Can Storage Devices be Power Adaptive?

Dedong Xie
University of Washington
dedongx@cs.washington.edu

Theano Stavrinos
University of Washington
thst@uw.edu

Kan Zhu
University of Washington
kanzhu@cs.washington.edu

Simon Peter
University of Washington
simpeter@cs.washington.edu

Baris Kasikci
University of Washington
baris@cs.washington.edu

Thomas Anderson
University of Washington
tom@cs.washington.edu

Abstract

Power is becoming a scarce resource for data centers, raising the need for *power adaptive* system design—the ability to dynamically change power consumption—to match available power. Storage makes up an increasing fraction of total data center power consumption. As such, it holds great potential to contribute to data center power adaptivity.

To this end, we conduct a measurement study of power control mechanisms on a variety of modern data center storage devices. By changing device power states and shaping IO, we achieve a power dynamic range of up to 59.4% of the device's maximum operating power. We also study power control trade-offs, including throughput and latency. Based on our observations, we construct storage device power-throughput models and discuss the implications on power adaptive storage system design.

CCS Concepts: **• Hardware → Enterprise level and data centers power issues**; **• Information systems → Storage power management**.

Keywords: Data center, storage system, power management

ACM Reference Format:
Dedong Xie, Theano Stavrinos, Kan Zhu, Simon Peter, Baris Kasikci, and Thomas Anderson. 2024. Can Storage Devices be Power Adaptive?. In *16th ACM Workshop on Hot Topics in Storage and File Systems (HOTSTORAGE '24), July 8–9, 2024, Santa Clara, CA, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3655038.3665945

1 Introduction

Power management has become a focus of modern data center operations. On short timescales, data centers oversubscribe power to support higher load within existing power infrastructure [18, 39]; adaptation here needs to occur in milliseconds. In the medium term, power availability can vary due to internal factors like power rail failures [9] and external factors like the share of renewable energy in the power supply and the impact of weather on grid demand and supply [1, 2]. On longer timescales, the increased computational and storage demands of data centers are outpacing the ability of grids to supply highly reliable power to meet those demands [14]. For all three reasons, data center operators increasingly must actively manage power and contribute to demand response programs [20, 40].

We believe that now is the time to build power-adaptive storage systems to help address these issues. A power-adaptive storage system can adjust power consumption on varying timescales to match available supply[1]. The fraction of power consumed by storage in the data center is increasing. When hard disk drives (HDDs) were prevalent, storage made up around 10% of the total data center power draw [29]. Flash-based solid-state drives (SSDs) now make up a large and increasing share of storage in data centers to meet the need for low-latency, high-bandwidth, and high-access rate storage [22, 43, 44]. While SSDs consume similar or less power at idle than HDDs, their peak active power can be 2× the peak active power of hard disks for the same storage capacity [44]. Furthermore, the active power of data center SSDs has more than septupled in the last decade: early SSD generations consume about 3.4W [24] and current devices consume about 25W [27]. As such, understanding the power characteristics and control mechanisms of modern storage devices is becoming important as a step towards controlling power usage in the data center.

To that end, we carry out a measurement study of modern data center storage devices. We explore the potential of shaping IO (e.g., sequential and random IO, IO chunk sizes) and using device power control mechanisms to tune storage device power usage as a function of available power. We build a measurement infrastructure to study the impact of IO size and queue depth on power usage for both random and sequential IO. With these measurements, we construct a power consumption model of each device. Our models quantify the power/performance trade-off, allowing storage system designers to make informed choices about how to respond to power reduction events while respecting the performance guarantees provided to users.

[1]Power adaptivity is related to but different from power proportionality, the design of storage systems whose average power use scales up and down with workload intensity [3, 34, 36, 37].

This work is licensed under a Creative Commons Attribution International 4.0 License.

HOTSTORAGE '24, July 8–9, 2024, Santa Clara, CA, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0630-1/24/07.
https://doi.org/10.1145/3655038.3665945

Specifically, we find that device power control and shaping IO can halve a device’s idle power and enable a power dynamic range of up to 59.4% of the device’s maximum active power. However, applying these mechanisms blindly can cause device throughput to drop precipitously, to 1/25 of maximum. We thoroughly study these trade-offs and analyze the implications for the storage system, such as IO traffic shaping, power-aware IO direction to a subset of active devices, and leveraging asymmetric IO.

