Data Center Challenges and Their Power Electronics

Philip T. Krein

***Abstract*—Data centers use more than 1.5% of all electricity in China and the U.S., with continuing growth. This paper reviews the power hierarchy levels within modern data centers. It considers energy consumption and power electronics challenges across all levels of a data center, including building distribution, dc architectures, and conversion down to the board level. Power electronics plays a central role throughout the hierarchy, and emerging approaches are described. Strategies that enhance center energy efficiency, both in terms of overall center operation and in terms of computation performance, are discussed.**

***Index Terms*—Data centers, dc distribution, dc-dc converters, server racks, power utilization factor, dc power architecture, digital power.**

I. Introduction

MODERN data centers are superficially similar to mainframe computer rooms from the 1960s, but have grown in scale and in number. They represent a fast-growing load component within the electric power grid. In [1], it is reported that in 2015, more than 400,000 data centers in China alone were consuming about 1.5% of national electricity supply. In the U.S., consumption in 2014 reached 70 TWh, about 1.8% of total supply [2]. Life-cycle costs of data centers today are often dominated by energy expense [3].

Although there is vast literature about overall data center performance, and large-scale industry efforts to enhance high-level efficiency [3], [4], the attributes associated with power architecture and power electronics across the layers in a data center are less widely discussed. This paper draws on [5], with updates across the hierarchy. Opportunities for improvements across the layers are summarized, and work in these areas is reviewed. The impact of power electronics is emphasized.

Based on the recommendation to "follow the power" [6], it is important to track energy flow in a data center from the external utility substation through building distribution to the individual racks, and then on to the server blades, memory and computing boards, and finally to the low-level information technology (IT) loads. Each level in this hierarchy brings its own challenges and opportunities for energy savings.

Manuscript received March 27, 2017. The work was supported in part by the Grainger Center for Electric Machinery and Electromechanics at the University of Illinois.

Philip T. Krein is with Grainger Center for Electric Machinery and Electromechanics University of Illinois, Illinois, 61801, USA (e-mail: krein@illinois.edu).

Digital Object Identifier 10.24295/CPSSTPEA.2017.00005

II. The Hierarchy Levels in a Data Center

Data center power can be treated according to levels of function:

- Utility level. Independent feeds, cogeneration, and external renewables are a few examples of challenges outside a data center itself.
- Building level. This is the level at which the primary utility feed enters the facility. Challenges include the choices between low voltage (e.g. 480 V or 600 V) input and medium voltage input, and whether the main building feed should be converted immediately to dc.
- Backup level. Many architectures have uninterruptible power supplies (UPS) and fuel-driven generators to provide power at the building level. Some distribute backup at lower levels.
- Rack level. Groups of racks are likely to use broader thermal management, but power delivery to individual racks is typical.
- Board level. Within a blade, power may be managed for memory boards, computing motherboards, and other local IT equipment.
- Chip level. Within a board, individual point-of-load (PoL) dc-dc converters provide low-voltage power for chips or chipsets. It is typical to see three to five different voltage values at this level of power management.
- Internal level. Some ICs, especially multi-core processors have on board energy management.

Each level involves particular power requirements and opportunities. Although an optimized system might couple attributes at several levels to achieve the best results, this paper will emphasize the individual levels and the challenges that each one entails.

A. Utility Level

At the utility level, many of the design aspects are linked to reliability. Reliability is linked to tiers [7], and at the highest Tier IV level, multiple independent utility feeds are typical. A data center might also be part of a microgrid. Some of the system architecture opportunities for power electronics include active switch management of multiple feeds [8] and integration into microgrids [9].

From a thermal perspective, the "energy" in the form of information leaving a data center is insignificant, and IT equipment ultimately converts 100% of incoming energy

into heat. This implies opportunities for combined heat and power (CHP or co-gen) utility-level architectures [10]. However, CHP has only been explored in a limited way for data centers. One issue is that the waste heat from IT devices is at relatively low temperature (rarely above 50°C) and therefore is not useful for steam generation. It can be used in district heat systems [11]. Following some early work [12], the trend has been to emphasize co-gen based on other devices such as microturbines [13]. A more practical situation locates a data center near a power plant that can also supply steam for absorption chillers.

As data center energy use continues to grow, numbers of data centers climb, and reliability demands increase, utilities are likely to have a larger role in siting and planning [14]. This has been typical in the past for large manufacturers, and data centers are approaching similar load levels.

B. Building Level

At the building level, there have been extensive studies that address the *power usage effectiveness* (PUE), defined as the ratio of total energy into a facility to total energy consumed by IT equipment within the facility. PUE must be greater than one under any circumstances and can readily exceed two given conventional heating, ventilation, and air-conditioning (HVAC) systems operating in hot environments. Realistically, PUE will exceed one substantially even if HVAC requirements are minimal, since a useful data center has lighting, security systems, and other non-IT functional loads.

An extreme low in PUE can be achieved with liquid immersion cooling in unstaffed data centers. In [15], a PUE value of 1.02 is reported for an insulating boiling immersion cooling approach. However, such a low number strongly suggests incomplete computation. For example, the power into a data center is delivered to power electronics, and only indirectly to IT equipment. The loss in power conversion is not really energy delivered to IT loads. Given an overall power conversion efficiency as high as 95% (unlikely), a PUE below 1.05 is not physically valid unless power supplies are considered to be IT load. This also ignores any energy consumed within uninterruptible supplies and even building switchgear and wiring.

Google reports a much more realistic company-wide PUE of 1.12 [16], and also reports there that

> "When measuring our IT equipment power, we include only the servers, storage, and networking equipment. We consider everything else overhead power. … Similarly, we measure total utility power at the utility side of the substation, and therefore include substation transformer losses in our PUE."

The distinction between IT and non-IT equipment is important when PUE values are used to incentivize system improvements.

