The path to megawatt server racks and DC microgrids

Artificial intelligence presents new challenges for data center operations. In particular, power supply and distribution must be completely reinvented to meet growing demands. How can switching to 800 V DC supply voltages help?

By Sam Abdel-Rahman, Senior Principal System Architect at Infineon Technologies

Image: zhou yan/stock.adobe.com



Artificial intelligence (AI) is playing an increasingly important role in countless fields of application. Providers are expanding their data centers to meet this growing demand. More specifically, training ever-larger AI models requires increasingly powerful computing capabilities, including the clustering of up to 100,000 processors into a single virtual machine.

These growing demands are also creating new challenges in power supply at every level – from individual processors to entire data centers. Modern processors need to operate at ever higher load currents and handle strong transient load steps. Power requirements are expected to reach up to 10,000 A per processor within this decade, representing a tenfold increase compared to requirements seen today. AI server racks will require power levels beyond 1 MW. This again represents a tenfold increase in rack power compared to the current standard. Entire data centers will reach power levels in the gigawatt scale, which will necessitate a different infrastructure and new ways of power distribution within the data center. Furthermore, as data centers are turning into substantial consumers of electricity, buffering of load profiles and providing ancillary grid services will become essential.

In addition to these requirements, the power supply for AI data centers must be designed to be as efficient as possible in order to conserve energy despite this enormous energy demand, especially in the face of climate change. Driven by the rising power draw of modern GPUs and the need to cluster more of them per rack, it is clear that traditional power distribution architectures will struggle to support gigawatt-scale data centers without a fundamental redesign. Just as semiconductors have always enabled data processing, storage, and transmission at the heart of data centers, they will also play an increasingly important role in power supply systems in the future. With their help, high power density and efficiency can be combined.

DC microgrids

Achieving ultimate efficiency in an AI data center requires a holistic optimization of the entire power distribution, from energy generation to consumption. As the power demand approaches the gigawatt scale, a new infrastructure for power distribution is required to optimize energy efficiency and, consequently, operational expenditure. DC microgrids are the most likely solution to shape the future of AI data centers. They represent a rethinking of how power is managed within the infrastructure of data centers.

In today's 48-V architecture, the server racks are powered by AC voltage, with AC/DC conversion taking place individually in each rack. Centralized power generation directly from the MV AC grid (10 to 35 kV AC) and distribution via high-voltage DC removes these AC-DC power supplies from the power flow. Thus, within the server rack, only DC-to-DC conversion is required, allowing more efficient and

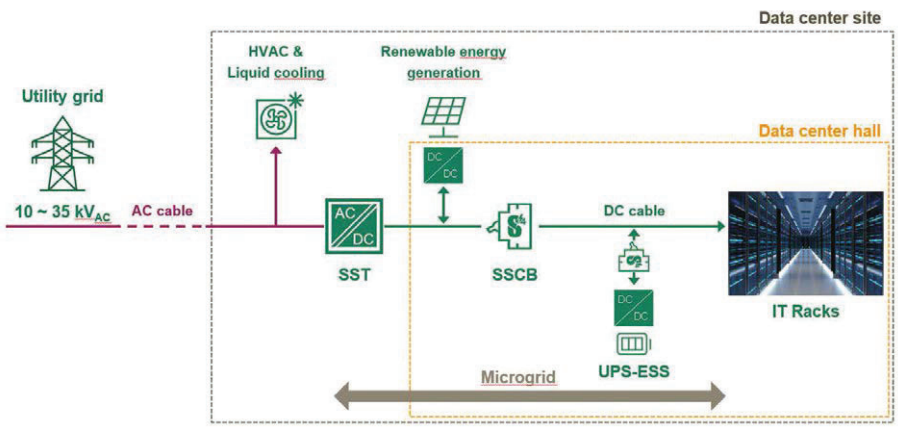


Figure 1: Architectural changes of the power flow for evolving energy demands of AI data centers. Infineon power semiconductor solutions drive efficiency in every power stage, no matter whether it is DC-DC, AC-DC, or DC-AC (Image: Infineon Technologies).

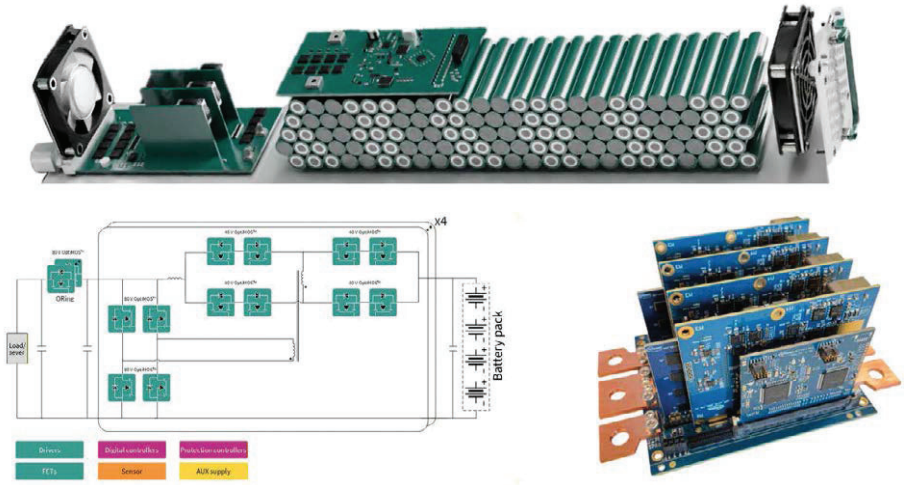


Figure 2: A sample circuit for a 12 kW BBU (Image: Infineon Technologies).

more compact power conversion that may be integrated at the server board level, as outlined above. The DC microgrid becomes the central infrastructure element of the data center, providing energy on an HV DC bus to the server racks as a future-proof scalable solution. Figure 1 illustrates such a scenario.

