Technical Article

Data centers evolve to meet AI's massive power needs

Brent McDonald, systems and applications engineer, Texas Instruments

With large language models revolutionizing how we access data, artificial intelligence (AI) advancements are disrupting how industries and societies use data center computing resources. Instead of typing certain keywords into a search engine, we are getting to the point where we can ask AI a question, much like we would ask a person, and get a detailed response. Of course, this is just the tip of the iceberg of what AI can do. AI is also writing code, generating pictures and videos, and transcribing and summarizing meetings. All of these AI functions require significant increases in power to enable them.

Delivering this much power and ensuring that AI can reach its full potential requires reimagining how the IT server rack is structured in a data center, and how best to derive and deliver that power. In this article, I'll examine the derivation and delivery of data center power to the server functions doing the computing, why the power distribution architecture needs to change to meet rapidly evolving AI computing and power requirements, and how to make this possible.

Figure 1 shows the IT server rack-level power requirements needed as a function of time. Figure 1 projects that by 2028, an IT rack will need 1.5MW of power, which is 10 times more power than what server racks are using today.

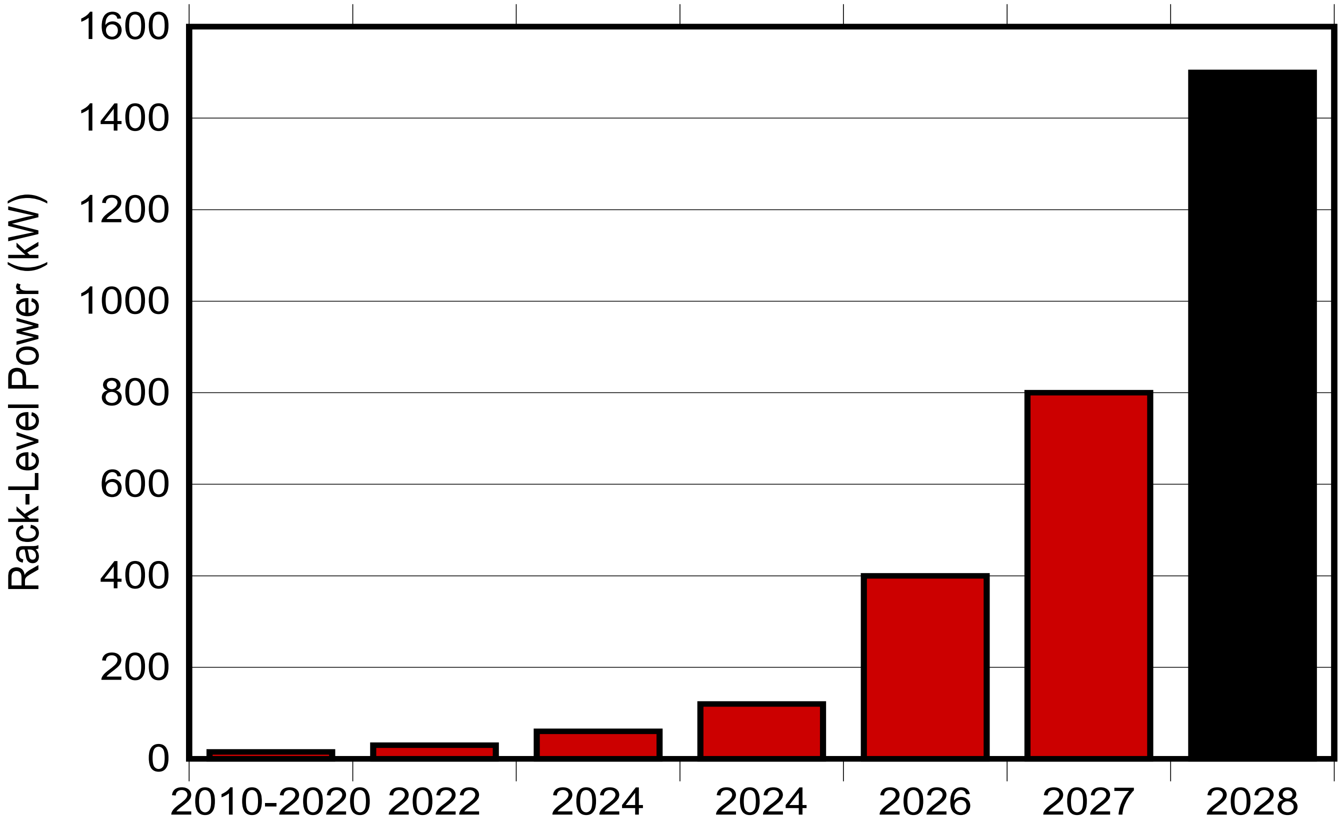

Figure 1. Rack-level power requirements

A brief history

In order to appreciate the magnitude of the changes occurring in the power delivery network inside data centers and servers, let's briefly review the current architecture. Figure 2 shows the first-generation power distribution architecture, which has dominated servers and data centers since the 1990s to today. The top-left area of Figure 2 shows three-phase AC power coming from the AC power grid. This power is transformed from a "medium-voltage" of approximately 13kV to an AC line-to-line voltage of 480V. An uninterruptable power supply (UPS) buffers this voltage.