Liquid immersion in data centers can take at least two forms. As in [15], the individual boards or blades can be immersed in coolant. Advanced hydrofluorocarbon liquids and mineral oils have been used for this [17] (claiming PUE values in the range of 1.02 to 1.03). The alternative is to immerse a complete rack or data center itself, possibly with conventional air-based heat transfer into the surrounding liquid. An undersea approach based on this has been announced by Microsoft [18].

Building-level emphasis on PUE has encouraged advanced HVAC work. Just a few examples include [19], in which enthalpy wheels are evaluated for data center benefits, and [20], in which rack-based cooling is proposed. In [21], a comprehensive overview of power architectures and interconnections to renewables is presented. The work also discusses power electronics architectures and cooling alternatives. Much of the discussion addresses UPS efficiency and operation.

In the context of power electronics, the discussion and use of direct dc distribution within data centers is a key trend. By the time of the workshop [6] held in February 2007, there were already some demonstration projects and active industry engagement. In [22], architectures are compared in depth across conventional ac distribution with conversion at the rack level, building-level 48 V dc distribution, building-level 400 V dc distribution, and others. The work identifies 400 V building-level dc distribution as advantageous. Several studies, including [23], [24] have reported that dc distribution at the building level enhances efficiency and also raises reliability and availability. It is important to recognize, however, that efficiency improvements are on the order of one percentage point or so with best-practice present designs.

Results from related areas, including mainframe computer systems, supercomputers, and mobile power systems have bearing on questions of dc distribution. A broad discussion of the issues and the power conversion requirements can be found in [25]. This work suggests that dc distribution was reported to have advantages even as early as the ENIAC mainframe system. In [26], architectures and power electronics designed for electric ship applications are suggested as a basis for dc applications in data centers. In [27], a supercomputer powered directly from a 380 V dc solar energy bus is described.

One important advantage of building-level dc distribution is that megawatt-scale rectifiers can be designed with much higher efficiency that rack-level kilowatt-scale devices [28]. Building-level conversion also simplifies reliability design. Redundancy at the building level is relatively straightforward, and it is generally accepted that single points of failure need to be avoided.

The trend toward dc distribution in data centers is part of a larger discussion of dc distribution within low-energy buildings [29]. Early efforts related to data centers and more broadly to buildings were conducted by Lawrence Berkeley National Lab. These activities have culminated in an international industry group, the Emerge Alliance [30]. The energy savings are one aspect, even if the amounts are not large [31]. At present, dc distribution in a data center is not confined to research. Commercial systems are in place [32].

It is important to notice that PUE in particular, and also dc distribution practices, do not by themselves drive power in-

novation in data centers. Even though PUE is a limited metric, it does give linkage to some efficiency improvements, particularly those associated with HVAC. Starting from relatively high PUE numbers (above 1.5), several strategies for energy savings have been suggested by Pouchet [33]:

- Optimize air flow throughout the building.
- Optimize air flow within each rack.
- Extra attention to building outer envelope and room sealing.
- Economizers in the heating, ventilation, and air-conditioning (HVAC) system.

A combination of cooling from ambient air and inexpensive energy has driven many data center locations [34].

C. Backup Level

In conventional practice, a UPS is provided at the building level, and represents an important element in overall efficiency. The basic electrical distribution scheme in a typical ac data center is shown in Fig. 1. Notice in particular the stack-up of power electronics. Just behind the building feeder, an online UPS (also termed a "double conversion" UPS) provides back-to-back conversion, with an input-side rectifier, regulated battery bus, and output-side inverter. This provides high-quality power conditioning and backup. At the rack level, ac may be connected to individual server blades, or an active rectifier supports a dc bus for the complete rack, usually at 48 V. At the board level, a relatively low voltage is used to send power across board traces. The final point-of-load (POL) converters supply the digital chip sets. Other POL converters produce supplies for communication links, disk drives, diagnostic devices, sensors, and other IT infrastructure. Building-level HVAC and lighting are not shown.

The advantage of an online UPS in this context is high power quality and excellent immunity to short-term changes on the grid side. Not shown is a transfer switch that can connect in local fuel-driven generation if the utility feed or feeds are lost. The disadvantages are that the UPS must have inherently high reliability, and that the multiple conversion stages involve power loss.

Offline or "standby" UPS circuits are also in wide use in data centers. Various hybrid configurations are employed to speed the process of transferring the energy source from the grid to a battery set [35]. Recent innovations include smaller intermediate converters that support the load during the transfer interval [36]. This can also be supported with passive filters combined with small converters [37]. Grid-interactive standby architectures allow fast source transfer without having the UPS in the continuous energy conversion path [38].

Notice that in the online UPS case, power is converted at least four times from the ac mains to the IC level. Given estimates such as a 97% efficient UPS process, a 95% efficient rack-level rectifier, and POL converters at typically 93% or less, overall about 85% of energy, at best, reaches the final IT loads. This would limit the PUE to more than 1.17 even without considering HVAC – provided IT power does not include power supply input.

A direct dc architecture, in which the "dc UPS" could be as simple as a master regulated dc bus, is shown in Fig. 2. In this case, the building rectifier efficiency exceeds 98%, the rack level might reach 97%, and the POL efficiency is still 93%. The total is about 88%, and the PUE will be more than 1.13.

