

The Role of PSU Design in Meeting AI Data Center Power and Efficiency Needs

By Llew Vaughan-Edmunds, Senior Director, Product Management & Marketing,
Navitas Semiconductor



Al is expanding quickly. As OpenAl CEO <u>Sam Altman said</u> at June's AMD Advancing Al conference, it's growing faster than the infrastructure can handle.

Meta has recently announced a 5 GW capacity data center that will be the <u>size of Manhattan</u>, yet such is the scale of growth that several larger hyperscalers <u>already exist</u>. The number of data centers deployed globally is now approximately 7,000, which is roughly <u>double</u> what it was a decade ago.

And while technology brings huge benefits – <u>especially in fields such as healthcare diagnostics</u> – its power consumption, especially during model training phase, is huge. By next year, data centers are <u>expected to break through 1,000 TWh</u> globally, which is a <u>doubling in just four years</u> vs 2022's 460 TWh. The US alone is likely to reach this 1,000 TWh figure on its own by 2030.

This represents 2% of all energy consumed globally (3% in the US) and an analysis of 2022 International Energy Agency (IEA) data shows data centers already consume more energy than all but nine countries (including every single EU state). And an extrapolation to next year suggests only China, the US, India, Russia and Japan will use more.

Featured in PCIM Magazine - September 2025





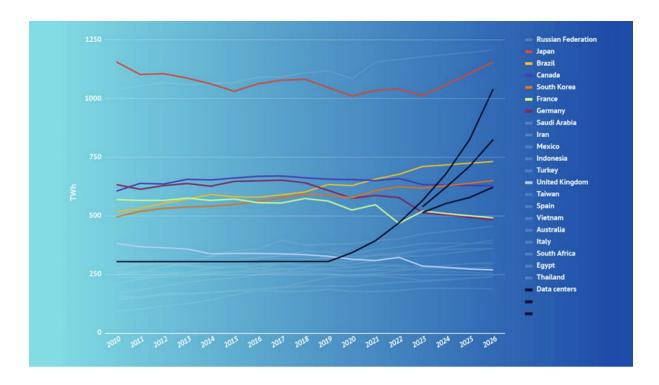


Figure 1: The global share of energy consumption by data centers will grow to be the 6th biggest in 2026.

Source: Navitas based on IEA data

And, by 2030, it is expected that <u>data centers will account for 8%</u> of global electricity usage, with the growth wholly driven by the demands of AI.

To address this, data center operators are making considerable investments in clean energy infrastructure – including large-scale renewables and even small-scale nuclear facilities. And, of course, using the energy as efficiently as possible is absolutely vital both from environmental and operational expense perspectives.

More Efficient Processors

Data centers are looking to more efficient processors (measured in kW per FLOPS). The evolution of the NVIDIA GPUs and servers highlights this trend well.

For example, NVIDIA's DGX H100 (Grace Hopper) is capable of 7X the computing speeds of the previous generation (DGX A100) but only uses 1.5X the power – enabling 0.32 kW per petaflops.

Featured in PCIM Magazine - September 2025





The Blackwell DGX B200 server system takes this further and delivers 72 petaflops at just 0.2 kW per petaflops.

Rumors are starting to energy for NVIDIA's that suggest its Kyber racks could deliver up to <u>15 exaflops</u> of <u>FP4 inference</u>, but power consumption isn't yet clear.

Processors are getting significantly more efficient, but Al's growth still requires ever greater power densities

Source: Goldman Sachs

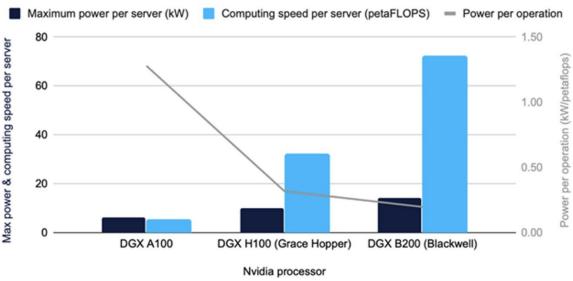


Figure 2: Al servers are becoming significantly more efficient, but the rise of Al usage requires greater power densities to be delivered. Note – figures quoted are per server, not per processor.

Beyond the processor, efficiency advances are being made through algorithm improvements, with China's DeepSeek enabling the training of frontier AI models to be made with <u>far less computing power</u>.

