheating, which often leads to substantial water usage. The water consumption for cooling these massive data centers can rival that of small cities, putting additional pressure on local water resources.
- 3. Electronic Waste (E-waste): The rapid pace of technological advancement means that hardware in AI data centers can quickly become obsolete. This results in significant amounts of e-waste, which, if not properly managed, can lead to environmental pollution due to the toxic substances commonly found in electronic components.

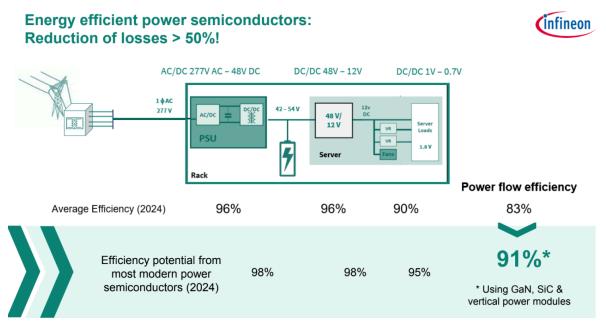
Wevolver Team: We already talked about the substantial power demands in Al data centers, can you quantify the benefits of using state-of-the art power management solutions there?

**Adam White:** "Addressing the power demands in Al data centers through enhanced energy efficiency is critical. We're facing many challenges, but at Infineon, we continuously innovate to provide effective solutions. Let's look at where we start as a semiconductor company:

In an average data center, about 17% of energy is lost as power delivery network losses during various conversion steps. This is a significant figure, but we are actively reducing these losses through innovation and optimization of our semiconductor technologies.



We focus intensely on each step of power conversion within the data center. By utilizing highly efficient power semiconductors, we've managed to approximately halve these power delivery network losses from 17% down to about 9%. This improvement in energy efficiency—from an average power flow efficiency of 83% to now reaching up to 91%—is not just about energy savings. It represents a critical advancement in our ability to manage and mitigate the environmental impact of these facilities and to bring down the TCO of operating AI data centers. By boosting energy efficiency, we can significantly minimize power losses, which in turn leads to a substantial reduction in CO2 emissions. Additionally, this also results in decreased heat waste, ultimately lowering the demand for cooling and conserving precious water resources.



Source: Infineon Technologies

This progress demonstrates our commitment to 'thinking power first.' By optimizing each conversion step with the latest in semiconductor technology, we ensure that AI data centers can meet their growing computational demands more sustainably. This approach helps reduce operational costs and plays a crucial role in our global efforts toward energy conservation and sustainability."

Wevolver Team: Which levers are most promising to increase energy efficiency in Al data centers?

**Adam White:** "At Infineon, we identify several key levers that are most effective in increasing energy efficiency in AI data centers. Here are the main areas we focus on:

1. Advanced Semiconductor Materials: Alongside our high-quality Silicon products, we employ cutting-edge Wide Bandgap (WBG) materials like Silicon Carbide (SiC)



and Gallium Nitride (GaN). These materials are crucial for reducing losses in power conversion processes and can handle higher voltages and temperatures, which are common in high-performance computing environments.

- Optimized Power Architectures: Our approach includes developing innovative
  power architectures that optimize the entire energy flow from the grid to the core.
  This involves improving power density and reducing power losses at every stage of
  the power delivery network, thus enhancing the overall power flow efficiency.
- 3. Advanced Packaging Techniques: We also focus on advanced packaging solutions that enhance thermal management and reduce resistive losses. Techniques like chip embedding and integrated magnetics help minimize the physical space required for power components, which in turn improves the thermal performance and efficiency of power modules.
- 4. **Smart Control and Monitoring:** Implementing intelligent control systems and software that can dynamically adjust power usage based on workload requirements is another crucial lever. These systems help in reducing power wastage by ensuring that power supply is closely aligned with actual processing needs, thereby optimizing the power usage effectiveness (PUE) of data centers.

These levers are not just about incremental improvements; they represent a fundamental shift in how we think about and manage power in data centers. By continuously pushing the boundaries in these areas, we aim to lead the way in sustainable AI development and operation."

Wevolver Team: Besides energy efficiency, what other areas is Infineon focusing on to power AI innovatively?

**Adam White:** "Beyond energy efficiency, Infineon is deeply invested in addressing two crucial areas of significant importance: TCO and system robustness.

