Electronic Design

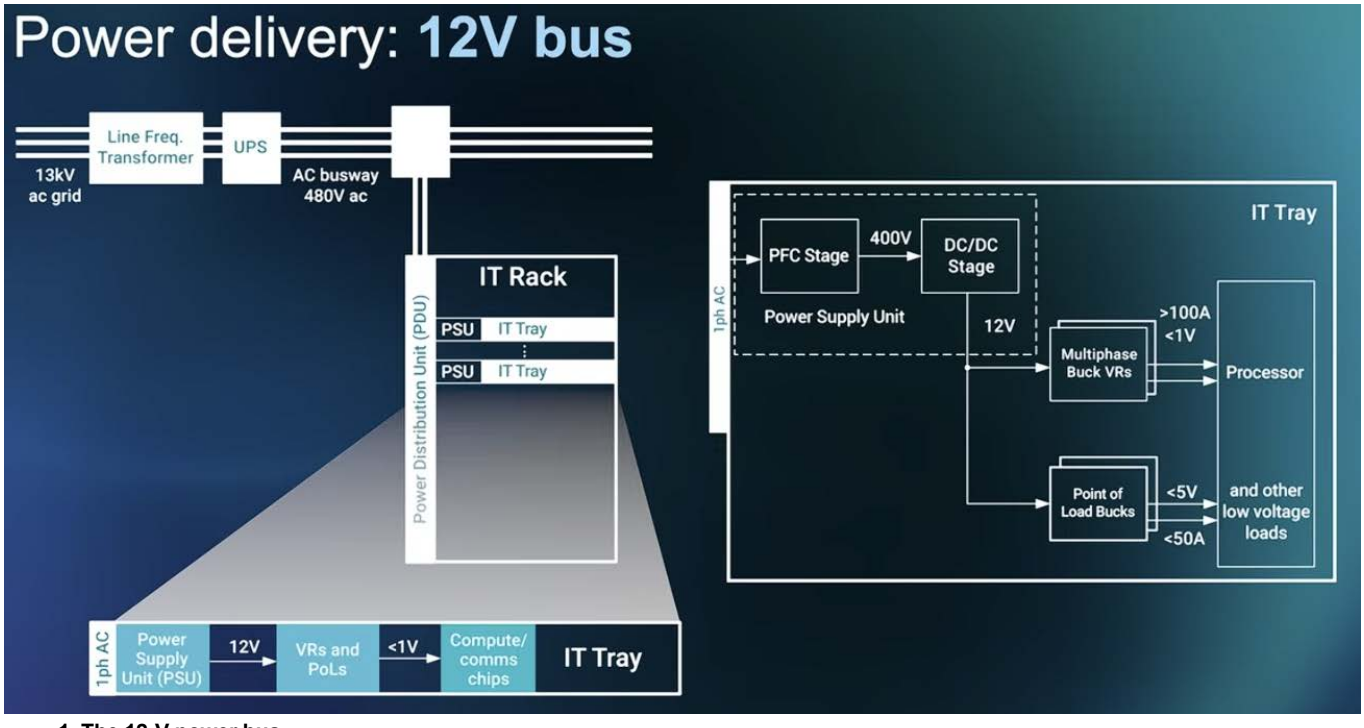
ERNEST WORTHMAN, Contributing Editor

Engineering High-Performance, Reliable, and Efficient Data Center Solutions

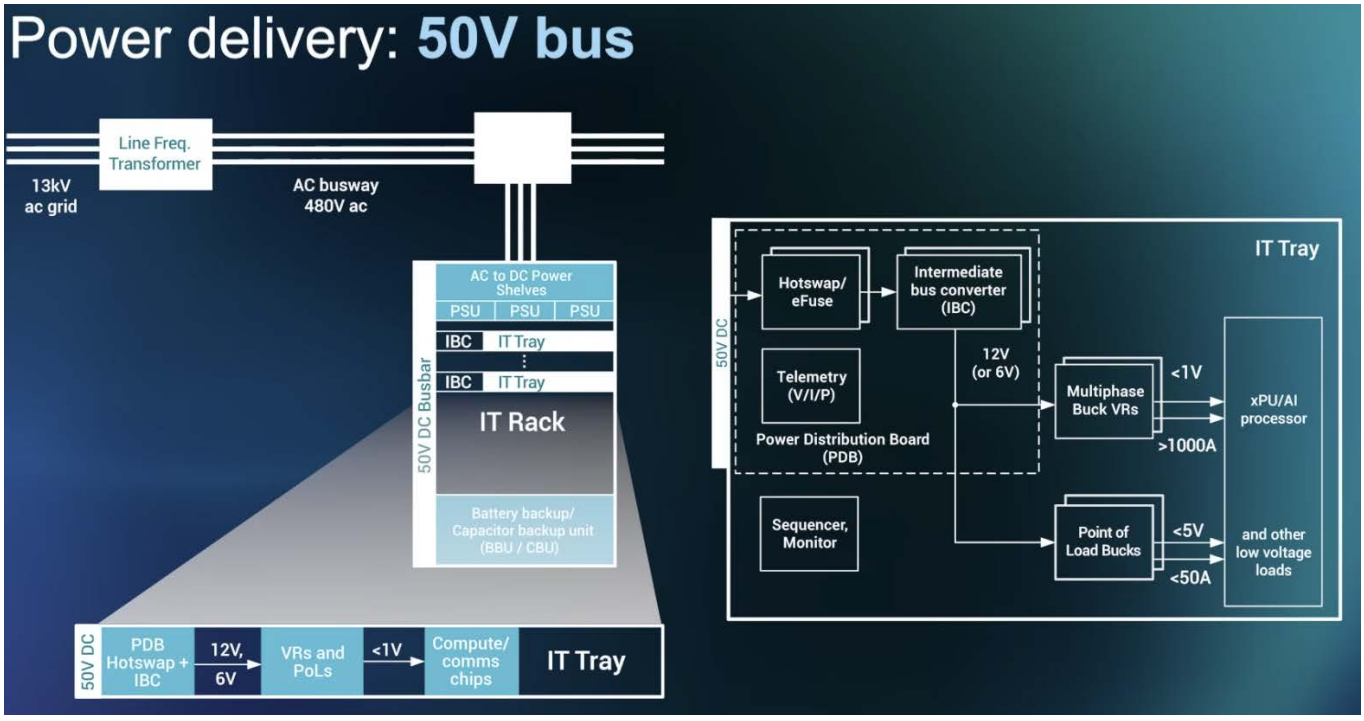
Sponsored by Texas Instruments: This article takes a deep dive into the latest power-supply developments for data centers and how advanced MCUs are implementing GaN semiconductors to meet the higher power demands.

Over roughly the past decade, awareness has grown about the aging power infrastructure and how it could soon hit a critical tipping point. Much of it is 50 to 75 years old and has been struggling to keep up with rapidly rising power demands. By any definition, it's operating at close to full capacity, and the expansion of electric vehicles (EVs), data centers, smart homes, crypto mining, and other extreme power-consuming technologies adding even more pressure.

Perhaps at the top of the pyramid is the mega-data center. Such installations, driven by AI, place massive and unprecedented pressure on the electric grid. By 2030, data centers



1. The 12-V power bus.



2. The 50-V power bus.

are predicted to consume 9% to 12% of the nation's total electricity. A single modern AI data center can use as much electricity as 100,000+ homes.

The exponential increase in artificial-intelligence and machine-learning (AI/ML) processors is placing a proportional impact on power consumption. Today's processors for these platforms are consuming kilowatts of power, a tenfold increase or more over mainstream processors in use today.

One major direction is densification of processor configurations. The emergent philosophy in mega-data centers is to cluster as many processors as possible to handle the enormous data loads. This means lots of CPUs and associated interconnects. However, the more hardware packed into a system, the higher the latency becomes and the more heat is generated, creating a critical design case. A new solution is needed to manage the orders of magnitude in chip density and interconnect.

Bus Architecture Evolution

One obvious vector is to increase the power rails in the data center power-handling system. Traditionally, the grid power to the center may be up to the tens of kilovolts range. The center converted that to some lower voltage, typically a 480-V AC bus from which voltages are further converted to the digital range for processors and other digital hardware, usually from 5 to 12 V *(Fig. 1)*.