Clearly the advances being made in both GPU and algorithm are hugely important, but we should be aware of Jevon's Paradox. This observes that making systems more efficient often leads to an overall increase in use. And improvements are therefore needed across the whole system. This leads to the next big issue: power supply efficiency.

CARBON NEUTRAL NEUTRAL NO.



Standards Driving Efficiency

The 80 PLUS framework from <u>CLEAResult</u> has now been in place for over 20 years and gives an industry benchmark for PSU efficiency.

Not only does certification require specific efficiency levels to be met, but it also avoids PSU manufacturers being able to provide only a single efficiency figure, which could mask poor performance outside of a narrow peak efficiency range.

The original 2004 <u>80 PLUS</u> specification required PSUs in computers and servers to reach at least 80% energy efficiency at 10%, 20%, 50%, and 100% of rated load, and have a true power factor of at least 0.90 in 50% loads.

Over time, the framework evolved to include more stringent levels for higher performance power supplies: Bronze, Silver, Gold, Platinum and eventually Titanium in 2011. This increased the required efficiency at higher loads to 96%. It also added a requirement for 90% efficiency at just 10% loads and moved the PFC requirement down to 20% loads.

Earlier this year, 80 PLUS added its first standard upgrade since Titanium, with the new leading standard, Ruby, increasing peak efficiency to 96.5% at 50% of load (277 V and 480 V internal redundant supply frameworks). It also adds a minimum efficiency requirement at 5% loads (90% efficiency), and extends the PFC requirement to all loads, with a PFC of 0.9 for 5% loads and up, and 0.96 from 20% loads.

80 PLUS Certification % of Rated Load	277V / 480V Internal Redundant				
	5%	10%	20%	50%	100%
80 PLUS					
80 PLUS Bronze		80%	82%	85% PFC ≥ 0.90	82%
80 PLUS Silver		82%	85%	89% PFC ≥ 0.90	85%
80 PLUS Gold		85%	88%	92% PFC ≥ 0.90	88%
80 PLUS Platinum		88%	90%	94% PFC ≥ 0.95	91%
80 PLUS Titanium		90%	94% PFC ≥ 0.95	96%	91%
80 PLUS Ruby	90% PFC ≥ 0.90	91% PFC ≥ 0.90	95% PFC ≥ 0.96	96.5% PFC ≥ 0.96	92% PFC ≥ 0.96

Figure 3: 80 PLUS requirements for 277 V / 480 V internal redundant PSUs (Source: CLEAResult)

Featured in PCIM Magazine - September 2025





These may appear relatively small improvements, just one- and half-point increases versus Titanium, however the shift from 96% to 96.5% represents a 12.5% reduction in losses.

This means that as global data center electricity consumption reaches 1,000 TWh, Ruby-standard PSUs can save 5 TWh. And based on the US's standard electricity source mix and using EPA eGrid 2023 data tables (see here, and here) this converts to just over 2 million tons of CO2.

Reaching Ruby

To meet such stringent targets, designers must consider the best topology for each conversion stage and the materials used for the primary power switches.

The use of wide bandgap (WBG) materials, such as silicon carbide (SiC) and gallium nitride (GaN), is critical in and enable the highest possible power densities and efficiencies.

In addition to the semiconductor, designers will need to work with industry standards for compatibility and interoperability when designing server PSUs, notably the Open Compute Project (OCP) Data Center – Modular Hardware System (DC-MHS) specifications that standardize interfaces and form factors for equipment targeting data centers.

M-CRPS (Common Redundant Power Supply) is an OCP standard that standardizes 1U-high internal PSUs. More recently this has started to be superseded by the Open Rack v3 (ORv3) power shelf, which incorporates six standard-form-factor PSUs per shelf.

Standardization has also happened on supply voltages, with M-CRPS PSUs now tending to favor 54 V over the traditional 12 V output to reduce current, and power shelves tending to standardize on 48 V.

This need for high efficiency within a tightly constrained footprint means an LLC converter is now essential for PSU designs and this is especially true when basing these AI data center PSUs on WBG technologies.

The thermal performance, leakage inductance, and power density of these LLC converters can all be enhanced through the use of planar transformer technology with flat windings embedded within the PCB layers.

Combining this with the use of high-frequency GaN devices also gives a low voltage-second product that eliminates core saturation and further reduces overall losses.

