EMPower: The Case for a Cloud Power Control Plane

Jonggyu Park
University of Washington

Theano Stavrinos
University of Washington

Simon Peter
University of Washington

Thomas Anderson
University of Washington

ABSTRACT

Escalating application demand and the end of Dennard scaling have put energy management at the center of cloud operations. Because of the huge cost and long lead time of provisioning new data centers, operators want to squeeze as much use out of existing data centers as possible, often limited by power provisioning fixed at the time of construction. Workload demand spikes and the inherent variability of renewable energy, as well as increased power unreliability from extreme weather events and natural disasters, make the data center power management problem even more challenging.

We believe it is time to build a power control plane to provide fine-grained observability and control over data center power to operators. Our goal is to help make data centers substantially more elastic with respect to dynamic changes in energy sources and application needs, while still providing good performance to applications. There are many use cases for cloud power control, including increased power oversubscription and use of green energy, resilience to power failures, large-scale power demand response, and improved energy efficiency.

ACM Reference Format:

Jonggyu Park, Theano Stavrinos, Simon Peter, and Thomas Anderson. 2024. EMPower: The Case for a Cloud Power Control Plane. In *Proceedings of 3rd Workshop on Sustainable Computer Systems (HotCarbon'24)*. ACM, New York, NY, USA, 8 pages.

1 INTRODUCTION

For a sustainable computing future, the age of abundant power in the cloud is nearing its end. A rapid increase in cloud application energy demand due to artificial intelligence, as well as the tailing off of energy efficiency gains from Dennard scaling [7], has put power management at the center of cloud operations. In some regions, such as Northern Virginia and Ireland [10], data centers already draw more than 10% of grid power. That portion is projected to continue to grow, even relative to the increased supply needed to support decarbonization of transportation and building heating and cooling.

As a growing and already major power consumer, data centers will increasingly have to balance fluctuating power demand and supply. In this scenario, there are three primary challenges for data center power management (Figure 1a):

(1) Solar and wind power—needed to support increased energy use by data centers, vehicles, and homes [24]—has volatile swings in power supply [65].
(2) Power infrastructure is a significant capital expenditure; operators increasingly oversubscribe power to lower costs [41, 62, 70]. However, oversubscription reduces data center resilience to demand spikes.
(3) Extreme weather events lead to blackouts and brownouts [16]. Data centers have limited power reserves to be resilient to such situations, but they deplete quickly if power demand is not managed well.

To address these challenges, data centers need to react quickly to changes in power demand and supply while respecting application SLAs, power infrastructure capacity, and grid stability. To this end, we propose to build a *power control plane* for cloud data centers. EMPower (*E*lastic *M*anagement of *Power*) would observe and control power demand at a fine granularity and over short timescales (on the order of seconds) by making it *software-defined*. The key is to *gracefully trade off* power, performance, and application quality of service (QoS) over time. Our approach leverages the fact that application QoS requirements often allow for slack. This slack will allow EMPower to conserve power during a power event by shedding and consolidating load, power-switching hardware components, and migrating critical workloads to less power-intensive processors, within QoS parameters. Meanwhile, non-critical load would be shifted to times with ample power supply (cf. Figure 1b).

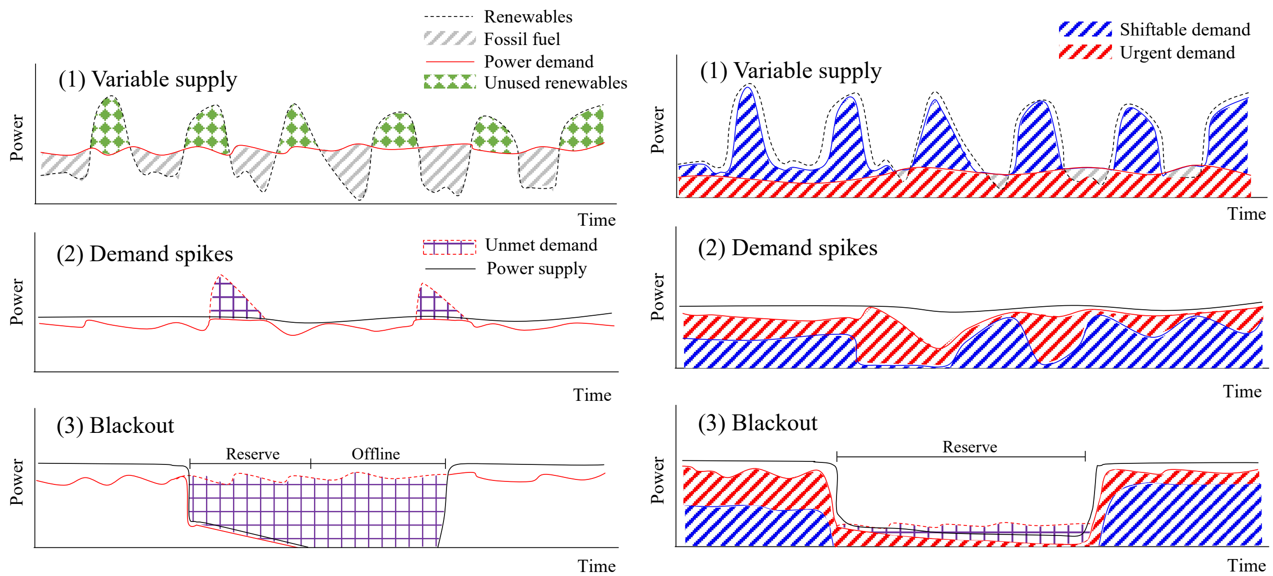
Existing methods for addressing power-related challenges have been conservative, offering a narrow control range [38, 41, 43, 62, 70]. For instance, Google introduced a hardware-agnostic power capping system named Thunderbolt [43] that aims to reduce QoS violations while safely allowing power oversubscription. Thunderbolt regulates CPU power draw by either limiting bandwidth or deactivating cores, balancing QoS with available power. However, by focusing on CPUs, such systems support only a small power control dynamic range. Moreover, power attribution is too coarse-grained to accurately determine how applications draw power. Similarly, application QoS is often specified at the relatively coarse granularity of a virtual machine. As a result, it is challenging to determine which application loads to control and by how much.