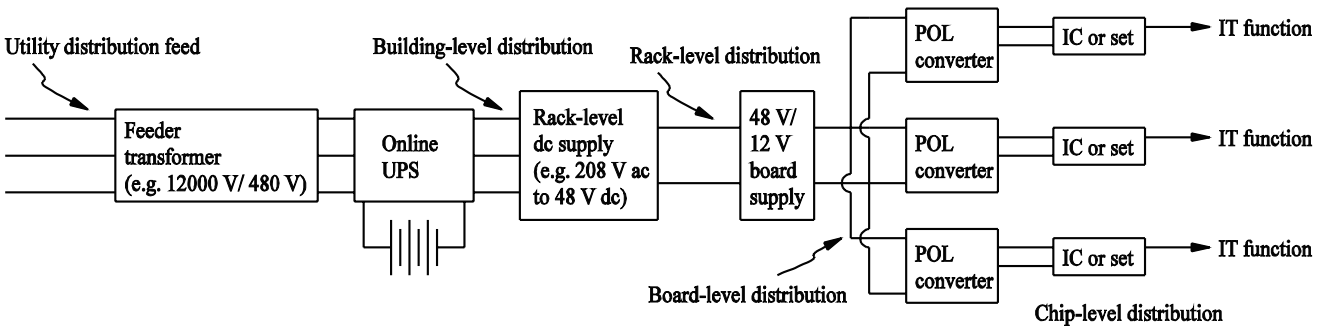

Fig. 1. Typical conversion sequence in an ac data center architecture [5].

Given these limitations on power loss and performance, how, for example, does Google achieve system-wide PUE of 1.12 [16]? Part of the answer is a *distributed UPS* architecture [38]-[40]. In the initial 2009 announcement [38], Google reported the use of small 12 V batteries directly within each individual server. The result has higher backup performance than a building-level online UPS, since it is distributed among all computing loads, but avoids power conversion layers built into conventional UPS architectures. Strictly speaking, the overall system efficiency is the same as in Fig. 2 because there is still conversion at the building level, rack level, and POL level, but the system is simpler and adapts better over a wide load range. The effective PUE improves by a few points. Many other users have adopted the distributed approach.

D. Rack Level

Servers in a data center are organized into conventional rack mounts, or in high-density blade configurations. A 42-unit (42U) rack is typical. In an ac data center, each rack might be provisioned with three or four three-phase circuit sets, provid-

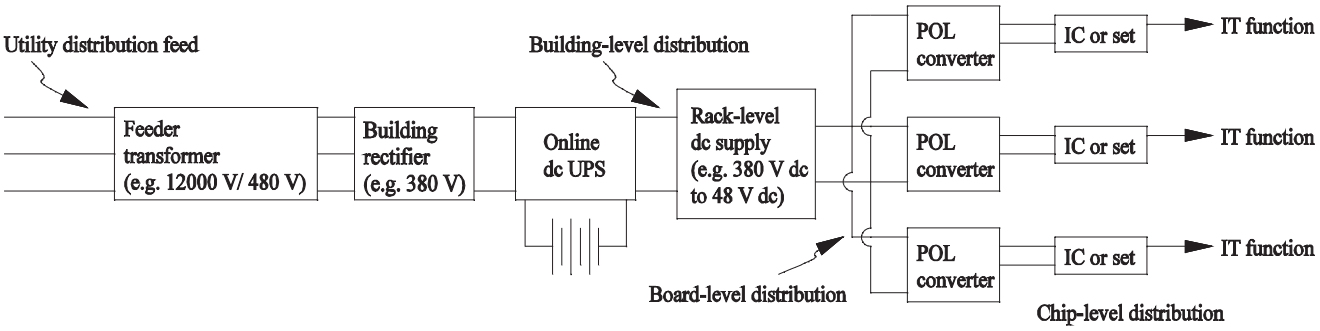

Fig. 2. Typical conversion sequence in a dc data center architecture.

ing up to 25 kW per rack. In a dc data center, nominal dc voltage such as 380 V is delivered directly to each rack, providing up to about 30 kW. In each case, power is often delivered directly to server-level power supplies.

An important challenge in rack provisioning is to manage various rating stackups. Standards and codes limit the loading level of any circuit. Servers are likely to have dual redundant supplies. The supplies themselves are oversized, and servers do not always operate close to 100% power. The combination generates considerable confusion when rack-level power provisioning is designed [42]. Worst-case design leads to extreme over-provisioning [43].

Rack distribution offers several opportunities for advanced power electronics. For example, supplies within a rack should be sequenced to avoid combined inrush peaks or other short-term system stresses. Faults should be managed to avoid propagation. Protection at the rack level is especially challenging with dc distribution [44]. Supplies must be efficient over a wide load range, given the limited full power operation of servers. Several approaches have been proposed to broaden the range of high efficiency [45]-[47], but rack-level power considerations are not quite the same as board-level issues discussed below.

The ongoing discussion about 48 V power for data centers [48] mainly considers board-level innovations. However, 48 V rack distribution is discussed [22], and direct board 48 V power would eliminate one power conversion level in the chain.

A more radical rack-level approach is to connect individual servers in *series* [49]-[51]. Because communications links are electrically isolated, alternative electrical connections can be supported. A series stack of server boards or blades eliminates a layer of conversion and substantially increases system efficiency [52], [53]. An oversimplified version of the approach would connect 32 servers in series, and connect them directly from a 380 V bus. This requires diligent load balancing, but avoids a conversion step and enhances system efficiency. In [54], it is shown that the approach can raise this part of the power conversion process from a state-of-the-art 96% efficiency to more than 99.8% efficiency. Some of the ideas appeared earlier [55], but the architectures have much in common with older battery balancing techniques [56]-[58].

Better integration of power and thermal management at the rack level has been discussed as a power saving tool [59], [60]. In [59], local rack cooling is linked to computing loads. The intention is to make total system power track the computational load. In [60], methods for server power interconnection within a rack, to reduce redundancy but maintain power reliability, are presented. The gains in energy performance from interconnection do not approach those of a complete series connection, but linkages between system power and computing power are important. This is linked to the concept of *energy proportional computing*, as described by Google [61].

E. Board Level

Individual server boards often employ one or two power supplies, known in the data center industry as *power supply units* (PSUs), within the server enclosure. Two are used when direct "1+1" redundancy is required. As in [38], many modern designs use battery backup directly at this level. As in [48], there is continuing discussion of 48 V supplies at the board level. Recent results [62] deliver 48 V directly to end-use POL converters.