We summarize our contributions as follows:

- We conduct a systematic study of power consumption of three enterprise-grade SSDs and one enterprise-grade HDD handling various microbenchmarks.
- We discuss the implications of our observations on the design of power-adaptive storage systems.
- We present the design of a power measurement system that can work with enterprise-grade SSDs and HDDs to collect device power draw with a sub-10 ms period and an accuracy within 1% relative error.

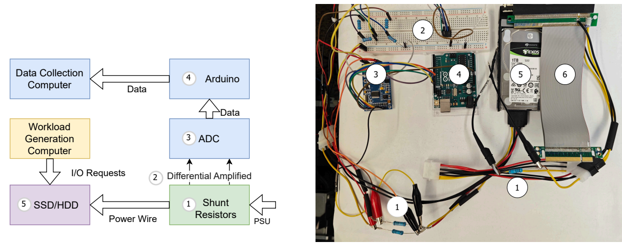
2 Power Characteristics of Storage Devices

The dynamic range of storage device power can be considerable. For example, in a storage server with 16 SSDs, each SSD can have an idle power of 5W and an active power of 23W (e.g., the Samsung PM1743 [27]). The total idle storage device power is 80W and the active power can be up to 368W. This range is comparable with the power dynamic range of the host server without storage devices.

Modern storage devices have built-in power control mechanisms [11, 25]. These mechanisms include low-power idle modes and, for SSDs, caps on the operating power of the device. Storage device power can also be modulated through storage IO operations issued by the host. These mechanisms are important tools for building power-adaptive storage systems, but they come with performance trade-offs that must be considered. We explore their effect for HDDs and SSDs—the most prevalent data center storage devices—in this work. In this section we give a brief overview of how power is consumed in these storage devices.

Hard disk drives. To achieve low IO latency, HDD platters must rotate constantly; most disks support a single, constant rotation rate. As a consequence, HDDs have a narrow power dynamic range during normal operation. HDDs support low-power states that typically flush any buffered data in the device’s on-board DRAM and spin the disk down, i.e., halt platter rotation [16]. However, IO issued to a spun-down disk experiences orders of magnitude higher latency than a normal IO while the disk spins back up.

Solid-state drives. Because SSDs do little work while in an idle state, they tend to have a lower idle power than HDDs. The high capacity and high throughput served by SSDs means that their absolute power consumption is typically higher than HDDs. Further, techniques to increase SSD density such as stacking and richer encodings tends to increase maximum SSD power demand.



(a) Design. **(b)** Experimental setup.

Figure 1. Power measurement infrastructure.

Like HDDs, SSDs can be put into a low-power standby mode, which uses one-tenth of the power of the device at idle [25], or powered down entirely. Because SSDs have no mechanical parts, transitioning into and out of low-power or powered-down modes takes milliseconds, and hence has a much smaller impact on latency than it does for HDDs [12]. An SSD may also have different power states, each of which caps the device’s average power draw at a specified limit within any 10-second period. For instance, the Samsung PM1743 SSD has a typical read power of 23W and a typical write power of 21.1W [27]. It can be power-capped to consume a maximum of 9W, around 40% of its uncapped maximum power draw and 1.8× its idle power draw of 5W. The host selects a power state through the NVMe power control interface [11]. Power capping may reduce throughput and increase latency depending on the workload.

3 Power Measurement Study

The goal of our study is to determine to what extent storage devices can be power adaptive. To answer this question, we first determine the power dynamic range of a variety of representative storage devices, including SSDs and HDDs. We then explore different ways to control power consumption within this range and determine the effectiveness and trade-offs of each method.

Measurement infrastructure. In order to measure the power draw of a storage device, we must separate its power usage from the computer’s other components, e.g., the CPU and DRAM. Existing power reporting mechanisms like IPMI do not break out storage device power. Thus, we design our own measurement infrastructure as shown in Figure 1.

To support a wide range of drives (5), we build our toolkit for three prevalent interface types: SATA, PCIe, and NVMe (M.2, U.2, E1.S, etc.). SATA provides easy physical access to the power wires, which we instrument directly. For PCIe devices, we use a PCIe riser card (6) with external power supply wires that fully power the device, and instrument the external power supply wires. Because all NVMe interfaces

Label	Protocol	Model	Measured Power Range
SSD1	NVMe	Samsung PM9A3	3.5-13.5W
SSD2	NVMe	Intel D7-P5510	5-15.1W
SSD3	SATA	Intel D3-P4510	1-3.5W
HDD	SATA	Seagate Exos 7E2000	1-5.3W

Table 1. Evaluated storage devices.

are pin-compatible with PCIe, we use a PCIe transformer card to transform NVMe into PCIe (i.e., U.2 to PCIe card).