When the AC grid loses power, the UPS uses local batteries and an inverter function to keep the data center servers running long enough for the backup generators to take over, using either an automatic transfer switch (ATS) or a static transfer switch (STS). The 480V line-to-line AC voltage is equivalent to a $277V_{AC}$ line-to-neutral voltage.

After delivery of the three phases of the $277V_{AC}$ to the IT server rack, a power-supply unit (PSU) performs power factor correction (PFC) and generates a regulated 12V output for distribution to the server IT trays. This 12V distribution voltage for first-generation architectures powers various loads, voltage regulators and other point-of-load regulators (PoL), creating voltages to power the processors, memory and communication integrated circuits used throughout the server trays. This architecture worked well when the total rack power was around 10kW to 20kW. As demand for more computing power has increased, however, so has the power required to handle these computing functions.

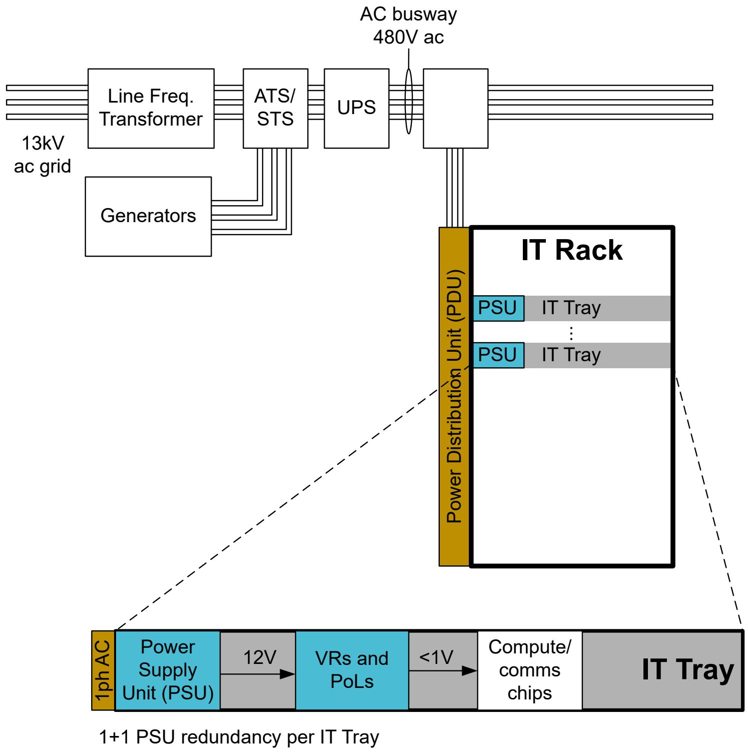

Figure 2. First-generation traditional rack servers

Figure 3 shows the next evolution in the data center power distribution architecture. Starting from the upper-left area of Figure 3, this architecture begins with the same medium-voltage input power source. Like the first-generation architecture, transformers convert from three-phase 13kV to a 480V_{AC} line-to-line voltage. This architecture does not use a UPS but instead sends the equivalent 277V_{AC} line-to-neutral voltage directly to local PSUs inside the IT rack. These PSUs are no longer dedicated to each server tray but rather combined in a single power shelf. In this context a power shelf is simply a shelf of power supplies with their outputs sharing the common load demands from the IT equipment.