To push data center power control well beyond existing capabilities, EMPower will incorporate several novel power-saving mechanisms and policies by leveraging the capabilities offered by emerging development models and modern hardware. For example, disaggregated memory presents a unique opportunity to decouple application state from heterogeneous compute cores with minimal overhead. EMPower can leverage disaggregated memory for aggressive consolidation of compute across servers and accelerators, while shutting down unused components to expand the power control dynamic range. These selections will be made in real time, guided by our policies.

To realize EMPower, we require a hardware/software co-design of next-generation data centers, workloads, and their run-time systems. We identify five challenges to realizing EMPower, which form the basis of a research agenda for power-adaptive cloud systems:

(1) An effective power control plane must scale to include most of a data center's hardware and applications.

HotCarbon'24, July 9, 2024, Santa Cruz, CA



(a) Without EMPower. **(b) With EMPower.**

Figure 1: Data center power management challenges.

(2) There is currently no mechanism for applications to convey fine-grained service-level agreements (SLAs) to operators, forcing operators to be conservative when deciding how to respond to power demand.
(3) Power control policies must be automatic and robust over both long and short timescales (seconds or less).
(4) The range of power-controllable hardware devices must be expanded to unlock the full power control dynamic range available in data centers.
(5) Power instrumentation and control mechanisms today operate at the wrong granularities, making it difficult to identify opportunities for efficiency and to respond to fluctuations in power availability.

We expect that EMPower will dramatically improve the energy efficiency of data centers, enable more renewable energy use, reduce the time to recover from power outages, and allow data centers to outlive power disruption events by leveraging software-defined power control. For example, EMPower may allow individual cloud data centers to handle more load by enabling further oversubscription of available power beyond what can be safely achieved today. EMPower's power instrumentation insights would allow developers to focus on code debloating to improve software energy efficiency. By quickly shedding load and power-switching associated hardware resources, EMPower aims to make data centers resilient to power supply variability, including power disruption and green power availability. Finally, EMPower aims to keep critical applications in operation in a power crisis and gracefully reduce data center power demand during supply shortages.

2 CLOUD POWER CONTROL: WHY NOW?

With the commercial success of internet-scale applications and cloud computing, cloud infrastructure has grown rapidly. Estimates place data centers as responsible for 1–2% of aggregate worldwide electricity consumption [36, 59] and project that data center power draw will grow to 10% of global electricity use by 2030 [36, 46, 51]. In many power grids, data centers are already major load contributors. For example, in Northern Virginia, data centers account for 12% of power draw (2022), and are predicted to reach 22% in 2032 [20, 21]. In Ireland, data centers account for 14% of national electricity use (2022) [10] and may be 30% by 2029 [23]. In response, the Ireland national grid manager recently canceled more than 30 planned data center projects to preserve the stability of the grid [37]. These are just the leading edge. With continued cloud and artificial intelligence growth, the power draw of data centers is expected to be a large factor for many regional grids [31, 36, 51]. Models suggest many grids will not be able to meet datacenters' projected power demand growth; building new datacenters in certain regions may only be possible with degraded availability [45]. A consequence of this rapid growth is that data centers will need to operate under tight and variable power envelopes to be allowed access to grid power.

Due to the high cost of provisioning peak power, some hyperscale cloud data centers already oversubscribe their power infrastructure [41, 62, 70]. With oversubscription, more servers are placed on a circuit than can be fully powered at peak load simultaneously. To prevent overwhelming power infrastructure, providers deploy power-capping systems to automatically shed non-critical load in overload situations [43]. These generally are designed to make adjustments within a small dynamic range.

However, power demand and supply variability can occur suddenly and with large swings. For example, Google observed a 30× increase in compute demand for some applications during the first quarter of 2020 due to the pandemic-induced spike in home-office use [9]. Provisioning enough power to fulfill the demand of newly deployed servers to handle the spike was a major challenge. On the supply side, renewable energy is becoming a primary power source [2, 24]. Wind and solar installations have large swings in power production around their nominal generating capacity [65]. Even without renewables, an increase in natural disasters has led to more blackouts and brownouts; observed grid failures worldwide are 4× above IEEE expectations for commercial power systems [25], and failure frequency is trending upward [16]. The problem may also become self-made: as the largest data centers become increasingly power proportional, large load swings [44] introduce the possibility of grid-destabilizing power demand changes.

3 CHALLENGES

Power variability has traditionally not been a focus for system designers. Consequently, existing systems have a small dynamic range of power control, as well as coarse-grained instrumentation and load control. Challenges include high idle power draw in servers, power instrumentation only at the chassis and CPU socket levels, load control at the virtual machine level, and missing integration with accelerators, such as GPUs and SmartNICs. We describe these challenges in this section and explain how they make it difficult to support efficient power control at scale. Building a power control plane requires us to overcome them.

Limited power control dynamic range. Cloud hardware traditionally has a small dynamic range for power draw. Server power control features, such as dynamic voltage and frequency scaling (DVFS) and running average power limit (RAPL), allow only limited control over CPU, GPU, and memory power [11, 42, 58]. Even when no application is running, cloud servers draw a large amount of idle power that cannot be controlled with RAPL or DVFS. We measured idle power on a variety of servers, including on CloudLab [19]. Servers' idle powers ranged from 58 to 220 Watts. For servers with GPUs, idle power can be as high as 600 W. For EMPower to be effective, we believe it is necessary to make servers more power-proportional—*i.e.*, more efficient at any utilization—to increase the data center's power control dynamic range.

