WHITE PAPER
FLEX POWER MODULES

Optimizing Power Electronics for AI, Machine Learning and Cloud Applications

This white paper explores various power architecture strategies for use in AI/ ML data centers and their trade-offs, emphasizing that efficiency in power rail generation is crucial to achieving the objectives of minimizing temperatures and energy costs, while maximizing data throughput in a minimum footprint.

Table of Contents

Introduction

New data center applications such as AI, machine learning, cryptocurrency mining and cloud computing have accelerated the requirement for more processing power in ever-more compact form factors. This demand adds to the baseload from more traditional activities such as video streaming. An inevitable consequence is a significant increase in energy demand. According to the International Energy Agency (IEA), electricity consumption from data centers – including those used for AI and cryptocurrency – could double by 2026. Data centers are major drivers of growth in electricity demand in many regions. After consuming an estimated 460 terawatt-hours (TWh) globally in 2022, data centers' total electricity consumption could reach more than 1,000TWh in 2026. This demand is roughly equivalent to the electricity consumption of Japan (1). These figures are expected to rise further, presenting both environmental and commercial incentives to enhance the efficiency of power delivery networks within data centers.

---

[1] https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks

Data center cabinet power is heading towards 200kW

In a typical data center today, the challenge is to supply around 30-40kW to each cabinet from the AC utility supply, ultimately providing sub-1VDC to the most power-hungry CPU and GPU end-loads. These systems also require higher voltages at lower power levels. For example, the Nvidia H100 AI accelerator, which contains 80 billion transistors, has a thermal design power (TDP) of 700W, and the latest Blackwell B200 is rated at 1000W. In comparison, current CPUs such as the Sky Lake/ Cascade Lake types have a TDP of 200W. Looking ahead then, cabinet power could exceed 200kW due to further improvements in processing capabilities and higher packing density of server blades, with currents at the lowest voltages sometimes exceeding 1000A and reaching even higher peaks (2).

To manage such extreme power levels, the power distribution network (PDN) must be optimized, and various metrics can be identified: energy efficiency, power density of converters, rack space utilization, floor space utilization, cost, upgradeability, serviceability, uptime and more **(Figure 1)**.

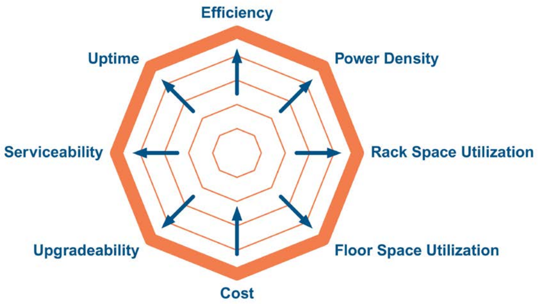


Figure 1: Data center power distribution network metrics and their direction of travel

Intermediate Bus Converters in a data center PDN

At the rack level, many PDN architectures have been proposed, adopted, and often superseded over the years which were argued to be optimum at the time. However, as technical requirements and priorities have been updated, the ideal architecture has also evolved. The parameters that have shifted markedly that now need to be considered include:

- Efficiency for cost and energy savings
- Voltage drops in connections from higher current levels at lower voltages
- Static and dynamic variations in voltage with greater step load changes and di/dt
- Noise from increased switching frequency
- Safety and functional isolation arrangements
- Cooling methods for increased rack power levels

Because resistance in connections significantly contributes to power loss, it is more practical to distribute power at a higher voltage. This approach results in lower current for a given power level and reduces power dissipation (I2R). For safety, it is recommended to use a nominal voltage below the Safety Extra Low Voltage (SELV) limit of 60V. Consequently, voltages like 48V or 52-54V are increasingly adopted, particularly where statutory insulation and isolation requirements are less stringent. Historically, these voltages were and still are widely used in IT systems. A battery could be connected to this rail for back-up, and then DC/ DC converters reduce the voltage to the formerly popular load levels of 12V or perhaps 5V in a Distributed Power Architecture (DPA). However, as end-load voltages have decreased to today's lows of around 0.5V, creating efficient DC/DC converters for such high conversion level from 48V becomes challenging. One solution is to implement an Intermediate Bus Architecture (IBA) which divides the DC/DC conversion into two stages and includes an optimized intermediate voltage for enhanced overall efficiency as illustrated in **Figure 2.**