SSTs and SSCBs: Semiconductors for DC microgrids

In such a scenario, the emerging technology of solid-state transformers (SSTs) will play a decisive role. SSTs are capable of receiving power from the MV AC grid at voltage levels of 10 kV AC to 35 kV AC and providing a regulated high-voltage DC distribution required by the server racks. Their relative compactness and lower weight in comparison to conventional MV transformers allow them to be placed close to server racks for minimal transmission losses. SSTs are expected to provide 2 to 10 MW of power per unit, offering benefits in efficiency, power density, and scalability.

An SST is typically built as an Input-Serial/Output-Parallel (ISOP) system, in which multiple series-connected converter cells form a string that is connected to each phase of the MV Grid. MV AC grid voltages are country-specific and may range from 10 kV AC to 35 kV AC. The SST provides a rectification stage and an isolated DC-DC stage, with all outputs feeding into one DC bus at, for example, 800 V DC. Each converter cell may use either 2-level or 3-level topologies. Infineon supports these applications with a broad portfolio of CoolSiC MOSFETs and IGBTs covering voltage classes from 750 V up to 3300 V. Especially the high-voltage class devices will reduce the complexity of the SST system by decreasing the number of required subsystems.

SSTs provide voltage regulation and disconnect options from the grid in case of failures. Downstream solid-state circuit breakers (SSCB) will be key elements for safety and reliability. With their fast turn-off capabilities, failures can be isolated at relatively low fault currents. CoolSiC JFET devices will be the technology of choice for these tasks. Combining SSTs with SSCBs and subsequent DC-DC stages enables a fully semiconductor-based power conversion chain from grid to core.

Distributed Energy Resources

A further advantage of the DC microgrid architecture is the ability to directly couple distributed energy resources (DERs). These diverse power sources, such as solar, fuel cells, or batteries, can be connected on a high-voltage DC level without any DC-to-AC and AC-to-DC power conversion. Further auxiliary systems, such as HVAC and chillers for liquid cooling, could also be adapted to a DC input and reduce losses due to minimizing the number of power conversion stages.

BBUs: Battery Backup Units

Batteries, in particular, are highly beneficial in AI environments, as they can bridge power outages and cushion consumption spikes caused by intensive computations without placing additional strain on the grid. A battery backup unit (BBU) is used for this purpose, which is composed of a battery management system (BMS) and a DC-DC converter. The BMS allows safe and high-performance operation of Li-ion batteries. It is essential and performs state-of-charge (SOC) and state-of-health (SOH) estimations. The BMS circuit is coupled with a DC-DC converter to provide a regulated output voltage (Figure 2).

For these applications, Infineon offers suitable DC-DC converters, with high power density, efficiency, and thermal performance in the same form factor as today's low voltage BBUs. This is enabled by 650 V/1200 V SiC technology. Built around a multi-level, multiphase architecture, they cut magnetic volume through stacked, interleaved, and coupled non-isolated Boost (discharger) and Buck (charger) stages, converting directly from a battery stack to an 800 V DC bus.

PSUs: From 48 V to high-voltage DC

In today's era of pre-DC-microgrids, 48-V DC power has become the standard. With the rising power consumption of modern GPUs and the need to cluster more GPUs per rack, power levels per rack will rise to 1 MW and beyond. These power levels are pushing the 48-V ecosystem past its limits, so that higher DC voltages are required here as well.

Traditionally, every rack contains power delivery components, backup power, and IT payload, leveraging single-phase power supply units (PSU) of increasing wattage. As rack power levels exceed 250 kW, the architecture is shifting towards dedicated sidecar racks for power delivery and backup power, along with a transition to three-phase AC power delivery (Figure 3). Finally, as rack power levels continue to rise, data centers are expected to adopt a fully centralized high-voltage DC power

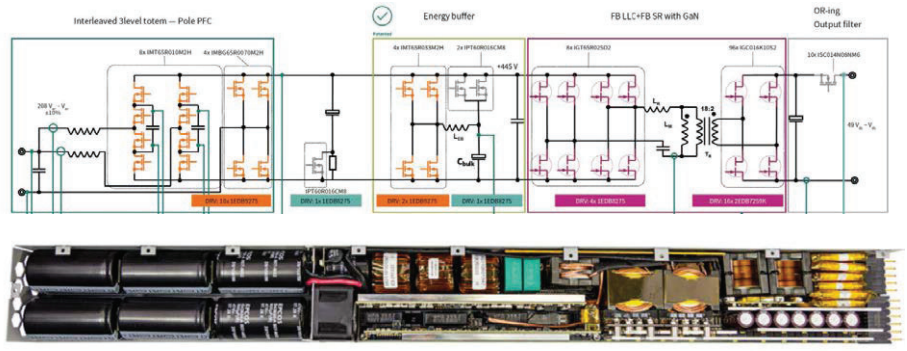


Figure 3: A 12kW 50V PSU for IT racks (Picture: Infineon Technologies).

