electronic design
The Authority on Emerging Technologies for Design Solutions

Next-Generation Servers Require Robust Power Platforms

Electronic Design
Stephen Oliver
Stephen Oliver, Contributing Technical Expert
Mon, 2013-07-22 09:03

Computing-resource utilization is accelerating in part due to continuing exponential growth in mature applications that form the wide-area network (WAN) data-traffic baseline such as e-mail, search, and file sharing, plus newer uses like high-resolution media delivery, cloud storage, and cloud computing. Adding to the growing load are so-called big data applications that manipulate enormous data sets, often originating from geographically disparate data sources, and their analyses, which similarly must be available worldwide and often in real time.

Related Articles

- Tap Online Design Tools To Improve Your Power Subsystem Designs
- Move Thermal Issues To The Front Of Your Design Cycle
- 400-V DC Power Transformation Saves Upfront "Green" Supply Costs

Market forces drive incremental data-access and computing costs down with increasing use. In server environments, this means that processing efficiencies and densities must increase throughout the system-design hierarchy, from core to cabinet and beyond to the data center infrastructure.

To keep pace with these trends, power system densities must increase faster than IT load densities. The complete power architecture, from power entry to point of load, must deliver the highest attainable power efficiency, not only to keep energy costs in check but also to mitigate already difficult thermal design challenges and meet demanding uptime goals.

Older Web-based services largely deliver existing data with comparatively little computational load. Yet many newer services require much greater data manipulation to support complex image rendering and large-scale data processing, complicating data center designs. Server use then expands and demands heterogeneous computational resources to facilitate applications with differing hardware requirements.

Power To The Processors

Power requirements for server-class computing have evolved. While legacy blade designs make do with typical power budgets on the order of 225 W (aggregating processor, memory, storage, interfacing, etc.), contemporary high-workload designs can almost double that load to around 400 W largely due to significant increases in CPU/GPU processor and memory energy use.

Increasingly popular, high-voltage power distribution improves site energy efficiency by reducing current in distribution feeds and their corresponding I^2R losses. Power schemes such as the Factorized Power Architecture continue this theme by

allowing designers to route power on the blade itself at 48 V instead of 12 V, reducing onboard losses by as much as a factor of 16.

Advanced power conversion topologies such as those embodied by Vicor's PRM and VTM modules implement 48- to 1-V conversion with a single converter stage, delivering efficiencies as high as 93% including board losses. In so doing, these topologies virtually eliminate the bulk filter capacitance needs while delivering excellent load-step response and low ripple. This direct 48-V to point-of-load approach is a high power-density or current-density approach (A/in.^{2}), allowing designers to optimize the layout (or minimize "squares") at the system's highest current feeds, further reducing board losses.

Flexible, scalable power platforms are increasingly important as point-of-load voltages continue to fall and currents increase. While standard server processors from AMD and Intel may draw around 90 W at 1.2 V, more advanced graphics processing units (GPUs) from Nvidia draw approximately 120 W at 1.1 V and next-generation Intel Xeon Phi (Knights Landing) processors draw around 225 W to accelerate scientific and financial computing, data analysis, imaging, and complex modeling applications. While a 145-W voltage regulator at approximately 1.8 V feeds Intel's Haswell chip, though, this design uses on-board dc-dc converters to supply approximately 1 V to the processor core.

Paralleling processor roadmaps, supply voltages for DRAM are also falling with successive technology generations. DDR3 memory devices require 1.35-V and 1.5-V rails. DDR4 devices require a 1.2-V supply. While memory power efficiency has increased (DDR3 progressed from approximately ~1 Gbyte/W to approximately 3 Gbyte/W with DDR4 promising an additional 40% benefit), functional demand of memory per processor has risen so dramatically that total memory power draw is increasing.