Coarse power instrumentation granularity. Currently, power instrumentation exists primarily at a coarse granularity, such as the full chassis through IPMI [33] or at the CPU socket level through RAPL [57]. However, in modern cloud environments, resource multiplexing is an essential mechanism for improving resource utilization. Existing methods for measuring power cannot attribute power draw to individual applications or processes multiplexed on the same hardware. Per-application or per-process power measurements thus remain elusive, making it difficult to identify inefficiencies in software and to fully realize elastic power control. At the same time, fine-grained power instrumentation must be handled carefully, as power draw is a common vector in side-channel attacks [47, 72].

Coarse power control granularity. Today's power control mechanisms are often coarse grained, limiting their utility. For example, while fine-grained and elastic microservice development models are emerging, many services still run in heavy-weight virtual machines (VMs). Shedding load involves shutting down entire VMs, and migration involves moving an entire VM's state among servers, which can take minutes [1]. We have to provide lighter-weight load control for VMs and containers to enable finer-grained, per-microservice power control.

Limited integration with accelerators and IO devices. Accelerators like GPUs offer limited software support for power instrumentation and control [58]. IO devices such as NICs and storage drives may have no established mechanisms at all. Understanding and controlling power in accelerators and IO devices is important for two reasons. First, such components contribute significantly to the power draw of servers [22, 52]. Second, accelerators, especially GPUs, contribute an increasing amount to the overall energy consumption of a data center. Integrating these devices into power control decisions is thus important for increasing the data center power control range.

Scalability. A perennial challenge of data center infrastructure design is scale. A data center power control plane must process information from power instrumentation and actuate power control over millions of heterogeneous processors and accelerators, applications, and hardware devices. It must do so in a timely manner, without violating SLAs and under bursty power budgets. For EMPower to be successful, we must design it from the ground up to be scalable, with low-overhead and low-latency measurement and actuation mechanisms, down to OS- and VM-level CPU, accelerator, memory, and IO scheduling, socket allocation, and process/VM assignment.

4 A CLOUD POWER CONTROL PLANE

We propose to build EMPower, a scalable cloud power control plane. EMPower will feature operating system mechanisms for power instrumentation and control that realize fine-grained and scalable power control policies. EMPower will address the challenges outlined in Section 3 by integrating

(1) server shutdown to widen the available power control dynamic range,
(2) fine-grain power instrumentation via performance counters and power models down to the process and procedure call level,
(3) novel OS mechanisms to provide fine-grained power control for microservices via modern hardware/software interfaces,
(4) new cloud infrastructure stacks that support low-power processing, and
(5) hierarchical power instrumentation and control that can operate at scale.

EMPower's proposed design is illustrated in Figure 2. Each physical server maintains power draw records. An EMPower controller collects this data through the network hierarchy, along with a total power budget from the grid. With this aggregated information, along with application SLAs from cluster schedulers, the EMPower controller establishes finer-grained power budgets and disseminates them via the network hierarchy. Switches may subdivide their budgets hierarchically, taking into account the power budget

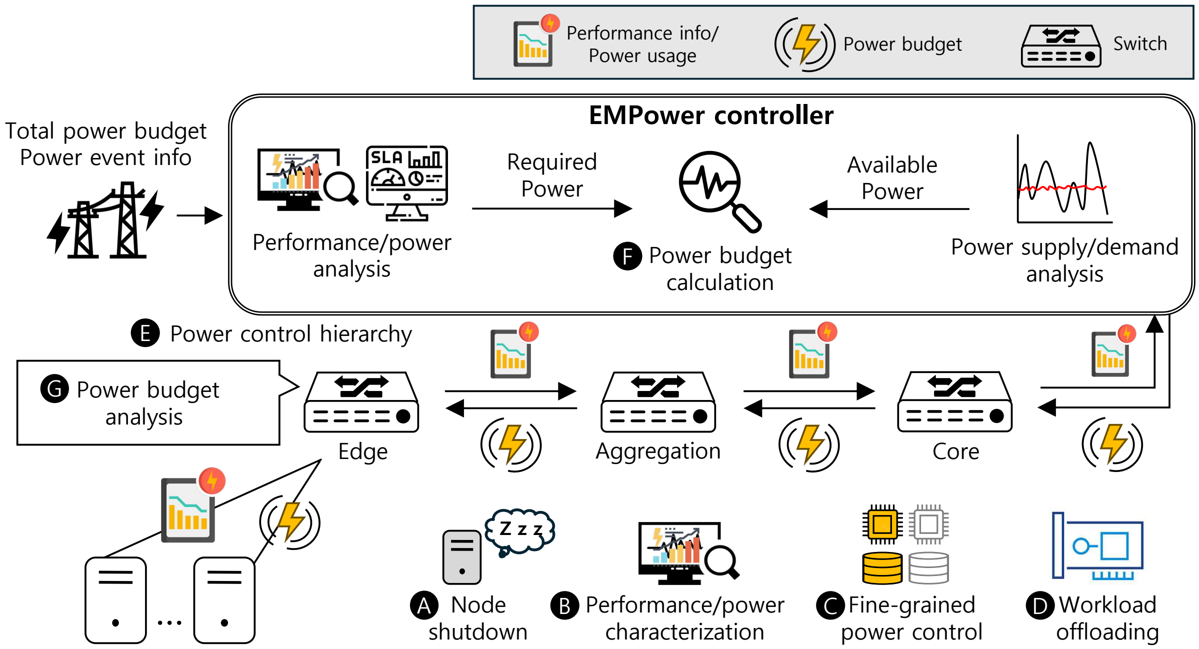


Figure 2: EMPower overview.

and workloads. Finally, the servers control process and VM load and implement power-saving mechanisms, such as node shutdown. We detail our approach in this section.