- 1. Total cost of ownership (TCO): TCO is a comprehensive financial estimate that encompasses all costs associated with acquiring, operating, and maintaining an AI server over its lifespan. Given that AI data centers consume substantially more power than traditional data centers and the power demand is increasing rapidly, the operational costs can be significantly higher. The cost of an AI server can be up to 30 times higher than that of a traditional server, mainly due to the intense power and cooling requirements. Infineon's innovations are designed to mitigate these costs by enhancing the energy efficiency and longevity of our components, thereby reducing both energy consumption and the frequency of maintenance or replacements needed.
- 2. **Robustness and Reliability:** The robustness of our systems is of great importance, especially as AI server systems grow in complexity and the stakes of operational



failure rise. Downtime or system failures in AI data centers can be extraordinarily costly, with 40% of firms interviewed in an ITIC survey from 2020 indicating losses greater than \$1 million per hour of downtime. You can imagine that this number will significantly increase in the future as AI servers need to handle more and more LLM training computations. If your server is failing whilst training such a model you must start all over again, meaning that the work of months could be gone. This is consuming significant amounts of electricity, time and financial resources. Interestingly, about 35% of these failures are attributed to issues with power component quality. To combat this, Infineon focuses on delivering high-reliability components that ensure continuous operation. Our products are rigorously tested to withstand the demanding conditions of AI data centers, enhancing the overall system performance and significantly reducing the risk of costly downtimes and reworks. Ultimately this also contributes to reducing CO2 emissions and e-waste.

These areas are critical as they directly influence Al's financial and operational viability. By innovating in TCO reduction and system robustness, Infineon helps Al data centers achieve higher performance, efficiency, and success, ensuring that our solutions provide both technical superiority and economic and operational benefits."

Wevolver Team: Can you give some tangible examples of your innovations that demonstrate Infineon's capabilities in powering AI data centers efficiently?

**Adam White:** "Please allow me to give you two notable examples that showcase our capabilities and innovations in efficiently powering Al data centers:

1. High-Density Power Modules: Recently, we launched the TDM2254xD series dual-phase power modules, which have set new benchmarks in terms of power density and efficiency. These modules integrate our latest OptiMOS™ MOSFET technology with a novel packaging that significantly enhances their electrical and thermal performance. This allows data centers to manage higher power loads more efficiently, reducing both energy consumption and the physical footprint required for power management.

For more details, read our official press release on these ultra-high current density power modules: <u>Infineon Launches Ultra-High Current Density Power Modules to Enable High-Performance AI Computing</u>





High-density, dual-phase, 160 A power modules: <u>Infineon OptiMOS™ TDM2354xD and</u> TDM2354xT

2. Innovative power supply units (PSU) Solutions: We have developed a range of PSUs that cater specifically to the needs of AI data centers, with capabilities ranging from 3 kW to 12 kW. These PSUs utilize a blend of our SiC, GaN, and traditional silicon technologies to achieve optimal efficiency. For instance, our latest 12 kW PSU can support server racks with high power demands while maintaining efficiency levels above 97.5%, significantly reducing the power loss compared to traditional solutions.

Wevolver Team: Looking ahead, what is the future outlook for powering AI technologies? Given the permanence of AI in our lives and the ongoing innovations, how do you see the evolution of solutions for efficiently powering AI?

**Adam White:** "Al's role in our lives is now a constant; it is here today and here to stay and continues to expand in its applications and influence. At Infineon, our commitment to innovation remains firm because we recognize the importance of sustaining this momentum. There are already many excellent solutions in place, reflecting the industry's capacity to address the challenges that come with powering Al systems. Yet, the journey doesn't end here.

The relevance of efficiently powering AI cannot be overstated—it is and will continue to be a critical area of focus for us. Efficient power management is not just about enhancing performance or reducing costs; it's about enabling the sustainable growth of AI technologies



in a way that is compatible with our environmental responsibilities. After all, there's no AI without power. This reality drives us to keep advancing our technologies, ensuring that as AI evolves, our solutions for powering it efficiently and effectively evolve as well.

Looking ahead, we anticipate a landscape rich with opportunities for further breakthroughs. Our goal is to lead the way in making Al operations more sustainable and efficient worldwide. By working closely with the world leading customers and ecosystem providers, we will shape the industry offering best-in-class solutions, which ultimately support our vision of digitalization and decarbonization.

This article was first published on Wevolver.com