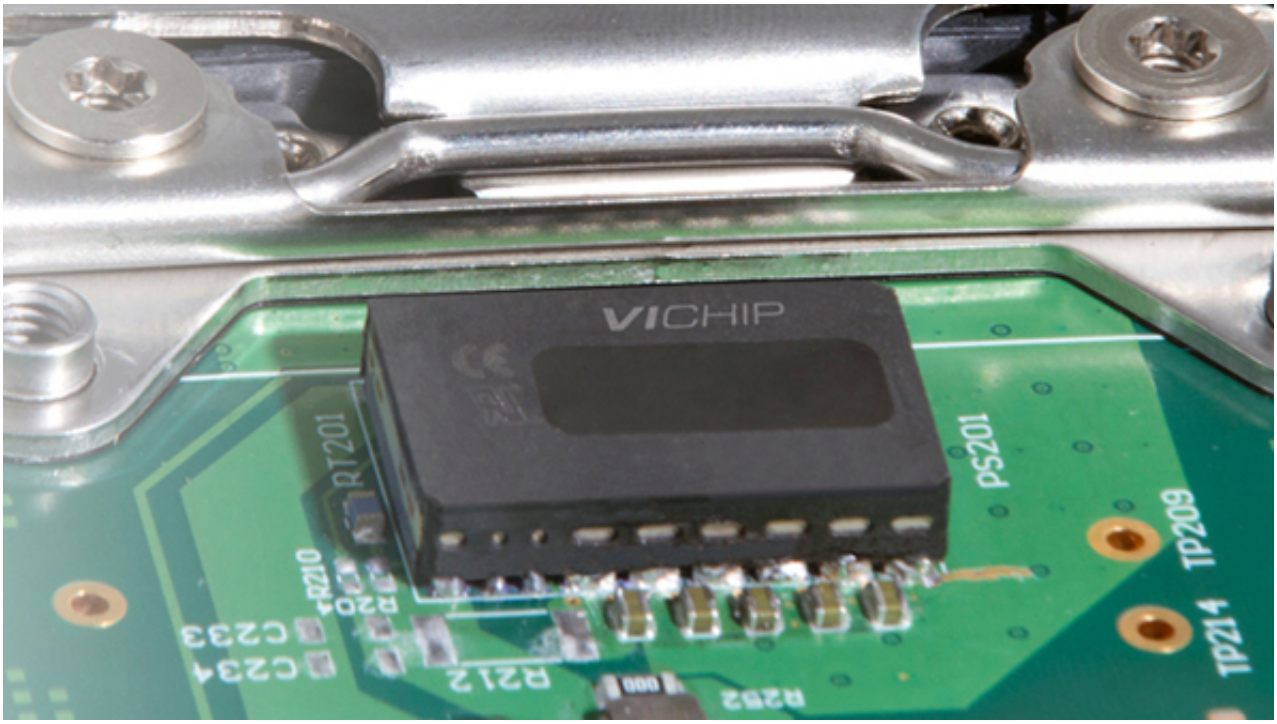
Powering The Rack

At the rack level, diverse workloads translate to a broad range of energy requirements, with associated power distribution, power conversion, and cooling challenges. The Open Compute Project (OCP), driven largely by Facebook, is developing a hardware specification for motherboards, switches, and cabinets that will result in a power draw of 14 kW per rack. At the same time, high-performance data analysis, rendering, and modelling applications require 30 to 50 kW per rack. Specialized supercomputing solutions have reached 100 kW in the same data center footprint. Here, questions of liquid cooling versus air cooling and 12-V or 48-V or 400-V/380-V dc power distribution and conversion must be addressed.

In all of these cases, advances in power packaging are pointing the way for such aggressive programs. For example, exploiting packaging such as Vicor's ChiP (Converter housed in Package) technology, designers can convert from 380 V dc to sub-1-V processors at an end-to-end efficiency of over 90% *(see the figure)*. The powertrain uses 380- to 48-V bus converters at 98% efficiency and greater than 2000 W/in.^{3} that can be placed within the rack on the backplane or even on the motherboard itself, eliminating traditional "silver box" power supplies and freeing valuable real estate for more computing performance. PRM regulator and VTM current multiplier modules provide direct 48-V-to-load conversion at greater than 200 A/in.^{2} at the processor socket, again opening more space for additional memory, interfacing, and storage.

Large-Scale Computing

Over the wide range of computing applications and along the complete powertrain from 380-V dc power distribution to power-hungry, sub-1-V processors, modular, scalable power architecture is required. High-efficiency, high-power-density power converters enable new levels of computing energy efficiency (MFLOPS/W) and compute density (MFLOPS/sq ft), allowing system designers to easily power new digital technologies to meet operators' evolving needs.



Stephen Oliver *is vice president of the VI Chip product line for Vicor Corp. He has been in the electronics industry for 18 years, with experience as an applications engineer and in product development, manufacturing, and strategic product marketing in the ac-dc, telecom, defense, processor power, and automotive markets. Previously, he worked for International Rectifier, Philips Electronics, and Motorola. He holds a BSEE from Manchester University, U.K., and an MBA in global strategy and marketing from UCLA. He holds several power-electronics patents as well.*

Source URL: http://electronicdesign.com/communications/next-generation-servers-require-robust-power-platforms