Large power control dynamic range via server shutdown. EMPower will remotely shut down and start up entire servers via built-in board management controllers (A) to reduce unnecessary idle power (*e.g.*, with the intelligent platform management interface (IPMI) [33] or Redfish [18]). To start and stop servers without disrupting services, we plan to leverage disaggregated memory to store virtual machine, operating system, and process snapshots with techniques such as background compressed write-through. Disaggregated memory, enabled by technology such as Compute eXpress Link (CXL) [15], allows low-latency load/store access to a server's shared memory pool even after the server is powered down, by leveraging its own independent power source. The additional power required to support disaggregated memory is marginal, as many servers can attach to a shared memory pool. This increases the flexibility in scheduling workloads across machines [14, 68] and enables rapid power demand adjustments. By reducing idle power, EMPower can significantly increase the power control dynamic range.

Fine-grained power instrumentation. To effectively control power, we need to understand how application software draws power at the level of applications, processes, and even remote procedure calls. To meet the requirement, we propose power draw models, APIs, and instrumentation tools (B) that account for application-level power draw across all relevant data center hardware, as well as across the entire software stack. This accounting would be similar to the *perf* tool [60], which profiles applications' CPU usage for performance debugging. Our instrumentation would report power draw at a similar level of detail, by instrumenting performance counters and leveraging models to translate performance to power draw. The fine-grained power instrumentation tool would use machine learning techniques, accounting for the electrical characteristics of modern CPU architectures, to estimate the power draw of various computing entities, including processes and containers. Power prediction models would enable EMPower to use the tool for power adaptive container scheduling and more efficient power capping. For security, EMPower will maintain this information and it will be accessible only to cloud operators.

Coordinated and fine-grained power control via modern hardware interfaces. To realize fine-grained power control, we must control CPU, IO, and memory load at a per-process level. Hardware/software support for fine-grained load control is increasingly available. Interfaces such as Intel's memory bandwidth allocation architecture [26] and cache allocation technology [54] expose control of per-process memory bandwidth and cache utilization, while modern NICs [32] and SSDs [55] can limit IO bandwidth to control IO load at a fine granularity. As hardware becomes more power-proportional, it is important to utilize these knobs. Unfortunately, current operating systems do not exploit such hardware-provided load control mechanisms in concert. We aim to develop OS policies and mechanisms to leverage these load control techniques to manage server power draw.

A practical illustration of this approach is to strategically confine applications to fewer server sockets, while commensurately limiting other resources (C), to reduce power draw when permitted by the workload. For example, EMPower may reduce the power draw of multiple servers by shutting down one socket, reducing the number of active memory and SSD channels in tandem with the reduced compute load, to balance system resources and reduce idle power drawn by these resources. When a workload is memory intensive, EMPower may decide to leave memory bandwidth at full capacity to allow the workload to finish within its SLA.

Cloud infrastructure software stack for low-power processing. We plan to redesign the cloud software infrastructure stack with support for low-power modes. For example, to be able to shut down servers (A), cluster managers need to treat server shutdown as a new operational mode, distinct from a failure. To use low-power processing, microservice runtimes need to support transparent migration of applications to low-power processors, such as SmartNICs (D). Similarly, disaggregated cloud services, such as storage, need to support these processors. Existing work develops prototype runtimes and storage services leveraging low-power options (e.g., iPipe [48] and E3 [49] for microservices, and LineFS [40] for storage). We plan to extend them to support low-latency application and service migration when power budgets change. We also plan to redesign many other cloud services, such as network communication, locking, and load balancing, to support low-power operation.

Leveraging hierarchy for power instrumentation and control scalability. Hierarchical aggregation and budgeting (E) would enable fine-grained power measurement and actuation to scale. EMPower would extend existing network control planes to aggregate power measurements and SLA information at the server, rack, and pod (set of racks) level, leveraging programmable switches in the data center network topology [4] to do so at microsecond granularity. Where power distribution does not correspond to the network structure, EMPower would use a virtual network hierarchy. The aggregated data would be forwarded to a central power controller that takes power supply measurements, executes a global power control policy, and then distributes power budgets back to switches and servers for effective power actuation, local to each server.

In detail, in our proposed design EMPower calculates the available power budget considering both power supply and demand. Meanwhile, EMPower also estimates the required power to guarantee application SLAs using the SLA and power information. Using the available power budget and required power, EMPower determines per-pod power budgets (F) and delivers them to the corresponding switches through the data center network hierarchy. Pod-level switches in turn subdivide and deliver the budget to racks. Rack-level switches finally determine per-server budgets, depending on the power budget and performance information (G). Individual servers execute power-saving mechanisms to meet their power budgets; OSes and hypervisors may further subdivide the server budget to per-socket, per-core, and per-application budgets.

5 USE CASES

There are many use cases for cloud power control, including increasing power oversubscription, improving resilience to power failures, implementing large-scale power demand response, improving energy efficiency, and preferring use of green energy.

Power oversubscription. Demand for cloud resources is increasing. One solution to this increased demand is to build more datacenters at significant delay and cost. An increasingly common alternative is to oversubscribe the data center's power infrastructure by provisioning more servers than the power infrastructure could support if those servers were running at full utilization [43]. Oversubscription deploys more compute resources by leveraging the fact that the actual power draw of data centers is often less than the maximum power draw of the deployed equipment [62]. However, oversubscription can threaten to push a data center's power draw beyond its supply unexpectedly if there are spikes in demand.

Continuously controlling the load and power demand to prevent overloading power equipment is handled by power capping systems [41, 43, 70, 71]. Compared to existing systems, we anticipate that EMPower will widen the power control dynamic range by shutting down servers and migrating workloads to low-power processors, as well as providing finer-grained control over power demand and supply. These capabilities should allow EMPower to increase data centers' ability to oversubscribe power infrastructure and control power demand much closer to the given power supply envelope.