At the board level, power supplies previously had various designations such as "silver," "gold," or "platinum," and there have been various industry initiatives for improvement [63], [64]. Wide ranges for high efficiency are more important than full-load efficiency [45], especially in redundant supply architectures in which the maximum output is 50% unless a failure occurs. In fact, dual redundancy benefits from high efficiency in the load range below 50% rather than higher.

Most innovations at this level have come from incremental improvements in power supplies. This is particularly true because most architectures employ a supply that delivers 12 V, 5 V, 3.3 V, and other voltages similar to those in a desktop computer. It draws power either from a single-phase ac input or a direct 380 V dc bus. Some of the work applies to rack-level power or to multiple boards (as in a blade enclosure) [65].

F. Chip Level

At the "chip level," the power province of POL convert-

ers, high efficiency is challenging. This is because voltages are low, currents are high, and additional requirements such as fast transient response trade off against low losses. Typical solutions use two-stage dc-dc converters with an initial 12 V to 5 V stage followed by a multi-phase 5 V to 1 V stage [66]. Other configurations can take advantage of high voltage conversion ratios [67]. Some broader concepts are addressed in [68], [69], in which the hierarchy issues are considered.

Some of the earliest work on POL converters considered 48 V input [70]-[72], even though implementation is much more recent [62]. In [70], the concept of multi-phase dc-dc converters is introduced. In [72], this is extended to a four-phase two-stage converter for 2 V output. Comprehensive work on design based on 48 V POL applications was presented in [73], [74]. The two-stage concept is presented in depth in [75]. Recent designs, such as [76], are mature and ready for applications.

In POL circuits and at the chip level, droop control methods can be employed for enhanced power management. Voltage droop (termed *adaptive voltage positioning*) seeks to keep POL output impedance resistive, such that increasing load drops the supply voltage linearly. This is a helpful innovation since finite output impedance is unavoidable, and resistive characteristics have advantages given dynamic disturbances. In digital loads, the effective load is resistive, since energy proportional to an internal capacitance C_{int} is lost during each digital clock cycle. This means the load power is

$$P_{load}=f_{clock}\,C_{int}\,V^2, \tag{1}$$

the same as replacing the digital load with an effective resistance R_{eff}. Voltage droop helps the stabilize operation against disturbances.

Adaptive voltage positioning or scaling can also refer to an adaptive process by which the supply voltage at the chip level is decreased as much as computational performance allows [77]. In an idealized version, a computational load runs a test protocol, and requests step-by-step reduction in supply voltage until excessive error rates are encountered. The chip itself might store a map of voltage vs. load to deliver error-free performance at the lowest possible power.

Frequency droop for digital loads is not a power supply strategy, but clock frequency adjustment or droop, as in (1), can manage load requirements in a dynamic manner. This is part of the industry concept of *dynamic voltage and frequency scaling* (DVFS) to rapidly adjust the power requirements of digital loads at the chip level [78]. The term DVFS appears in more than 1000 papers on IEEExplore. Summaries can be found in [79], [80].

G. Internal Level

At the internal level, within a chip, a recent trend is to bring dc-dc conversion into the chip itself. The general power-system-on-chip approach has been explored for a long time [81], [82], but commercial implementation is recent. The Intel "Haswell" processor concept was one of the first large-scale digital loads to incorporate internal power management [83], [84]. As digital supply voltages drop below 1 V, on-chip management becomes increasingly important.

Modern digital chips are likely to comprise many computation cores or internal modules. If these can be connected in other ways, such as in series [49], [55], large voltage savings become possible. For instance, a chip with 61 cores, connected in series, can operate directly from a 48 V bus at a nominal 0.79 V per core, eliminating all layers of point of load conversion. Such a connection does not eliminate benefits of DVFS, even though it involves more comprehensive system-level approaches to power conversion and management [85]. The series concept applies across many levels, with the highest benefits for core interconnections [86]. However, since electrical isolation is rare below the board level, series connected racks are likely to have more impact in the short-term than series connected cores.

The internal level is well-suited to certain types of software-based data center energy management. Individual cores can be switches on and off, or computing loads can be actively balanced among cores to best spread thermal stresses. A potential innovation is cores that report or even predict their power requirements back to on-chip power converters. The converters can anticipate requirements and make dynamic adjustments to deliver the best possible performance.

III. System-Level Management

Even though power distribution and conversion in a data center are inherently hierarchical processes, in the end the power is consumed by IT loads and combined management of hardware and software provides the greatest possible benefits. The PUE metric is less useful at this level. There are two aspects to consider:

1) Is a data center dynamically managed to keep its energy requirements as low as possible?
2) What about reduction of energy used per unit of computation?

For the first item, energy proportionality is important, but the complete system architecture also matters. For the second, the linkage between power electronics and digital design is important.

Strategies for keeping energy requirements as low as possible at the system level tend to be unique to an implementation. For example, for the highest-performing thermal designs, low energy consumption means good balance. A center operating at 10% of maximum design load in this case functions best if all internal elements also function at 10%. For power electronics, this is not typical, since there are overhead requirements that begin to affect efficiency as power decreases. For point-based thermal designs, typical practice is to run racks as close to 100% power as possible, setting other equipment to an idle state. This works well in systems designed for high full-load efficiency, but the high loading can diminish reliability and leads to hot spots and other potential failure mechanisms.

For reduced energy per computational unit, voltage reduction reduces this value in a nonlinear manner, but there are limits. Computation cores operating at near-threshold or sub-threshold conditions use much less power than cores running at conventional supply levels. However, power conversion down to 0.3 V to 0.5 V levels associated with this reduction are extremely hard to produce; series connections become important, even though they bring more complicated management issues. The clock rates also must be reduced, trading off performance in a multi-dimensional space. As pointed out in [87], the linkages between digital-side "low-power design" and system-level considerations in data centers are complicated.