CARRON NEUTRAL: Navitas



Designing for Power Density

Using this topology and approach, Navitas has designed a single output 54 V CRPS, using high-power GaNSafe ICs and Gen3-Fast SiC MOSFETs with an output power rating of 4.5 kW, and achieving an industry-leading power density of 137 W/in³.

With this design, a wirewound transformer could handle the 83 A current. However, to meet the size constraints of the OCP standards, the switching frequency must be (at least) 300 kHz to reduce the size of the output capacitors and magnetics. And this means silicon is no longer suitable for OCP-compliant designs as this is beyond the level at which silicon can function.

At this point it should be noted that particular care has to be paid to the gate drive as the thresholds are lower than standard silicon MOSFETs and voltage spikes or ringing can reduce device reliability. Navitas' high-power GaNSafe ICs used in this reference design include control, drive, and protection, eliminating the risk of damage.

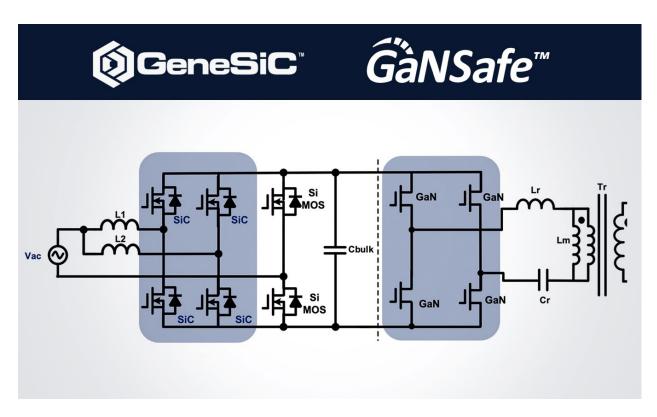


Figure 4: Navitas' 4.5 kW, 54 V CRPS achieves 137 W/in³ using SiC and GaN devices.





In the reference design, a bridgeless interleaved totem-pole PFC stage is used as there are far fewer losses than with a conventional bridge rectifier. The design is based upon Navitas Gen-3 Fast SiC MOSFETs as they have very low reverse recovery and switching losses. Therefore, the PFC stage can operate in continuous conduction mode with loss levels that simply could not be achieved with silicon.

Navitas' GeneSiCTM devices are based upon a 'trench-assisted planar' structure that enhances performance through an $R_{DS(ON)}$ that is 20% lower than competing devices, as well as being more stable over extended temperature ranges.

As a result, the reference design delivers over 97% efficiency at 50% load.

Evolution from 2.7 to 12 kW

Of course, power density is vital, but to meet the needs of Blackwell- and Rubin-based servers, absolute power delivery is vital too.

A recent analysis of Blackwell architectures has been undertaken by <u>AMAX Engineering</u>, which focuses on designing IT infrastructure solutions, with a particular focus on AI. It concluded there was already a "readiness gap" emerging, with Blackwell configurations requiring a level of power that (at time of writing) more than 95 percent of the world's data centers were unable to support.

As such, it concluded that, if unaddressed, this gap will lead to a "two-tier" data center ecosystem. And advances in PSU power delivery play a critical role in preventing this readiness gap.

Advances are happening rapidly as a result of evolutions in GaN and SiC as well as in digital control technologies. This has led to the <u>industry's first 12 kW supply</u> for ORv3 specifications and Open Compute Project (OCP) quidelines, which was launched in May of this year.





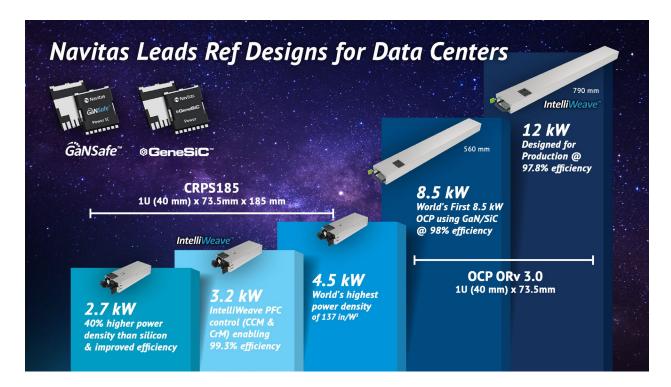


Figure 5: Navitas PSU Roadmap

To put this 12 kW PSU in context, for Navitas output power has increased more than four-fold in not much more than two years, starting with the 2.7 kW production-ready PSU design based on GaNFast ICs. This was superseded in August 2023, with a 3.2 kW design achieving a PFC efficiency of 99.3%, using the IntelliWeave™ digital platform control, power density of 100 W/in³ and a total system efficiency above 96.5%.