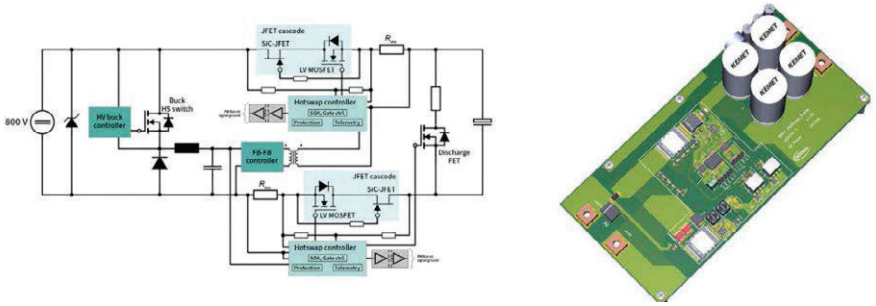


Figure 4: The hot-swap XDP711 with 6 kW from Infineon (Image: Infineon Technologies)

distribution architecture, enabling efficient power distribution across the entire facility.

eFuses and hot-swap functions

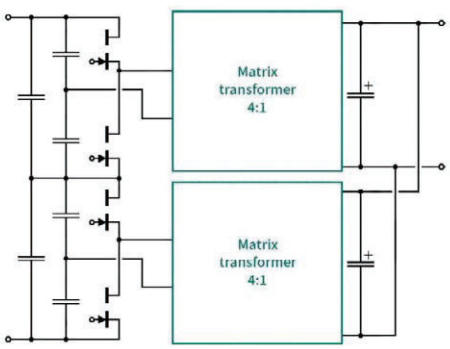
Future server boards will consequently operate directly from 800 V or ±400 V, respectively. This requires several new functions, such as pre-charging of server boards before connecting them to the HV DC bus and discharging of switches to avoid hazardous voltages when a server board is removed from the IT rack (Figure 4). In addition, eFuse functionality is required to interrupt the supply voltage in case of failures. This is an essential safety aspect, which must be implemented on every server board.

IBCs: Supplying the server board

The server board also requires a transition. Despite the high-voltage power supply in the rack, GPU supply voltages remain below 1 V. This means the voltage conversion must take place on the server board, within a space-constrained environment. This can be achieved using either a two-stage configuration via 12 V (800 V to 12 V) or a three-stage configuration of 800 V via 54 V and 12 V (800 V to 54 V to 12 V). While a three-stage approach with 800 V-to-54 V as the first stage, followed by an Intermediate Bus Converter (IBC) and VRM power stages or backside vertical modules, reduces losses in the power delivery network and supports mezzanine card solutions (with the IBC and VRM stages both being placed on the mezzanine card), a two-stage solution with 800 V-to-12 V eliminates an entire power conversion stage and may save precious space on the motherboard. IBCs, in particular, must meet high expectations in this scenario: maximizing end-to-end efficiency to limit thermal load, achieving very high power density near GPUs/XPUs, managing EMI in tightly packed systems, and accelerating design cycles despite increasing conversion complexity across multiple voltage domains (Figure 5).

From grid to core: the road ahead for AI data center power architecture

The future of AI is inextricably linked to the development of advanced power solutions. The increasing power demands of AI data centers, with server boards exceeding 10 kW, server racks surpassing 1 MW, and entire data centers approaching the gigawatt scale, pose significant technological challenges. To address these challenges, advanced semiconductor solutions are necessary to handle complex power conversion processes while improving energy efficiency, power density, robustness, and total cost of ownership. Infineon is taking a holistic approach to addressing these challenges by offering system solutions that span from the grid to the core. The development of corresponding power management solutions, including SSTs, SSCBs, IBCs, and PBUs, will be crucial for realizing high-voltage data center architectures of the future. eg



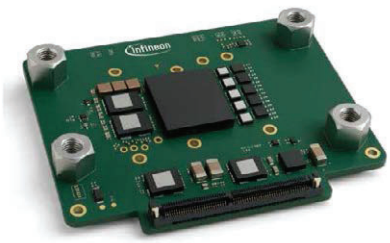


Figure 5: Example circuit of a HV IBC from 800 V to 50 V with 6kW (Image: Infineon Technologies)

>BROXING<
PowerClamp
100-600A for Pressfit
PowerClamp
30-150A for Soldering
SMD Contacts
10-400A
HIGH CURRENT CONTACT ELEMENTS for PCB and IMS
+ SWISS QUALITY +
Representatives and/or distributors wanted worldwide
info@broxing.com
PCIM 2026 Stand 6-234