There are typically six PSUs in each power shelf in a N+1 configuration to achieve redundancy. Adding power shelves achieves the total required power for the IT rack. The output of these power shelves is a 50V_{DC} bus that is distributed to each IT tray through a high-current busbar running along the back of the server rack. While some second-generation installations retain the UPS function, others will remove it (as shown in Figure 3) and replace it with a local battery backup unit (BBU) that keeps the 50V_{DC} bus active until power is restored or the backup generators can take over. In some cases, a capacitor shelf or capacitor backup unit (CBU) can help eliminate excessive voltage transients, current transients associated with any power disruptions. The 50V bus inside each IT tray goes to a local intermediate bus converter creates the 12V needed to power the system loads in the IT trays.

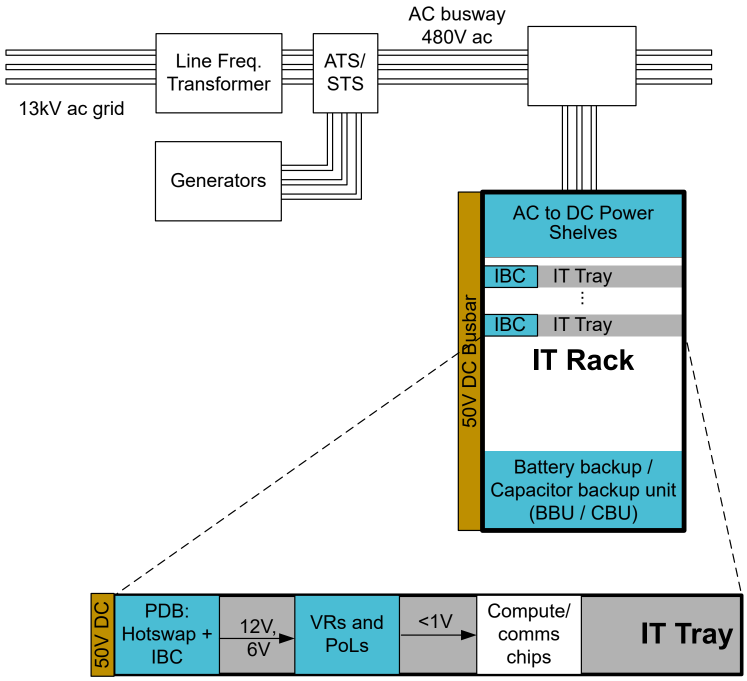

Figure 3. Second generation – cloud and AI computing

The second-generation architecture allows the IT rack to power loads beyond what the first generation can do. Realistic loads for the second generation are in the 100kW range. Once the total power required starts to hit around 200kW, the distribution losses become significant and make further power increases impractical.

AI data center power delivery

Data center racks responsible for running advanced AI models are now expected to eclipse 1MW of power in the 2028 time frame. Distributing this amount of power in a second-generation architecture would result in 20,000A of current, assuming a 50V busbar. The busbar required to deliver this much current is heavy, costly and very impractical. As a result, a higher-voltage 800V_{DC} or ±400V_{DC} bus for power distribution in new AI IT server racks will reduce the high-current busbar requirements from 20kA to 1.25kA. This order-of-magnitude reduction in current will help keep the overall power delivery efficiency high and enable the use of a lower-volume and lower-density copper busbar. Figure 4 shows this architecture.

The power shelf from the second-generation architecture is replaced with a sidecar, which takes the three-phase 480V_{AC} grid voltage as its input. The sidecar converts this to the 800V_{DC} or ±400V_{DC} bus and distributes it to one or more IT server racks. The sidecar also now houses the BBU. In addition to improving the power distribution efficiency, the third-generation architecture also makes more room in the IT rack for computing functions.

In some sense, increasing the computational density of the IT server rack is even more critical than the power distribution problem. AI-based IT racks use hundreds of processors to rapidly process the volume of computations necessary for AI to function at its best. These processors need to be able to communicate with each other in a high-density footprint. Removing a significant portion of the power conversion from the IT rack makes it possible to fit more processors in a smaller space. Now each IT tray in the rack takes as its input this 800V_{DC} or ±400V_{DC} bus voltage. The intermediate bus converter in the tray then converts that voltage for distribution on the IT tray. The distribution voltage could be 48V, 12V or even 6V depending on the chosen architecture.