Seven months later, in March 2024, this rose to 4.5 kW, which delivered the highest power density of any PSU. November last year saw the launch of the world's first 8.5 kW supply – created for OCP and ORv3 specifications. Like the 4.5 kW device, this implemented SiC and GaN, but implemented a 3-phase topology for both the PFC and LLC stages. As a result, it was able to reduce ripple current and EMI versus competing devices, while also delivering a peak efficiency of 98%.

The 3-Phase LLC topology is enabled by fourth-generation GaNSafe devices. GaNSafe is the world's safest GaN and has a short-circuit protection with a maximum latency of 350 ns, as well as a 2 kV ESD protection on all pins. The fourth generation integrates control, drive, and sensing, as well as critical protection features to further enhance reliability and robustness, with the features controlled via 4-pins to allow the package to be treated like a discrete GaN FET (no V_{CC} pin is needed).

Featured in <u>PCIM Magazine</u> – September 2025





The 3-Phase interleaved continuous conduction mode (CCM) TP-PFC is powered by Gen-3 Fast SiC MOSFETs with a <u>trench-assisted planar technology</u> that offers leading performance over temperature for cool-running, fast-switching, and enhanced robustness.

And this leads us back to the 12 kW PSU, launched May 2025.

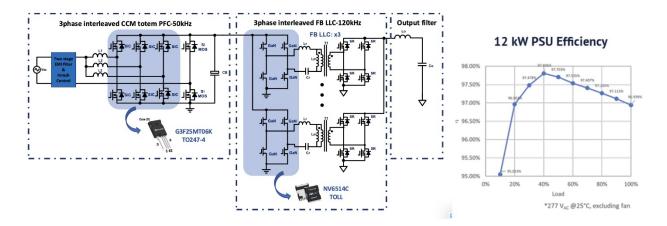


Figure 6: Navitas' 12 kW production-ready OCP PSU design delivers a power density of 85 W/in³, and an efficiency of 97.5% at 50% loads.

Again, this implements Gen-3 Fast SiC MOSFETs, and high-power GaNSafe ICs; and like the 8.5 kW PSU, also uses a 3-phase interleaved TP-PFC and FB-LLC topologies to ensure the highest efficiency and performance, with the lowest component count. But the design also implements a novel digital platform called IntelliWeaveTM.

IntelliWeave provides a hybrid control strategy for both CCM and critical conduction mode (CrCM) allowing use in both light- and full-load conditions. The technology ensures maximum efficiency while keeping a more simplistic, lower component count design that reduces power losses by up to 30% versus existing CCM designs.

The 12 kW OCP PSU has an input voltage range of $180-305\ V_{AC}$ and outputs at up to $50\ V_{DC}$ to ensure 12 kW at voltages above 207 V_{AC} , or 10 kW below this.

CARBON NEUTRAL NEUTRAL NO.



Conclusion

The explosive growth of AI has triggered an unprecedented demand for energy consumption. This is putting both organizations', and countries' commitments to reduce reliance on fossil fuels at risk. The increase in clean-energy investment is welcome, but reducing the energy used is absolutely vital. And this places immense pressure on the GPU manufacturers, on the cooling system developers and on the PSU developers.

To prevent an infrastructure bottleneck and mitigate environmental impact, PSU manufacturers face a dual challenge. They must not only meet increasingly stringent efficiency standards like 80 PLUS Ruby, where even fractional improvements save several terawatt-hours of consumption globally, but also dramatically increase the sheer amount of power delivered.

The recent, rapid escalation in PSU capabilities is a direct response to this pressure. Power delivery from a single unit has soared from a few kilowatts to 12 kW in just a couple of years, a rate of innovation essential for supporting next-generation AI servers. Ultimately, the ability of the PSU industry to continue this rapid evolution in both efficiency and power density is a fundamental requirement for sustaining the AI revolution.

CARRON NEUTRAL NAVITAS