Power resilience. Data center power supply is increasingly threatened by blackouts and brownouts from natural disasters—in particular climate events, which are becoming more frequent due to global warming—and failures of aging grid infrastructure [16, 17, 25]. This trend is especially salient for edge data centers which often receive power from only one utility [53]. EMPower could handle such events by keeping track of energy reserves, such as batteries and generators, and shedding an adequate amount of non-critical load to fit the power budget.

Power demand response. Demand response refers to adjusting power draw in response to changes in the power supply. Effective demand response has financial benefits and offers power resiliency. For instance, grid operators may increase the cost of power to incentivize reduced power when renewable energy makes up a small share of the energy mix or when the grid is particularly strained. Supporting power demand response is poised to become table stakes for new data center deployments. For instance, the Irish government has expressed a preference for energy-efficient and carbon-conscious data center developments [35]. Included in the preference are techniques to adapt to variable grid demands on power draw. By degrading non-critical load to reduce data center power draw at times of extreme grid-wide power demand, EMPower could support demand response to enable the deployment of data centers that preserve grid health and reduce operational costs.

Energy efficiency. Many cloud APIs layer inefficient implementations, wasting energy. EMPower could help debug software inefficiencies by providing a granular accounting of the power draw of individual components of the application stack. EMPower would profile power usage to help developers identify code and application architectures that may be streamlined. The power profiling information supplied by EMPower could also be used to compare system designs for energy efficiency.

Low-carbon computing. The computing industry is growing faster than green energy sources can be brought online. To date, the price of emitting carbon is too low to drive data center carbon efficiency. This will change over time and the tools we build for managing power draw can be used for minimizing data center carbon. For example, common low-carbon computing problems are to handle power supply swings caused by renewable energy and to estimate the carbon emissions of servers [3, 66]. EMPower aims to natively address power volatility (§2), and its power draw measurements could support carbon emissions estimation. As power control planes are adopted to address urgent power control problems, they can also facilitate the longer-term goal of low-carbon computing.

6 RELATED WORK

Power capping. Power capping systems [43] enable oversubscription [41, 62, 70] of data center power infrastructure. The key criterion for such systems is to uphold QoS guarantees, shedding load that would exceed a predefined power envelope. Further, server overload control [13] can preserve latency targets in non-power-capped scenarios. EMPower's proposed design builds on this work to enhance the power control dynamic range and provide elastic power control beyond oversubscription, including supply variability, resilience, and demand control.

Intermittent computing. Many system-level hardware and software techniques address continued operation under variable power supply, for instance for Internet-of-Things devices, edge computing, and energy-harvesting computing environments. The most extreme scenario is intermittent computing, where only a fraction of peak power is available for extended periods [52, 64, 69]. Some techniques include fast snapshotting and restoration during low-power events and the use of heterogeneous hardware with different power profiles. We adapt some of this work to operate beyond the server scale. In particular, EMPower will combine OS and networking techniques to enable fast reaction and control at rack scale and beyond.

Resource disaggregation. Resource disaggregation is a recent hot topic in the systems and architecture communities [28, 29, 63], supported by emerging hardware and protocols [15, 56]. Its primary purpose is to pool hardware resources, including memory and storage, to enable more efficient sharing and to raise utilization. We plan to build on recent resource disaggregation support for fine-grained power control, such as shutting down a server chassis, while keeping pooled memory online.

Power proportionality. Power proportionality is a requirement of EMPower, since it increases the power control dynamic range of the data center [50]. Existing work targets power proportionality on single servers, for instance using dynamic voltage and frequency scaling (DVFS) and hardware sleep states [5, 8, 34, 58, 61]. However, DVFS provides only a small power dynamic range [42], and these mechanisms are typically only scoped to a single server. Other work aims for system-wide power proportionality, i.e., across multiple nodes, often by leveraging the different power profiles and capabilities of heterogeneous hardware [6, 12, 50, 52]. We plan to enhance data center power proportionality by integrating low-power processors (e.g., SmartNICs [40, 49]), server hibernation, and fine-grained instrumentation and control.

Energy attribution. There have been several research efforts to measure power draw in cloud servers and applications [27, 30, 39, 67]. For example, EnergAt presents a thread-level, NUMA-aware energy attribution for CPU and DRAM in multi-tenant environments [30]. However, EnergAt uses up to 10% of an application's energy to determine its energy consumption, which is too high for continuous use, such as in EMPower. Other systems use performance counters, accessed through hardware interfaces or *perf* [60], to estimate the power draw at a container level [27, 39, 67]. However, such event monitoring from VMs is unavailable in cloud settings since it is a privileged task. To support monitoring in a cloud setting, systems often rely on CPU occupation time, which is inaccurate, or a customized hypervisor to estimate container-level CPU power draw. Moreover, the related work focuses on CPU and memory power draw without considering other components, including accelerators and peripherals. EMPower would collect power draw information beyond CPU occupation time, aiming to allow for precise attribution of energy to applications in a cloud computing environment. EMPower would avoid the security risks of existing approaches by supplying applications with appropriate power data aggregates, limiting the attack surface of side-channels.

7 CONCLUSION

We underscore the critical need for a power control plane in cloud data centers, driven by the end of Dennard scaling, rising power costs, increased use of renewables, increased extreme weather events, and sudden power demand surges. We propose EMPower to provide fine-grained, scalable control over data center power use, aiming to enhance data center elasticity in response to dynamic changes in energy demand and supply. New technologies, including disaggregated memories, low-power compute devices, programmable switches, and fine-grained development models, open the opportunity for EMPower. We envision several use cases for cloud power control, including increased power oversubscription, use of green energy, resilience to power failures, and improved energy efficiency.