In the large-scale computing literature, there is a growing body of work on system-level data center energy management. In [88], an early summary of management approaches is presented. In [89], [90], high-level simulation is used in an energy management process.

A critical observation is that energy saved at the point of end use, especially via reduction in energy required per unit of computation, provides extra dividends throughout a data center. Even if PUE is defined all the way at the point of final chip connection, with a value as low as 1.12, reduced consumption at the end offers extra benefits. In more conventional scenarios, in which PUE does not include power supply losses and impacts, reduced computation energy has much larger impact at the input. Dynamic optimization has a key role to play in long-term data center operation [91].

IV. Conclusion

Data centers are large energy consumers, analogous to heavy manufacturing industries. Both are key economic drivers, and energy costs are crucial in both arenas. Energy reduction in data centers requires attention to power distribution and conversion across all hierarchy levels. A challenging requirement is to combine hardware and software in operation to achieve system-level optimization. Power electronics plays a vital role in enabling intelligent system-level power management.

References

[1] L. Luo.(2015, March). Chinese Government calls for green data center catch-up. *DatacenterDynamics*. [Online]. Available: http://www.datacenterdynamics.com/content-tracks/design-build/chinese-government-calls-for-green-data-center-catch-up/93618.fullarticle

[2] A. Shehabi, et al., "United States data center energy usage report," Lawrence Berkeley National Laboratory, Berkeley, CA, Tech. Rep. LBNL-1005775, Aug. 2016.[Online]. Available: https://eta.lbl.gov/publications/united-states-data-center-energy

[3] S. Chen, S. Irving, and L. Peng, "Operational cost optimization for cloud computing data centers using renewable energy," *IEEE Systems Journal*, vol. 10, no. 4, pp. 1447-1458, 2016.

[4] M. Dayarathna, Y. Wen, and R. Fan, "Data center energy consumption modeling: A survey," *IEEE Communications Surveys & Tutorials,* vol. 18, no. 1, pp. 732-794, 2016.

[5] P. T. Krein, "A discussion of data center power challenges across the system," in *Proc. Int'l. Conf. Energy Aware Computing*, 2010.

[6] "'Follow the power,' a report on an energy efficiency workshop," EPRI/PSMA, Tech. Rep. 2007.

[7] U. Müller, K. Strunz, "Integrated reliability modeling for data center infrastructures: A case study," in *Proc. IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, 2012.

[8] M. Kim, A. Kwasinski, V. Krishnamurthy, "A storage integrated modular power electronic interface for higher power distribution availability," *IEEE Trans. Power Electronics,* vol. 30, no. 5, pp. 2645-2659, May 2015.

[9] A. Kwasinski, "Power electronic interfaces for ultra available dc micro-grids," in *Proc. Int'l. Symp. Power Electronics for Distributed Generation Systems,* 2010, pp. 58-65.

[10] V. Rudolf, et al., "Methodology for EBSILON simulation studies of on-site generation CHP systems for data centre," in *Proc. Int'l. Energy and Sustainability Conference* (IESC), 2016.

[11] F. Huang, J. Lu, J. Zheng and J. Baleynaud, "Feasibility of heat recovery for district heating based on cloud computing industrial park," in *Proc. Int'l. Conf. Renewable Energy Research and Applications* (ICRERA), 2015, pp. 287-291.

[12] Keith E. Herold, and R. Radermacher, "Integrated power and cooling systems for data centers," in *Proc. Intersociety Conf. Thermal Phenomena,* 2002, pp. 808-811.

[13] A. Escobar, J. C. Balda, C. A. Busada, and D. Christal, "An indirect matrix converter for CCHP microturbines in data center power systems," in *Proc. IEEE Int'l. Telecommunications Energy Conf.* (INTELEC), 2012.

[14] S. Everitt. (2014, May). Utilities and data centers do not mix – yet. *DatacenterDynamics*.[Online]. Available: http://www.datacenterdynamics.com/content-tracks/power-cooling/utilities-and-data-centers-do-not-mix-yet/86565.fullarticle

[15] Two-phase immersion cooling -- A revolution in data center efficiency. 3M news article.[Online]. Available: http://multimedia.3m.com/mws/media/1127920O/2-phase-immersion-coolinga-revolution-in-data-center-efficiency.pdf.

[16] Efficiency: How we do it.[Online]. Available: https://www.google.com/about/datacenters/efficiency/internal/

[17] D. Varma. Oil submersion cooling for today's data centers. *Green Revolution Cooling.*[Online]. Available: http://www.grcooling.com/wp-content/uploads/2015/06/GRC_WP-CLICK-Oil_Sub_DCc.pdf

[18] B. Cutler, S. Fowers, J. Kramer, and E. Peterson, "Dunking the data center," *IEEE Spectrum,* Mar. 2017.

[19] S. Mok, S. Kumar, R. R. Hutchins, and Y. K. Joshi, "Impact of a rotary regenerative heat exchanger on energy efficiency of an air-cooled data center," in *Proc. IEEE ITHERM Conf*., 2016, pp. 1182-1190.

[20] T. Gao, et al., "Innovative server rack design with bottom located cooling unit," in *Proc. IEEE ITHERM Conf*., 2016, pp. 1172-1181.

[21] X. Wang, M. Li, and Z. Wang, "Research and application of green power system for new data centers," in *Proc. IEEE Int'l. Telecommunications Energy Conference* (INTELEC), 2015.

[22] A. Pratt, P. Kumar, and T. V. Aldridge, "Evaluation of 400V DC distribution in telco and data centers to improve energy efficiency," in *Proc. IEEE Int'l. Telecommunications Energy Conference* (INTELEC), 2007, pp. 32-39.

[23] V. Sithimolada, and P. W. Sauer, "Facility-level DC vs. typical ac distribution for data centers: A comparative reliability study," in *Proc. TENCON, IEEE Region 10 Conference,* 2010, pp. 2102-2107.