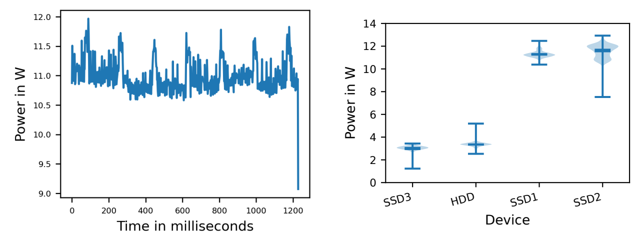
We measure power by instrumenting the power wires with shunt resistors (①). The shunt resistor R_{shunt} transforms the current signal I to a differential voltage signal, following the voltage drop on the resistor being $\Delta V = I \cdot R_{shunt}$. The power is then calculated by $P = U \cdot I$ where U is the voltage of the power wire. We use a 0.1Ω resistor along with a differential signal amplifier (②) to mitigate noise in the measurements.

The voltage signal can be measured by oscilloscope or analog-digital converters (ADCs). We use the ADC (③) to connect to our data logger to sample power data at millisecond scale. The ADC is set to sample the voltage signals at 1 kHz. An Arduino UNO (④) is used to configure the ADC and read the voltage values measured by the ADC, then send them to a data logging computer. With a 24-bit ADC, we achieve less than 1% of measurement error, sampling at millisecond-scale.

We use a Texas Instruments ADS1256 ADC and an Arduino UNO R3 for data collection and transmission. The workload generation computer is a Dell Precision Tower 5810 with Intel Xeon E5-1603 v3 CPU and 16GB RAM. This computer supports PCIe 3, which has limited bandwidth compared to some of our SSDs. We find that read bandwidth cannot always be saturated. However, this should not affect the outcome of our study.

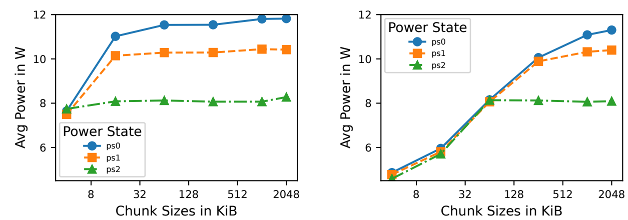
Storage devices. The storage devices used in the experiments are listed in Table 1, including the power range measured by our experiments. We cover recent storage devices that are common in data centers, including devices with SATA and NVMe interfaces, as well as HDDs and SSDs. The devices are all marketed for the data center and can be found within typical server configurations (e.g., by Supermicro [32]).

Workloads. To measure the power drawn by a storage device under different operating conditions, we test random and sequential reads and writes, variable IO chunk sizes, and variable IO queue depths. We generate workloads with `fio` 3.28 [4]. The IOs are submitted asynchronously and directly to the device, bypassing the OS page cache. By testing random-only and sequential-only workloads, we evaluate workload extremes to get the boundaries of power dynamic range. We test 6 different chunk sizes from 4 KiB to 2 MiB. We experiment with 6 different IO depths from 1 up to 128. Each experiment issues requests for one minute or until the requests total 4 GiB, whichever comes first.



(a) SSD1 power usage. **(b)** Power distribution.

Figure 2. Random write power use during one experiment (chunk size 256 KiB, queue depth 64).



(a) Queue depth 64. **(b)** Queue depth 1.

Figure 3. SSD2 random write average power under different power states.

3.1 Power Measurement Example

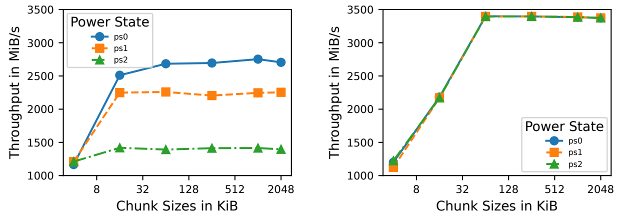
Figure 2a shows an example of the power measurements collected over a single experiment with our power measurement system. It shows that there is substantial variability in power usage over small timescales. Figure 2b shows a violin plot of the distribution of power measurements for various storage devices during the same experiment. Some devices have more power variability than others. Horizontal lines within the distribution show both median and mean, which nearly overlap. Without our measurement system's millisecond-scale sampling rate, these details in device power consumption could not be captured and analyzed.