---

[2] https://www.ramboll.com/en-us/insights/decarbonise-for-net-zero/100-kw-per-rack-data-centers-evolution-power-density

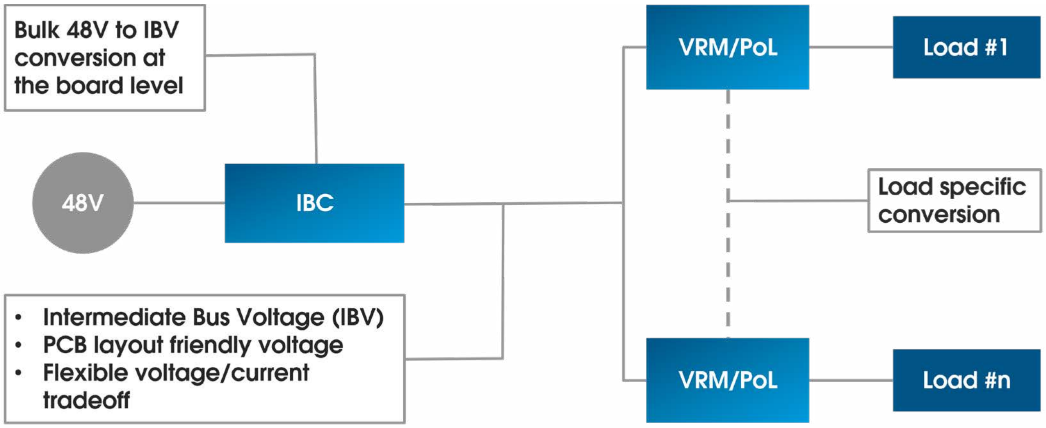


Figure 2: The intermediate bus architecture

The first DC/DC stage, known as an Intermediate Bus Converter (IBC), has traditionally been isolated to address safety concerns and also regulated, since legacy loads like Peripheral Component Interconnect Express (PCIe) cards, Dual Inline Memory Modules (DIMMs) and fans have run directly from the 12V supply and thus require precise voltage levels. An example of such a solution might be the BMR491 series from Flex Power Modules – a fully regulated Quarter Brick delivering up to 1,300W with 1,500V isolation from a 40-60V supply voltage.

However, in the data center environment, the 48/54V feed is not derived from batteries, but from isolated shelf level AC/DC power supplies. Consequently, the need for isolation in the board mounted DC/DC converters can be relaxed, as the safety requirement is transferred to the AC/DC PSU. This shift has led to further developments of non-isolated, regulated board-mounted DC/DC converters such as the BMR351 which, partly because of the removal of isolation, can now deliver up to 1,600W of continuous power while maintaining the same quarter brick footprint.

In addition, particularly for AI applications (and increasingly for standard cloud computing applications), the load voltages are derived from Voltage Regulator Modules (VRM). These modules typically accept a wide range of input voltages. With that in mind, the DC/DC converter can now also be unregulated, converting the input voltage to an output voltage at a direct ratio, such as 4:1. For example, a 48V input would yield a 12V output, but a range from 40-60V would result in an output ranging from 10-15V.

With the relaxation of design requirements, companies like Flex Power Modules are now able to offer higher power levels in significantly smaller packages. This leads to solutions such as the BMR313, which delivers 1kW of continuous power and up to 3kW of peak power from a tiny package measuring just 23.4 x 17.8 x 7.65 mm – approximately 80% smaller than a similarly rated quarter-brick package size. This significant leap in terms of power density explains why 90% of the AI industry is now choosing non-isolated, unregulated (fixed ratio) front-end IBCs.

IBC variations

These fixed ratio intermediate bus converters are available with different ratios between input and output voltages. Popular ratios currently include 4:1 and 8:1 but ratios such as 5:1, 6:1 and 10:1 are also becoming more prevalent [3]. For a nominal 48V input, these ratios result in nominal outputs of 12V, 6V, 9.6V, 8V and 4.8V respectively. For non-regulated parts, the ratio is fixed, and the output voltage varies in proportion to the input. For example, with an 8:1 part, the output varies from 5 to 7.5V as the input ranges from 40V to 60V. However, input transients also pass through in the same ratio.