This is referred to as grid-to-gate power conversion. So far, the design has been very successful for data centers; most of the racks are designed to consume the low 10s of kilowatts of power. However, it hits a wall with scalability. Scaling this architecture does little to reel in power demand as data centers grow from generic servers to cloud and AI centers, and it has limits.

For this reason, the 50-V bus architecture was developed. Though it may seem counterintuitive, this simply redesigned the power distribution system to make it more able to deliver the higher power *(Fig. 2)* demanded by AI processors. This architecture predominates the build-out of typical cloud and AI centers today and is scalable up to about 200 kW.

The newest architecture is an 800-V bus. This completely redesigns the rack *(Fig. 3)* with high-voltage components. It allows over 1 MW of power delivery.

Such an architecture brings next-generation microcontrollers and gallium-arsenide-based drivers (MCUs and GaN) to next-generation power-supply units (PSUs). As the density of the rack ramps up, these 800-V PSUs must be able to deliver greater than 1,000 A. The challenges here are multifold.

A major issue across the entire data center is conversion efficiency and power loss. For the racks, pushing large currents at low voltage across long board traces creates significant resistive (I^2R) losses that turn into wasted heat. Therefore, concentrated placement requires next-generation layout designs that optimize component placement and minimize interconnects and heat losses.

Primarily, it's necessary to address the components themselves in the PSU. This is accomplished with devices such as GaN field-effect transistors (FETs) and next-generation MCUs.

Let's Drill Down

The design of such PSUs requires an innovative MCU, such as Texas Instruments' F28xxx series. These deliver both

high power efficiency and fast transient response down to 5 A RMS. They integrate fast analog comparators with optimized power control. Thus, high power densities with large power capacity are possible because the fast analog comparators have peak current-mode control and slope compensation that delivers precise power conversion.