3.2 Power Measurement Results

Two types of mechanisms influence storage device power consumption: (1) in-device mechanisms (e.g., standby modes and device power states that cap power), and (2) IO shaping (i.e., IO chunk size and queue depth). We systematically study to what extent these mechanisms can control storage device power and their trade-offs.

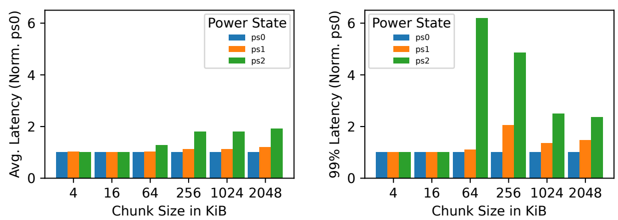
3.2.1 Power Capping. NVMe SSDs support power capping, where an active device limits average power consumption to a specification.

Benefit. Device power use cannot exceed the cap. For instance, SSD2 implements three power caps: ps0 limits maximum power to below 25 W (the maximum device power), ps1 to 12 W, and ps2 to 10 W. Figure 3 shows average



(a) Sequential writes. **(b)** Sequential reads.

Figure 4. SSD2 throughput under different power states (queue depth 64).



(a) Average. **(b)** 99th percentile.

Figure 5. SSD2 random write latency (queue depth 1).

power when these caps are in effect. As we saw in Figure 2, instantaneous power can differ from average power.

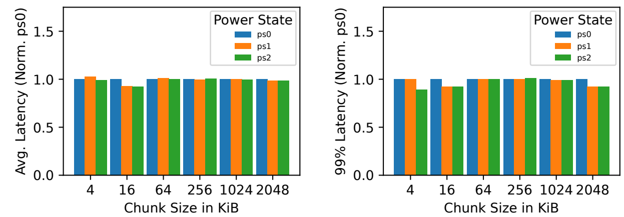
Trade-off. While power capping is efficient in controlling power usage, it hurts performance. Figure 4 shows throughput in ps1 and ps2. Write workloads suffer a significant drop compared to that of ps0. For instance, the performance of sequential writes in ps1 and ps2 is merely 74% and 55% of that of ps0, respectively.

The influence of power capping on read and write throughput differs. We compare sequential write (Figure 4a) and read (Figure 4b) throughput between power states. For sequential writes, there is a 26% (45%) drop in throughput from ps0 to ps1 (ps1 to ps2). For sequential reads, the trend disappeared: capping from ps0 to ps1 (ps1 to ps2) results in minimal drop in throughput. A similar trend in the impact of the power cap on IO throughput is also seen for SSD1. This difference makes a power cap more effective for read-intensive workloads.

As shown in Figure 5, average random write latency changes with the power cap by up to 2×. Tail latency increases dramatically. For ps2, tail latency increases by a factor of up to 6.19× relative to ps0.

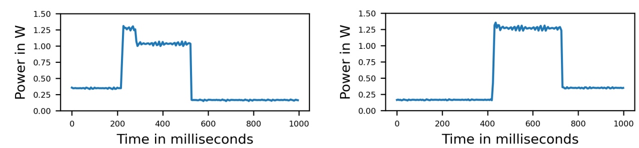
Non-trade-off. In terms of read workload latency, there is no noticeable difference in average latency and 99% latency between different power states as shown in Figure 6. This is because reads at queue depth 1 do not create enough load on the device to be power capped.

3.2.2 Low-Power Standby. To reduce power further, HDDs can spin down and SSDs can disable interface



(a) Average. **(b)** 99th percentile.

Figure 6. SSD2 random read latency (queue depth 1).



(a) Idle to standby. **(b)** Standby to idle.

Figure 7. 860 EVO SSD power use during standby transition. ALPM command issued at 200ms and 400ms.

functionalities [6], called standby. While ubiquitous in HDDs, standby is rarely supported in data center SSDs. To evaluate the potential for SSD standby mode, we experiment with a desktop SSD, the Samsung 860 EVO [26].

Benefit. Standby can double a device's power dynamic range. We measure our HDD's standby power consumption at 1.1 W, compared to 3.76 W at idle, saving 2.66 W. This is comparable with the savings between idle and active of 5.3 W.