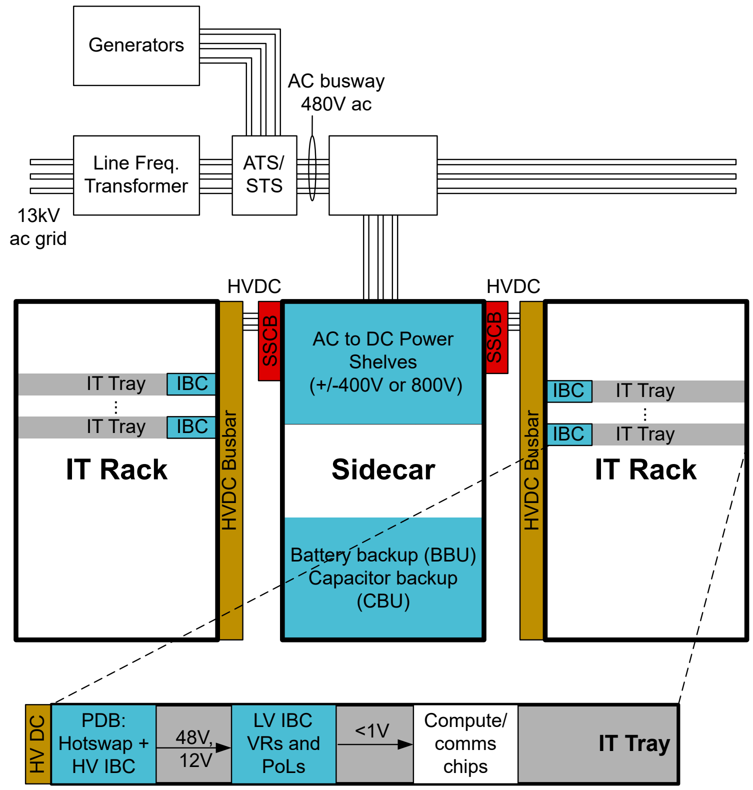

Figure 4. Third generation – AI computing DC distribution sidecar

What comes next?

While the third-generation architecture does a great job of improving the power distribution efficiency and significantly increasing the computation density inside the IT rack, it does so at the expense of taking up more space on the IT floor of the data center. As a result, the next step in the data center evolution is moving the sidecar AC/DC power conversion functions from the IT floor into a utility room.

Figure 5 illustrates a proposal for a fourth-generation architecture. In this architecture, the sidecar houses the BBU functions and the AC/DC function has been moved into a solid-state transformer (SST). In the first-, second- and third-generation architectures, the input voltage is the 13kV medium-voltage provided by the grid. It gets transformed into the three-phase $480V_{AC}$ distribution bus that then gets converted to the DC distribution bus voltage. The SST replaces both the 13kV transformers and the $480V_{DC}$ to $800V_{DC}$ or $\pm400V_{DC}$ power conversion. The SST achieves the PFC function, the voltage step-down and the conversion to DC in a single power conversion stage. The backup generators now need to connect at the medium-voltage node or at the output of the SST through an AC/DC converter. The net result is a higher-efficiency power distribution network and more space on the IT floor for computing.