[24] A. Kwasinski, "Evaluation of dc voltage levels for integrated information technology and telecom power architectures," in *Proc. 4th International Telecommunication - Energy special conference,* Vienna, Austria, 2009.

[25] D. J. Becker, and B. J. Sonnenberg, "DC microgrids in buildings and data centers," in *Proc. IEEE Int'l. Telecommunications Energy Conference* (INTELEC), 2011.

[26] A. Kwasinski, "Advanced power electronics enabled distribution architectures: Design, operation, and control," in *Proc. International Conference on Power Electronics/IEEE ECCE Asia,* 2011, pp. 1484-1491.

[27] J. Salazar. (2016, September). New Hikari supercomputer starts solar HVDC. Phys.org news.[Online]. Available: https://phys.org/news/2016-09-hikari-supercomputer-solar-hvdc.html.

[28] L. Schrittwieser, J. W. Kolar, and T. B. Soeiro, "99% efficient three-

phase buck-type SiC MOSFET PFC rectifier minimizing life cycle cost in DC data centers," in *Proc. IEEE Int'l. Telecommunications Energy Conference* (INTELEC), 2016.

[29] B. Nordman, and K. Christensen, "Dc local power distribution," *IEEE Electrification Magazine,* Jun. 2016, pp. 29-36.

[30] [Online]. Available: http://www.emergealliance.org/.

[31] P. Donovan. (2013, September). AC/DC: Highway to data center efficiency hell.[Online]. Available: http://blog.schneider-electric.com/datacenter/2013/09/09/acdc-highway-data-center-efficiency-hell/.

[32] D. Tuite. (2014, February). 400-V DC distribution in the data center gets real. *Electronic Design*. [Online]. Available: http://electronicdesign.com/power/400-v-dc-distribution-data-center-gets-real-0.

[33] J. Pouchet, "Data center power quality," in Power Sources Manufacturers' Association (PSMA), "Follow the Power — A Report on an Energy Efficiency Workshop from Generator to Integrated Circuit," PSMA/EPRI, Tech. Rep., Feb. 2007.

[34] S. Grundberg, and N. Rolander, "For data center, google goes for the cold," *The Wall Street Journal*, Sept. 2011.

[35] S. Karve, "Three of a kind [UPS topologies]," *IEE Review*, vol. 46, no. 2, pp. 27-31, 2000.

[36] M. Arias, M. M. Hernando, D. G. Lamar, J. Sebastian, and A. Fernandez, "Elimination of the transfer-time effects in line-interactive and passive standby UPSs by means of a small-size inverter," *IEEE Trans. Power Electronics*, vol. 27, no. 3, pp. 1468-1478, Mar. 2012.

[37] M. Arias Perez de Azpeitia, A. Fernandez, D. G. Lamar, M. Rodriguez, and M. M. Hernando, "Simplified voltage-sag filler for line-interactive uninterruptible power supplies," *IEEE Trans. Industrial Electronics,* vol. 55, no. 8, pp. 3005-3011, Aug. 2008.

[38] H. Tao, J. L. Duarte, and M. A. M. Hendrix, "Control of grid-interactive inverters as used in small distributed generators," in *Rec. IEEE Industry Applications Annual Meeting,* 2007, pp. 1574-1581.

[39] S. Shankland. (2009, April). Google uncloaks once-secret server. *CNet*.[Online]. Available: https://www.cnet.com/news/google-uncloaks-once-secret-server-10209580/

[40] V. Kontorinis, L. E. Zhang, B. Aksanli, J. Sampson, H. Homayoun, E. Pettis, D. M. Tullsen, and T. S. Rosing, "Managing distributed UPS energy for effective power capping in data centers," in *Proc. 2012 39th Annual International Symposium on Computer Architecture* (ISCA), 2012, pp. 488-499.

[41] S. Alanazi, M. Dabbagh, B. Hamdaoui, M. Guizani, and N. Zorba, "Joint resource scheduling and peak power shaving for cloud data centers with distributed uninterruptible power supply," in *IEEE Globecom Workshops,* 2016.

[42] How many amps are needed for a rack (48U) of servers? (2016, June). answered by C. Goolsbee. [Online]. Available: https://www.quora.com/How-many-amps-are-needed-for-a-rack-48U-of-servers.

[43] E. M. Landsman, "Scalable data center architecture for on-demand power infrastructure," in *Proc. IEEE Applied Power Electronics Conf.* (APEC), 2003, pp. 10-13.

[44] K. Tan, X. Song, C. Peng, P. Liu, and A. Q. Huang, "Hierarchical protection architecture for 380V DC data center application," in *Proc. IEEE Energy Conversion Congress and Exposition* (ECCE), 2016.

[45] Power-One, Inc., PFE1100-12-054NA 54 MM 1U PSU in the Server Environment. *Peritus Power white paper PFE1100-12-054N,* pp. 9-10. [Online]. Available: http://plugloadsolutions.com/documents/20101122_1U-54mm_whitepaper_9.pdf.

[46] J. W. Kim, D. Y. Kim, C. E. Kim, and G. W. Moon, "A simple switching control technique for improving light load efficiency in a phase-shifted full-bridge converter with a server power system," *IEEE Trans. Power Electronics,* vol. 29, no. 4, pp. 1562-1566, Apr. 2014.

[47] Y. Jang, and M. M. Jovanovic, "Light-load efficiency optimization method," *IEEE Trans. Power Electron.*, vol. 25, no. 1, pp. 67-74, Jan. 2010.

[48] R. Merritt, "Google, intel prep 48V servers," *EE Times,* Jan. 21, 2016.

[49] P. S. Shenoy, S. Zhang, R. A. Abdallah, P. T. Krein, and N. R. Shanbhag, "Overcoming the power wall: Connecting voltage domains in series," in *Proc. Int'l. Conf. Energy Aware Computing*, 2011.

[50] P. S. Shenoy, and P. T. Krein, "Differential power processing for DC systems," *IEEE Trans. Power Electronics*, vol. 28, no. 4, pp. 1795-1806, April 2013.