Acknowledgments. We thank the anonymous reviewers and our shepherd, Romain Jacob, whose feedback substantially improved the paper. This work is supported by National Science Foundation grants CNS-2104548 and CNS-2148209 and the University of Washington Center for the Future of Cloud Infrastructure (FOCI).

REFERENCES

[1] Improving azure virtual machine resiliency with predictive ml and live migration, 2018. https://azure.microsoft.com/en-us/blog/improving-azure-virtual-machine-resiliency-with-predictive-ml-and-live-migration/.
[2] Canada's energy industry technology trends in 2023, 2023. https://news.microsoft.com/en-ca/2023/01/18/canadas-energy-industry-technology-trends-in-2023/.
[3] Bilge Acun, Benjamin Lee, Fiodar Kazhamiaka, Kiwan Maeng, Udit Gupta, Manoj Chakkaravarthy, David Brooks, and Carole-Jean Wu. Carbon Explorer: A Holistic Framework for Designing Carbon Aware Datacenters. In *Proceedings of ACM International Conference on Architectural Support for Programming Languages and Operating Systems*, page 118–132, 2023.
[4] Mohammad Al-Fares, Alexander Loukissas, and Amin Vahdat. A Scalable, Commodity Data Center Network Architecture. In *Proceedings of the ACM SIGCOMM Conference on Data Communication*, page 63–74, 2008.
[5] AMD. AMD PowerNow! Technology. https://www.amd.com/content/dam/amd/en/documents/archived-tech-docs/white-papers/24404a.pdf.
[6] David G Andersen, Jason Franklin, Michael Kaminsky, Amar Phanishayee, Lawrence Tan, and Vijay Vasudevan. FAWN: A Fast Array of Wimpy Nodes. In *Proceedings of ACM SIGOPS Symposium on Operating Systems Principles*, pages 1–14, 2009.