The choice of which IBC to use often depends on the downstream VRM converter efficiency curve, to

---

[3] https://flexpowermodules.com/whats-driving-the-need-for-different-input-output-ratios-of-dc-dc-converters-in-distributed-power

minimize system losses. Several other factors also need to be evaluated: typical, peak, minimum operating input voltages and power levels, how much current is distributed on the board at different voltages, as well as the size and cost of the IBC variant and its placement. Higher conversion ratios imply lower VRM input voltages and higher associated current levels, and the VRM may or may not be more efficient under these conditions but resistive losses from the higher current may offset any gains if the IBC is not placed very close to the VRM. In a system with multiple VRMs and end loads fed from one IBC, it is practically impossible for the IBC to be in an optimum position for all loads.

Examining an example using available components from Flex Power Modules can provide some valuable insights. If we simplistically assume a load that requires 0.8V at 360A maximum (288W) with a source supply of 40-60V, 48V nominal, we could choose six paralleled 2-phase 80A rated VRMs type BMR510 with an input range of 4.5-15V which could be driven from IBCs with any ratios from 4:1 to 8:1. The efficiency of the VRMs can improve with a low input/output voltage differential at loads below around 75%, so an IBC with 8:1 ratio providing nominal 6V to each VRM might be initially considered, such as the BMR320 **(Figure 3)**. From their efficiency curves **(Figure 4)**, this combination would yield an overall efficiency at 360A of about 86.4%, with about 5.6W dissipated in each VRM and a 11.7W lost in the IBC. The VRMs have a combined input current of 53.5A. The IBC and each VRM are operating at around 80% of their maximum rating, which provides a comfortable engineering margin, particularly in liquid cooling environments common in AI applications.

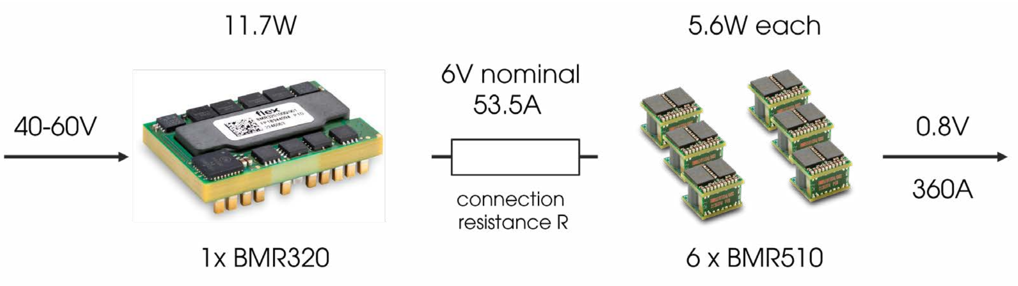
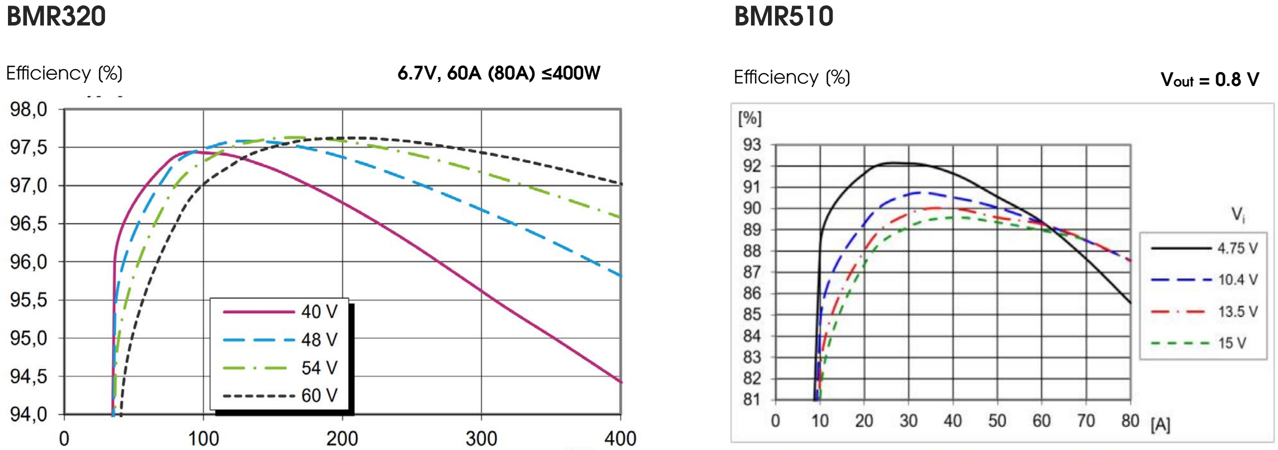


Figure 3: Arrangement with 8:1 IBC yielding overall 86.4% excluding connection resistance



BMR320

BMR510

Driver losses excluded, 2 phases, fsw=550KHz

Figure 4: BMR320 and BMR510 efficiency curves

We could now reflect on the same specification but using a 4:1 IBC such as the BMR313 **(Figure 5)** with its efficiency curves shown in **Figure 6**.