[51] J. McClurg, Y. Zhang, J. Wheeler, and R. Pilawa-Podgurski, "Re-thinking data center power delivery: Regulating series-connected voltage domains in software," in *Proc. IEEE Power and Energy Conference at Illinois* (PECI), 2013, pp. 147-154.

[52] E. Candan, P. S. Shenoy, and R. C. N. Pilawa-Podgurski, "A series-stacked power delivery architecture with isolated differential power conversion for data centers," in *Proc. IEEE Int'l. Telecommunications Energy Conference* (INTELEC), 2014.

[53] E. Candan, D. Heeger, P. Shenoy, and R. Pilawa-Podgurski, "Hot-swapping analysis and implementation of series-stacked server power delivery architectures," *IEEE Trans. Power Electronics,* to be published 2017.

[54] E. Candan, P. S. Shenoy, and R. C. N. Pilawa-Podgurski, "A series-stacked power delivery architecture with isolated differential power conversion for data centers," *IEEE Trans. Power Electronics,* vol. 31, no. 5, pp. 3690-3703, May 2016.

[55] S. Rajapandian, K. L. Shepard, P. Hazucha, and T. Karnik, "High-voltage power delivery through charge recycling," *IEEE J. Solid-State Circuits,* vol. 41, no. 6, pp. 1400-1410, Jun. 2006.

[56] S. T. Hung, D. C. Hopkins, and C. R. Mosling, "Extension of battery life via charge equalization control," *IEEE Trans. Industrial Electronics,* vol. 40, no. 1, pp. 96-104, Feb. 1993.

[57] G. L. Brainard, "Nondissipative battery charger equalizer," U.S. Patent 5,479,083, Dec. 26, 1995.

[58] C. Pascual, and P. Krein, "Switched capacitor system for automatic series battery equalization," in *Proc. IEEE Appl. Power Electronics Conf.* (APEC), Feb 1997, pp. 848-854.

[59] S. M. Yuan, C. Y. Chiang, M. J. Chen, and Y. L. Chen, "Automatic rack cabinet management system for data center," in *Proc. IEEE Int'l. Symp. Consumer Electronics* (ISCE), 2013, pp. 181-182.

[60] W. Kwon, and H. Kim, "Efficient server power supply configuration for cloud computing data center," in *Proc. Int'l. Conf. Computational Science and Computational Intelligence*, 2014, pp. 299-300.

[61] L. A. Barroso, and U. Hölzle, "The case for energy-proportional computing," *IEEE Computer*, vol. 40, no. 12, pp. 33-37, Dec. 2007.

[62] S. Jiang, X. Li, "Google 48 V power architecture," in Proc. *IEEE Applied Power Electronics Conf.* (APEC), 2017.

[63] C. Calwell, A. Mansoor, and R. Keefe, "Active mode power supply efficiency: key issues, measured data and the design competition opportunity," in *Proc. IEEE Applied Power Electronics Conference* (APEC), 2004, pp. 323-328.

[64] C. Calwell, and A. Mansoor, "AC-DC server power supplies: making the leap to higher efficiency," in *Proc. IEEE Applied Power Electronics Conference* (APEC), 2005, pp. 155-158.

[65] J. Aguillon-Garcia, and G. W. Moon, "A high-efficiency three-phase ZVS PWM converter utilizing a positive double-star active rectifier stage for server power supply," *IEEE Trans. Industrial Electronics,* vol. 58, no. 8, pp. 3317-3329, Aug. 2011.

[66] Y. Ren, M. Xu, K. Yao, Y. Meng, and F. C. Lee, "Two-stage approach for 12-V voltage regulator," *IEEE Trans. Power Electronics,* vol. 19, no. 6, pp. 1498-1506, 2004.

[67] J. W. Kimball, J.T. Mossoba, and P.T. Krein, "A stabilizing, high-performance controller for input series-output parallel converter," *IEEE Trans. Power Electronics,* vol 23, no. 3, pp. 1416-1427, 2008.

[68] C. Belady, and C. Malone, "Data center power projections to 2014," in *Proc. 10th Intersociety Conf. Thermal and Thermomechanical Phenomena in Electronics Systems*, 2006, pp. 439-444.

[69] X. Wang, M. Chen, C. Lefurgy, and T. W. Keller, "SHIP: Scalable hierarchical power control for large-scale data centers," in *Proc. 18th Int'l. Conf. Parallel Architectures and Compilation Techniques,* 2009, pp. 91-100.

[70] L. F. Casey, and M. F. Schlecht, "A high-frequency, low volume, point-of-load power supply for distributed power systems," *IEEE Trans. Power Electronics,* vol. 3, no. 1, pp. 72-82, Jan. 1988.

[71] B. A. Miwa, L. F. Casey, and M. F. Schlecht, "Copper-based hybrid fabrication of a 50 W, 5 MHz 40 V-5 V DC/DC converter," *IEEE Trans. Power Electronics,* vol. 6, no. 1, pp. 2-10, Jan. 1991.

[72] P. T. Krein, P. Midya, and U. Ekambaram, "A distributed low-voltage power converter," Univ. of Illinois, Urbana, Tech. Rep. UI-

LU-ENG-93–2563, 1993.

[73] M. Ye, P. Xu, B. Yang, and F. C. Lee, "Investigation of topology candidates for 48 V VRM," in Proc. *IEEE Applied Power Electronics Conference* (APEC), 2002, pp. 699-705.

[74] P. Xu, M. Ye, P. L. Wong, and F. C. Lee, "Design of 48 V voltage regulator modules with a novel integrated magnetics," *IEEE Trans. Power Electronics,* vol. 17, no. 6, pp. 990-998, Nov. 2002.

[75] Y. Ren, M. Xu, K. Yao, and F. C. Lee, "Two-stage 48 V power pod exploration for 64-bit microprocessor," in *Proc. IEEE Applied Power Electronics Conf.* (APEC), 2003, pp. 426-431.