[7] Thomas Anderson, Adam Belay, Mosharaf Chowdhury, Asaf Cidon, and Irene Zhang. Treehouse: A Case For Carbon-Aware Datacenter Software. In *HotCarbon: Workshop on Sustainable Computer Systems Design and Implementation*, 2022.
[8] Esmail Asyabi, Azer Bestavros, Erfan Sharafzadeh, and Timothy Zhu. Peafowl: In-application CPU Scheduling to Reduce Power Consumption of In-memory Key-Value Stores. In *Proceedings of ACM Symposium on Cloud Computing*, pages 150–164, 2020.
[9] Brian Barrett. How Google Meet Weathered the Work-From-Home Explosion. https://www.wired.com/story/how-google-meet-weathered-work-from-home-explosion/.
[10] John Campbell. Data Centres Used 14% of Republic of Ireland's Electricity Use. https://www.bbc.com/news/world-europe-61308747.
[11] Aaron Carroll and Gernot Heiser. An Analysis of Power Consumption in a Smartphone. In *Proceedings of USENIX Annual Technical Conference*, pages 1–14, 2010.
[12] Geoffrey Challen and Mark Hempstead. The Case for Power-agile Computing. In *Proceedings of USENIX Conference on Hot Topics in Operating Systems*, pages 1–5, 2011.
[13] Inho Cho, Ahmed Saeed, Joshua Fried, Seo Jin Park, Mohammad Alizadeh, and Adam Belay. Overload Control for us-scale RPCs with Breakwater. In *Proceedings of USENIX Symposium on Operating Systems Design and Implementation*, pages 299–314, 2020.
[14] Jonathan Corbet. Live Migration of Virtual Machines over CXL. https://lwn.net/Articles/931528/.
[15] CXL™ Consortium. Compute Express Link. https://www.computeexpresslink.org/about-cxl.
[16] DARPA/ISAT Workshop on Energy-Resilient Systems, December 2020.
[17] Jacqueline Davis. Data Center Operators Will Face More Grid Disturbances. https://journal.uptimeinstitute.com/data-center-operators-will-face-more-grid-disturbances/.
[18] DMTF. Redfish Developer Hub. https://redfish.dmtf.org/.
[19] Dmitry Duplyakin, Robert Ricci, Aleksander Maricq, Gary Wong, Jonathon Duerig, Eric Eide, Leigh Stoller, Mike Hibler, David Johnson, Kirk Webb, Aditya Akella, Kuangching Wang, Glenn Ricart, Larry Landweber, Chip Elliott, Michael Zink, Emmanuel Cecchet, Snigdhaswin Kar, and Prabodh Mishra. The Design and Operation of CloudLab. In *Proceedings of USENIX Annual Technical Conference*, pages 1–14, 2019.
[20] Dominion Energy. 2020 Virginia Integrated Resource Plan. https://www.dominionenergy.com/-/media/pdfs/global/2020-va-integrated-resource-plan.pdf.
[21] Dominion Energy. 2021 Update to the 2020 Integrated Resource Plan. https://www.dominionenergy.com/-/media/pdfs/global/company/2021-de-integrated-resource-plan.pdf.
[22] Xiaobo Fan, Wolf-Dietrich Weber, and Luiz Andre Barroso. Power Provisioning for a Warehouse-Sized Computer. *ACM SIGARCH Computer Architecture News*, 35(2):13–23, 2007.
[23] Robbie Galvin. Data Centers Are Pushing Ireland's Electric Grid to the Brink. https://gizmodo.com/data-centers-are-pushing-ireland-s-electric-grid-to-the-1848282390.
[24] Greenpeace. Clicking Clean: Who is Winning the Race to Build a Green Internet? https://www.greenpeace.de/publikationen/20170110_greenpeace_clicking_clean.pdf.
[25] C. Heising. IEEE Recommended Practice for the Design of Reliable Industrial and Commercial Power Systems, 2007. IEEE 493-2007.
[26] Andrew J Herdrich, Marcel David Cornu, and Khawar Munir Abbasi. Introduction to Memory Bandwidth Allocation. https://www.intel.com/content/www/us/en/developer/articles/technical/introduction-to-memory-bandwidth-allocation.html.
[27] Hubblo. Scaphandre. https://github.com/hubblo-org/scaphandre.
[28] Jaehyun Hwang, Qizhe Cai, Ao Tang, and Rachit Agarwal. TCP ≈ RDMA: CPU-efficient Remote Storage Access with i10 . In *Proceedings of USENIX Symposium on Networked Systems Design and Implementation*, pages 127–140, 2020.
[29] Jaehyun Hwang, Midhul Vuppalapati, Simon Peter, and Rachit Agarwal. Rearchitecting Linux Storage Stack for μs Latency and High Throughput. In *Proceedings of USENIX Symposium on Operating Systems Design and Implementation*, pages 113–128, 2021.
[30] Hongyu Hè, Michal Friedman, and Theodoros Rekatsinas. EnergAt: Fine-Grained Energy Attribution for Multi-Tenancy. In *Proceedings of Workshop on Sustainable Computer Systems*, pages 1–8, 2023.
[31] Kevin Imboden. 2022 Global Data Center Market Comparison. https://cushwake.cld.bz/2022-Global-Data-Center-Market-Comparison.
[32] Intel Corporation. Intel 82599 10 GbE Controller Datasheet. Revision 2.6.
[33] Intel Corporation. Intelligent Platform Management Interface Specification Second Generation v2.0. https://www.intel.com/content/www/us/en/products/docs/servers/ipmi/ipmi-second-gen-interface-spec-v2-rev1-1.html.
[34] Intel Corporation. Overview of Enhanced Intel SpeedStep® Technology for Intel® Processors. https://www.intel.com/content/www/us/en/support/articles/000007073/processors.html.
[35] Ireland Government Statement on the Role of Data Centres in Ireland's Enterprise Strategy. https://assets.gov.ie/231142/e108d6fa-c769-4286-8fb4-0e2ff07548fe.pdf.
[36] Nicola Jones. How to Stop Data Centres From Gobbling Up the World's Electricity. *Nature*, 561(7722):163–167, 2018.
[37] Peter Judge. EirGrid Pulls Plug on 30 Irish Data Center Projects. https://www.datacenterdynamics.com/en/news/eirgrid-pulls-plug-on-30-irish-data-center-projects/.
[38] Kostis Kaffes, Dragos Sbirlea, Yiyan Lin, David Lo, and Christos Kozyrakis. Leveraging application classes to save power in highly-utilized data centers. In *Proceedings of the 11th ACM Symposium on Cloud Computing*, pages 134–149, 2020.
[39] Kubernetes Efficient Power Level Exporter (Kepler). https://github.com/sustainable-computing-io/kepler.
[40] Jongyul Kim, Insu Jang, Waleed Reda, Jaeseong Im, Marco Canini, Dejan Kostić, Youngjin Kwon, Simon Peter, and Emmett Witchel. LineFS: Efficient SmartNIC Offload of a Distributed File System with Pipeline Parallelism. In *Proceedings of the ACM SIGOPS Symposium on Operating Systems Principles*, page 756–771, 2021.
[41] Alok Gautam Kumbhare, Reza Azimi, Ioannis Manousakis, Anand Bonde, Felipe Frujeri, Nithish Mahalingam, Pulkit A. Misra, Seyyed Ahmad Javadi, Bianca Schroeder, Marcus Fontoura, and Ricardo Bianchini. Prediction-Based Power Oversubscription in Cloud Platforms. In *Proceedings of USENIX Annual Technical Conference*, pages 473–487, 2021.
[42] Etienne Le Sueur and Gernot Heiser. Dynamic Voltage and Frequency Scaling: The Laws of Diminishing Returns. In *Proceedings of International Conference on Power Aware Computing and Systems*, page 1–8, 2010.
[43] Shaohong Li, Xi Wang, Xiao Zhang, Vasileios Kontorinis, Sreekumar Kodakara, David Lo, and Parthasarathy Ranganathan. Thunderbolt: Throughput-Optimized, Quality-of-Service-Aware Power Capping at Scale. In *Proceedings of USENIX Symposium on Operating Systems Design and Implementation*, pages 1241–1255, 2020.
[44] Liuzixuan Lin and Andrew A Chien. Adapting Datacenter Capacity for Greener Datacenters and Grid. In *Proceedings of ACM International Conference on Future Energy Systems*, page 200–213, 2023.
[45] Liuzixuan Lin, Rajini Wijayawardana, Varsha Rao, Hai Nguyen, Wedan Emmanuel Gnibga, and Andrew A. Chien. Can Datacenters Get the Power Needed to Meet the Explosive Demand for AI?, 2023.
[46] Liuzixuan Lin, Victor M. Zavala, and Andrew A. Chien. Evaluating Coupling Models for Cloud Datacenters and Power Grids. In *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, page 171–184, 2021.
[47] Moritz Lipp, Andreas Kogler, David Oswald, Michael Schwarz, Catherine Easdon, Claudio Canella, and Daniel Gruss. Platypus: Software-based power side-channel attacks on x86. In *2021 IEEE Symposium on Security and Privacy (SP)*, 2021.
[48] Ming Liu, Tianyi Cui, Henry Schuh, Arvind Krishnamurthy, Simon Peter, and Karan Gupta. Offloading Distributed Applications onto SmartNICs using iPipe. In *Proceedings of ACM Special Interest Group on Data Communication*, pages 318–333, 2019.
[49] Ming Liu, Simon Peter, Arvind Krishnamurthy, and Phitchaya Mangpo Phothilimthana. E3: Energy-Efficient Microservices on SmartNIC-Accelerated Servers. In *Proceedings of USENIX Annual Technical Conference*, pages 363–378, 2019.
[50] David Lo, Liqun Cheng, Rama Govindaraju, Luiz André Barroso, and Christos Kozyrakis. Towards Energy Proportionality for Large-scale Latency-critical Workloads. In *Proceedings of ACM/IEEE International Symposium on Computer Architecture*, pages 301–312, 2014.
[51] Eric Masanet, Arman Shehabi, Nuoa Lei, Sarah Smith, and Jonathan Koomey. Recalibrating Global Data Center Energy-Use Estimates. *Science*, 367(6481):984–986, 2020.
[52] David Meisner, Brian T Gold, and Thomas F Wenisch. PowerNap: eliminating server idle power. *ACM SIGARCH Computer Architecture News*, 37(1):205–216, 2009.
[53] Bruce Myatt and Russell Carr. Advanced Microgrids as a Resiliency Strategy for Federal Data Centers. https://datacenters.lbl.gov/sites/default/files/Designing%20and%20Managing%20Data%20Centers%20for%20Resilience%20-%20Demand%20Response%20and%20Microgrids_3Dec2019_0.pdf.
[54] Khang T Nguyen. Introduction to Cache Allocation Technology in the Intel® Xeon® Processor E5 v4 Family. https://www.intel.com/content/www/us/en/developer/articles/technical/introduction-to-cache-allocation-technology.html.
[55] NVM Express Workgroup. NVM Express 1.2.1. http://www.nvmexpress.org/wp-content/uploads/NVM_Express_1_2_1_Gold_20160603.pdf.
[56] NVM Express Workgroup. NVM ExpressTM over Fabrics Revision 1.1. https://nvmexpress.org/wp-content/uploads/NVMe-over-Fabrics-1.1-2019.10.22-Ratified.pdf.
[57] Srinivas Pandruvada. Running Average Power Limit. https://01.org/blogs/2014/running-average-power-limit-%E2%80%93-rapl.
[58] Pratyush Patel, Zibo Gong, Syeda Rizvi, Esha Choukse, Pulkit Misra, Thomas Anderson, and Akshitha Sriraman. Towards Improved Power Management in Cloud GPUs. *IEEE Computer Architecture Letters*, pages 1–4, 2023.