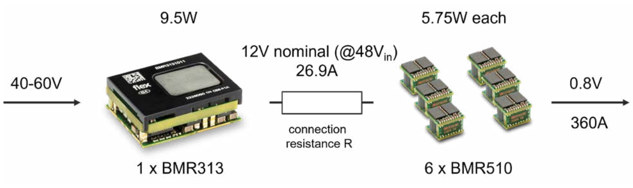


Figure 5: Arrangement with 4:1 IBC yielding 86.75% efficiency excluding connection resistance

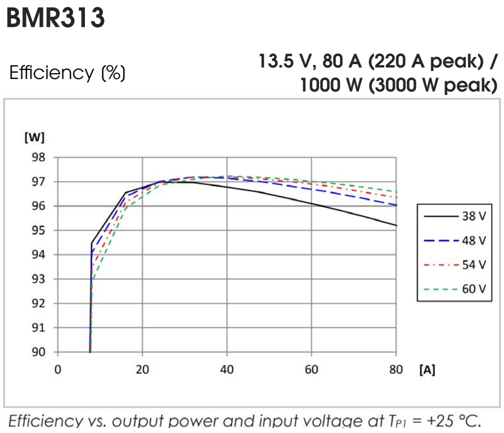


Figure 6: BMR313 4:1 IBC efficiency curves

Now, the VRMs run from 12V nominal input, but at approximately 80% load there is actually very little difference in efficiency from the 8:1 IBC, as the VRMs have been optimized for a wide range of input voltages. This lower ratio IBC operates at a slightly higher efficiency, around 97.1% compared to 96.5%, resulting in an overall efficiency of about 86.75%. While this might not sound like a significant improvement at a module level, it does offer substantial benefits in terms of reduced power dissipation and heat, especially when considering the scale of a modern-day data center composed of thousands of such server boards.

With the 4:1 IBC, each PoL dissipates about 5.75W but the IBC only dissipates 9.5W. The difference in dissipation between the two scenarios may not be large, but because the PoL input current is approximately halved, dissipation in connection resistance (R) reduces by a factor of four. For example, a resistance of just 5 milliohms for each of the total go and return paths with a current of 53.6A results in over 14W dissipation, dwarfing that of the modules. This decreases to 3.6W for a 12V bus voltage under the same connection arrangement, and probably even less as the copper runs cooler and the resistance is, therefore, lower. Similarly at 53.6A and 10 milliohms, over 0.5V drop is experienced.

The fixed ratio IBC provides superior bandwidth between input and output, which means that, with low source impedance (inductance), the 48V capacitors can be effectively used as an energy reserve for the VRM during load transients. Since increased length also contributes to higher inductance, placing the VRM close to the IBC is ideal for optimum performance, although board packaging constraints may not always allow this.

Significantly, although the 8:1 BMR320 has approximately the same footprint as the 4:1 BMR313, the BMR313 delivers more than double the continuous power rating, and, therefore, operates at about one-third of its capability in the described application. This is a key advantage of the ratio in DC/DC design, allowing for much greater power density in the 4:1 ratio DC/DC converters. The additional capability could be used to power other system components or to increase the number of paralleled PoLs for a larger single load.

To achieve these levels of power density, more effective cooling methods are needed than just forced air cooling. In modern systems, liquid cooling technologies are often employed which can more effectively dissipate the heat generated by such small form factors. In today's data centers serving AI workloads, direct-to-chip cooling is frequently used, where cooling fluid is delivered directly to the load via heatsinks with built-in piping. However, for the highest rack powers, full immersion cooling is also being considered, posing additional challenges for component suppliers. Flex Power Modules plans to address these topics in future articles on its website.

Another factor to note is the ability of these unregulated IBCs to handle higher peak powers effectively, compared to their regulated transformer-based brick type cousins. For example, the referenced BMR313 can deliver continuous power of 1kW and peak powers of up to 3kW, which far exceeds the capabilities of similarly rated parts like the physically much larger BMR350, which delivers 1.3kW continuously and 1.7kW of peak power. The key difference lies in the larger inductance in the output filter of the regulated brick, which slows its response to fast changes in output current.