[76] C. Fei, M. H. Ahmed, F. C. Lee, and Q. Li, "Two-stage 48 V-12 V/6 V-1.8 V voltage regulator module with dynamic bus voltage control for light-load efficiency improvement," *IEEE Trans. Power Electronics,* vol. 32, no. 7, pp. 5628-5636, Jul. 2017.

[77] V. Sharma, A. Thomas, T. Abdelzaher, K. Skadron, and Z. Lu, "Power-aware QoS management in Web servers," in *Proc. IEEE Real-Time Systems Symposium* (RTTS), 2003, pp. 63-72.

[78] M. Shojafar, C. Canali, R. Lancellotti, and E. Baccarelli, "Minimizing computing-plus-communication energy consumptions in virtualized networked data centers," in *Proc. IEEE Symposium on Computers and Communication* (ISCC), 2016, pp. 1137-1144.

[79] C. Isci, A. Buyuktosunoglu, and M. Martonosi, "Long-term workload phases: duration predictions and applications to DVFS," *IEEE Micro,* vol. 25, no. 5, pp. 39-51, Sept.-Oct. 2005.

[80] J. Luis Nunez-Yanez, M. Hosseinabady, and A. Beldachi, "Energy optimization in commercial FPGAs with voltage, frequency and logic scaling," *IEEE Trans. Computers,* vol. 65, no. 5, pp. 1484-1493, May 2016.

[81] A. J. Stratakos, S. R. Sanders, and R. W. Brodersen, "A low-voltage CMOS DC-DC converter for a portable battery-operated system," in *Rec.*, *IEEE Power Electronics Specialists Conf.* (PESC), 1994, pp. 619-626.

[82] M. Ludwig, M. Duffy, T. O'Donnell, P. McCloskey, and S. C. O. Mathuna, "PCB integrated inductors for low power DC/DC converter," *IEEE Trans. Power Electronics*, vol. 18, no. 4, pp. 937-945, Jul. 2003.

[83] G. Schrom, et al., "A 480-MHz, multi-phase interleaved buck DC-DC converter with hysteretic control," in *Proc. IEEE Power Electronics Specialists Conf.* (PESC), 2004, pp. 4702-4707.

[84] J. Hruska. (2013, May). Intel's haswell takes a major step forward, integrates voltage regulator. *HotHardware*. [Online]. Available: http://hothardware.com/news/haswell-takes-a-major-step-forward-integrates-voltage-regulator.

[85] E. Candan, P. S. Shenoy, and R. C. N. Pilawa-Podgurski, "Unregulated bus operation of server-to-virtual bus differential power processing for data centers," in *Proc. IEEE Applied Power Electronics Conf.* (APEC), 2017, pp. 1632-1639.

[86] P. T. Krein, R. H. Campbell, N. R. Shanbhag, "System and method for improving power conversion for advanced electronic circuits," U.S. Patent 9,116,692, Aug. 25, 2015.

[87] D. Meisner, and T. F. Wenisch, "Does low-power design imply energy efficiency for data centers?" in *Proc. IEEE/ACM International Symposium on Low Power Electronics and Design*, 2011, pp. 109-114.

[88] L. Benini, A. Bogliolo, and G. De Micheli, "A survey of design techniques for system-level dynamic power management," *IEEE Trans. Very Large Scale Integration (VLSI) Systems*, vol. 8, no. 3, pp. 299-316, June 2000.

[89] Y. Shi, X. Jiang, and K. Ye, "An energy-efficient scheme for cloud resource provisioning based on CloudSim," in *Proc. IEEE Int'l. Conf. Cluster Computing*, 2011, pp. 595-599.

[90] D. Meisner, J. Wu, and T. F. Wenisch, "BigHouse: A simulation infrastructure for data center systems," in *Proc. IEEE Int'l. Symp. Performance Analysis of Systems & Software,* 2012, pp. 35-45.

[91] C. Li, et al., "Towards sustainable in-situ server systems in the big data era," in *Proc. ACM/IEEE Ann. Int'l. Symp. Computer Architecture* (ISCA), 2015, pp. 14-26.

Philip T. Krein received the B.S. degree in electrical engineering and the A.B. degree in economics and business from Lafayette College, Easton, PA, USA, and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign, USA. He was an engineer with Tektronix, Beaverton, OR, USA, and then returned to the University of Illinois at Urbana-Champaign. He was a Senior Fulbright Scholar with the University of Surrey, Guildford, UK, from 1997 to 1998. From 2003 to 2014 he was a Founder and Director of SolarBridge Technologies, Inc., Austin, TX, USA, a developer of long-life integrated inverters for solar energy, and now part of Sunpower. He holds the Grainger Endowed Chair Emeritus in Electric Machinery and Electromechanics and is Director of the Grainger Center for Electric Machinery and Electromechanics at the University of Illinois at Urbana-Champaign. He is also Executive Dean of the Zhejiang University–University of Illinois Institute for engineering in Haining, China. He holds 38 U.S. patents with additional patents pending. His current research interests include all aspects of power electronics, machines, drives, electric transportation, and electrical energy, with an emphasis on nonlinear control approaches.

Dr. Krein is a Registered Professional Engineer in Illinois and Oregon. In 2001, he helped initiate the International Future Energy Challenge, a major student competition involving fuel cell power conversion and energy efficiency. In 2003, he received the IEEE William E. Newell Award in Power Electronics. He is a past President of the IEEE Power Electronics Society, and served as a member of the IEEE Board of Directors. He is Editor-at-Large of the IEEE Transactions on Power Electronics and an Associate Editor of the IEEE Journal of Emerging and Selected Topics in Power Electronics. In 2015-2016, he was Chair of the IEEE Transportation Electrification Community. He was elected to the U.S. National Academy of Engineering in 2016.