[59] Fred Pearce. Energy Hogs: Can World's Huge Data Centers Be Made More Efficient? *Yale Environment 360*, 2018.

[60] Perf Wiki. perf: Linux Profiling With Performance Counters. https://perf.wiki.kernel.org.

[61] George Prekas, Mia Primorac, Adam Belay, Christos Kozyrakis, and Edouard Bugnion. Energy Proportionality and Workload Consolidation for Latency-critical Applications. In *Proceedings of ACM Symposium on Cloud Computing*, pages 342–355, 2015.

[62] Varun Sakalkar, Vasileios Kontorinis, David Landhuis, Shaohong Li, Darren De Ronde, Thomas Blooming, Anand Ramesh, James Kennedy, Christopher Malone, Jimmy Clidaras, and Parthasarathy Ranganathan. Data Center Power Oversubscription with a Medium Voltage Power Plane and Priority-Aware Capping. In *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*, page 497–511, 2020.

[63] Yizhou Shan, Yutong Huang, Yilun Chen, and Yiying Zhang. LegoOS: A Disseminated, Distributed OS for Hardware Resource Disaggregation. In *Proceedings of USENIX Conference on Operating Systems Design and Implementation*, page 69–87, 2018.

[64] Navin Sharma, Sean Barker, David Irwin, and Prashant Shenoy. Blink: Managing Server Clusters on Intermittent Power. In *Proceedings of International Conference on Architectural Support for Programming Languages and Operating Systems*, page 185–198, 2011.

[65] Hans-Werner Sinn. Buffering Volatility: A Study on the Limits of Germany's Energy Revolution. *European Economic Review*, 99:130–150, 2017.

[66] Abel Souza, Noman Bashir, Jorge Murillo, Walid Hanafy, Qianlin Liang, David Irwin, and Prashant Shenoy. Ecovisor: A Virtual Energy System for Carbon-Efficient Applications. In *Proceedings of ACM International Conference on Architectural Support for Programming Languages and Operating Systems*, page 252–265, 2023.

[67] Spirals Research Group. PowerAPI. https://github.com/powerapi-ng.

[68] Dragan Stancevic. nil-migration: Nearly Instantaneous Live Migration of Virtual Machines, Containers, and Processes. https://nil-migration.org/.

[69] Sumanth Umesh and Sparsh Mittal. A Survey of Techniques for Intermittent Computing. *Journal of Systems Architecture*, 112:101859–101859, 2021.

[70] Qiang Wu, Qingyuan Deng, Lakshmi Ganesh, Chang-Hong Hsu, Yun Jin, Sanjeev Kumar, Bin Li, Justin Meza, and Yee Jiun Song. Dynamo: Facebook's Data Center-Wide Power Management System. In *Proceedings of the 43rd International Symposium on Computer Architecture*, page 469–480, 2016.

[71] Chaojie Zhang, Alok Gautam Kumbhare, Ioannis Manousakis, Deli Zhang, Pulkit A. Misra, Rod Assis, Kyle Woolcock, Nithish Mahalingam, Brijesh Warrier, David Gauthier, Lalu Kunnath, Steve Solomon, Osvaldo Morales, Marcus Fontoura, and Ricardo Bianchini. Flex: High-Availability Datacenters With Zero Reserved Power. In *Proceedings of ACM/IEEE Annual International Symposium on Computer Architecture*, pages 319–332, 2021.

[72] Zhenkai Zhang, Sisheng Liang, Fan Yao, and Xing Gao. Red alert for power leakage: Exploiting intel rapl-induced side channels. In *Proceedings of the 2021 ACM Asia Conference on Computer and Communications Security*, page 162–175, 2021.