Another practical difference to consider between 4:1 and 8:1 IBC voltages is the capacitor ratings on the bus. A bus voltage of 6V, varying from 5-7.5V can safely use 10V capacitors with a comfortable 25% headroom. In contrast, a 12V bus, which can vary up to 15V, necessitates the use of 25V rated capacitors with an unnecessarily generous 66% headroom, thereby increasing the size and cost of the components.

Positioning of IBCs and PoLs

Particularly if cooling is limited, there is a good case for using multiple lower-power IBCs to spread the current load, reducing each DC/DC hotspot temperature. I2R losses near the converters can also be reduced if each IBC can now be placed closer to its associated VRMs, with the lower resulting copper trace temperature further reducing losses. Although rarely quantified in data sheets, IBC losses typically rise with temperature due to the strong positive temperature coefficient of switch-on-resistances. This suggests that running multiple IBCs at cooler temperatures could yield efficiency benefits compared to a single IBC. You must also take into account the commercial implications of using multiple lower-powered devices in parallel and the space required for them, compared to using devices with greater power and capacity ratios.

Typically, the VRMs located near the end-load on the main board are arranged laterally, next to the CPU or GPU. This configuration can sometimes complicate cooling efforts and cause the wide, high-current power connections to block access to the signal lines. As previously discussed, this layout also makes it challenging to position the IBC as close as possible to the load on the top side of the customer's PCB.

An increasingly adopted solution to this issue is 'Vertical Power Delivery' (VPD). In this approach, the VRMs are configured vertically as a multi-phase, multi-rail solution immediately below the end load as illustrated in **Figure 7**. The terminations are designed to match the CPU or GPU pinout, minimizing connection length. These VPD solutions are necessarily fully customized to align with the power tree and ball grid array of the specific processor being powered. Flex Power Modules has developed solutions using 2- or 4-phase 'building blocks' at their core which allow for fast time to market and minimum design time when requirements change.

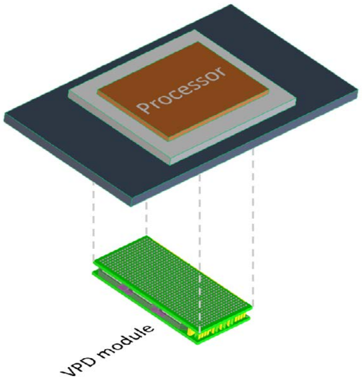


Figure 7: 'Vertical Power Delivery' for a processor

The VPD device typically serves as the power train of the VRM DC/DC converter, and a multi-phase controller is usually fitted elsewhere, although designs can also incorporate the controller within the module itself.

With this arrangement, multiple voltage rails can be generated with multi-phase control for a high level of integration and efficiency. Heat is extracted through a 'cold plate', typically with liquid cooling incorporated, fitted on the underside of the VPD device. Using such a vertical solution for the final stage conversion also frees up space on the top side of the PCB. Therefore, this design method allows the initial IBC conversion stage to be located closer to the load, significantly reducing distribution losses further.

Conclusion

The data center power delivery network requirements all interact when system architects seek an optimum solution. High efficiency is a noble goal, but it can be achieved in various ways that suit the system's mechanical and cooling arrangements. Undoubtedly, a holistic approach is necessary with careful evaluation of the device efficiency curves over the expected electrical and thermal operating conditions. What's also clear is that there is no 'one size fits all' approach when evaluating the optimum choice for fixed ratio conversion solutions, and careful appraisal is needed for each different application. Flex Power Designer (4) is a useful design software tool for power electronics engineers that can help to make this task easier and more effective.

---

[4] https://flexpowerdesigner.com

Connect with us

pm.info@flex.com

flexpowermodules.com

flexpowerdesigner.com

linkedin.com/showcase/flex-power-modules

twitter.com/flexpowermodule

flexpowermodules.com/wechat

youtube.com/flexintl

Flex (Reg. No. 199002645H) is the manufacturing partner of choice that helps a diverse customer base design and build products that improve the world. Through the collective strength of a global workforce across 30 countries and responsible, sustainable operations, Flex delivers technology innovation, supply chain, and manufacturing solutions to various industries and end markets.

For more information, visit flex.com.