Further, these high-power PSU MCUs use a 32-bit digital-signal-processing (DSP) architecture that can run up to 240 million-operations-per-second (MOPS) floating-point unit and trigonometric math for complex, time-critical calculations.

And, to round out the improved performance of the MCUs, their advanced instruction set drastically reduces the number of cycles required for complex math operations. Also leveraged is integrated analog, which handles pulse-width-modulation (PWM) signals entirely through hardware.

What it all boils down to is that this particular MCU/GaN driver combination ramps up the performance of the design, improves transient response, offers unprecedented levels of control and feedback, and offers the unique speed benefits of a hardware-based control loop while maintaining the flexibility of a digital controller.

GaN: The Great Enabler

Marrying the C2000 real-time MCUs with GaN FETs opens the door to new and innovative high-power designs. The combination of these GaN FETs and drivers markedly reduces switching losses and increases high power density when used with compatible MCUs, resulting in higher switching frequencies.

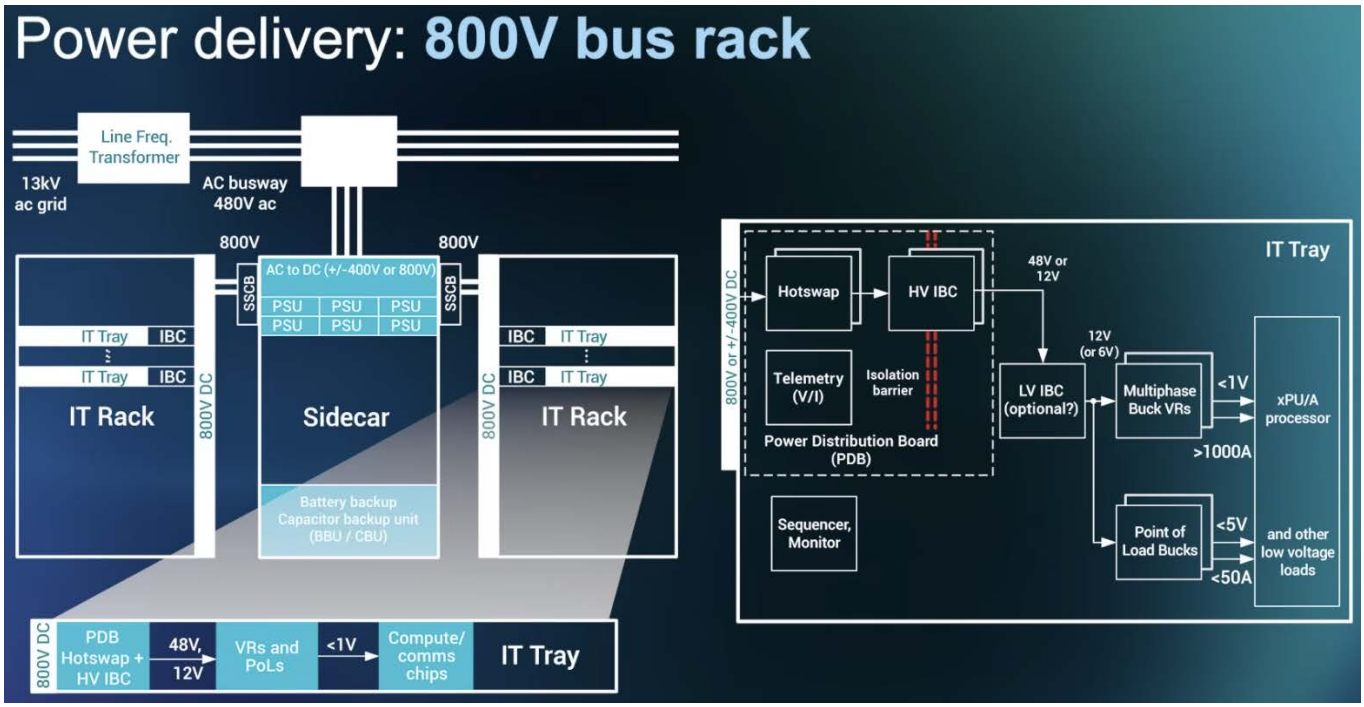
Traditionally, switching frequency losses have been the limiting factor with FETs. With this particular configuration, switching frequencies can be doubled, as high as 2.2 MHz, resulting in twice the speed and half the losses of discrete FETs.

While too lengthy to cover here, implementing this MCU-GaN configuration in the PSU has a slew of other benefits that significantly improve performance, reduce complexity, increase reliability and, of course, reduced power consumption and cost.

Conclusion

As the rollout of AI-enabled mega-data centers ramps up, keeping the grid power in step will not be possible. Of course, there will be innovations: solar, geothermal, wind, nuclear — even extreme concepts such as data centers in space, underwater, or a self-powered data center attached. But significant relief from these alternatives is still a long way off.

Therefore, the best immediate approach is to evolve the power demands of the target. The MCUs and GaN approach represent a timely solution.



3. The 800-V power bus.