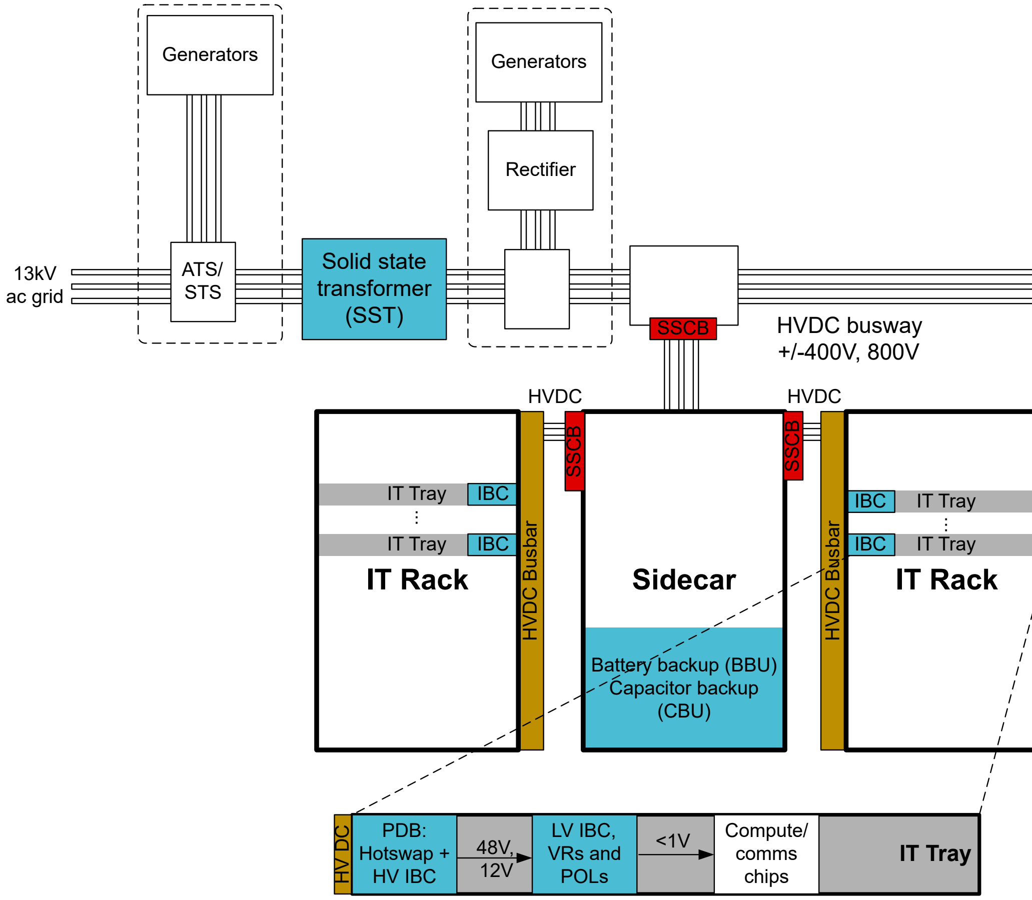

Figure 5. Fourth generation – AI computing SST and DC distribution

The technologies needed to make this happen

Each power distribution generation requires a large number of sophisticated power conversion functions. These functions include things such as PFC, $800V_{DC}$ or $\pm400V_{DC}$ DC/DC conversion, diode ORing, current sharing, hot swap, protection, control and power metering. Advanced semiconductors are the key to each of these functions being able to operate with the highest performance and efficiency possible. For example:

- Performing the PFC and generating the DC bus voltages requires real time microcontrollers [1].
- Enabling topologies such as inductor-inductor-capacitor (LLC) and PFC requires high-efficiency wide band-gap semiconductor switches [2].
- Supporting power metering, control and protection requires accurate current and voltage sensing [3].
- Powering the various isolated switches in the system requires small-form-factor, high-efficiency bias supplies [4] and gate drivers [5].

Conclusion

AI is changing how we interact with information and data. In order to address power conversion needs, data centers need new power distribution architectures.

Future installments in this series will examine the PSU in depth, discuss energy storage, and cover intermediate bus converter and voltage regulator trends, along with the primary technologies and semiconductors that enable them.

References

1. Texas Instruments. n.d. C2000™ real-time microcontrollers. Accessed July 29, 2025.
2. Texas Instruments. n.d. Gallium nitride (GaN) power stages. Accessed July 29, 2025.
3. Texas Instruments. n.d. Isolated ADCs webpage. Accessed July 29, 2025.
4. Texas Instruments. n.d. Isolated power modules (integrated transformer). Accessed July 29, 2025.

5. Texas Instruments. n.d. Isolated gate drivers. Accessed July 29, 2025.

Additional resources

- Check out the article, Five major trends in power-supply designs for servers.
- Read the following articles in Data Center Frontier, Grid-to-Gate: A Framework for Understanding Power-Management Challenges and High-Voltage DC Power: The Future of Data Center Power Architecture.
- Learn more in DataCenter Knowledge: High-Voltage DC: The Power Solution for AI Data Centers.
- Watch the video ±400VDC Rack Power System for ML AI Application.

Trademarks

All trademarks are the property of their respective owners.