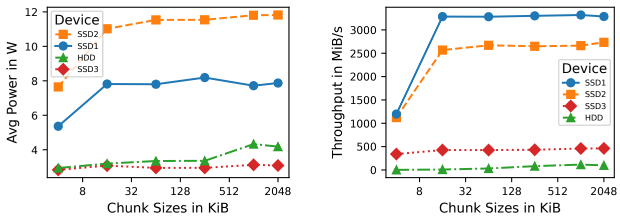
Similarly, SSD idle power is cut in half with standby (Figure 7). Aggressive Link Power Management (ALPM) [8, 17], a power management protocol for SATA devices, allows us to set standby mode. When we activate ALPM's lowest-power mode (SLUMBER), the Samsung 860 EVO SSD's power consumption decreases to 0.17 W from an idle power of 0.35 W.

Trade-off. Entering and exiting standby takes time. We observe that HDD spin-down and spin-up takes up to 10 seconds. SSDs can transition quickly—the EVO transitions within 0.5 seconds—but can consume additional power during the transition (Figure 7). In practice, requests issued while a device is in standby may incur additional latency, though for SSDs the added latency is often negligible [33, 34].

3.2.3 IO Shaping. IO size and queue depth provide the most fine-grained way to control active power, as shown in Figures 8 and 9.

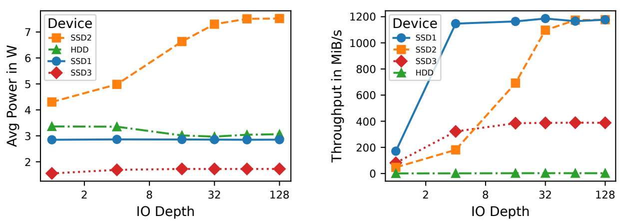
Benefit. Compared to chunk sizes of 2 MiB, 4 KiB chunks consume up to 30% less power. Compared to large IO depths of 64, an IO depth of 1 consumes up to 40% less power.

Trade-off. On the flip side, 4 KiB chunks have up to 50% performance loss, while an IO depth of 1 may provide only



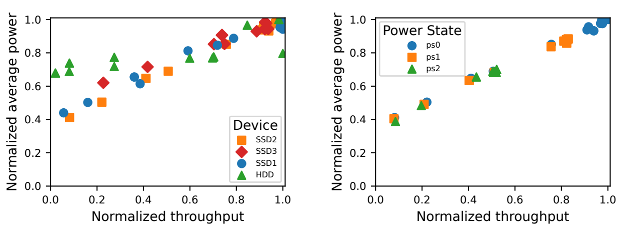
(a) Average power. **(b)** Throughput.

Figure 8. Random write power and throughput as chunk size varies (queue depth 64).



(a) Average power. **(b)** Throughput.

Figure 9. Random read power and throughput as queue depth varies (chunk size 4 KiB).



(a) Across storage devices. **(b)** SSD2 power states.

Figure 10. Power-throughput model for random write.

10% of the performance. Therefore, when considering IO shaping to control storage device's power consumption, the trade-off between power and throughput must be carefully considered to minimize impact on storage system QoS.

3.3 Power-Throughput Model

Putting together the experimental results, we can form a power-throughput model of each device under the explored power control mechanisms. To show one example of the storage power and throughput dynamic range, Figure 10 plots the normalized power and throughput with respect to each device's maximum average power and maximum throughput under a random write workload across all combinations of power control mechanisms (device power control and IO shaping). Each point in the plots represents one combination of a storage device or the device's power state, and the IO shaping on that device. The maximum power dynamic range is obtained by SSD2 with a range of 59.4% of its maximum power. The largest performance trade-off is observed in the HDD where throughput can drop to 4% of the maximum. Using the power-throughput relationship captured in these plots, a power-adaptive storage system that operates under certain power and performance constraints can find the configuration of power cap and IO shape for each device to meet these constraints.

For example, for SSD1, assume the device is operating at queue depth 64 and chunk size 256 KiB, which corresponds to the point in Figure 10a with 3.3 GiB/s write bandwidth (99.5% of maximum throughput) and 100% maximum average power at 8.19 W. For a power reduction of 20%, the model suggests a reduction of 40% in throughput, with a queue depth of 1 and chunk size of 256 KiB. The storage system can use this information to decide that it can curtail 40% × 3.3 GiB/s = 1.3 GiB/s in best-effort load to provide the same service for high-priority load under the 20% power reduction cap. In this case, it may only enter this configuration if it has 1.3 GiB/s of best-effort load available. As seen in the figure, the model generalizes across storage devices. In scenarios with multiple, heterogeneous devices, power-throughput models of multiple devices can be combined to derive the performance Pareto frontier of device configurations under a power budget.

4 Impact on Power-Adaptive Storage Systems

This study investigates mechanisms for controlling power usage of storage devices and the extent to which those mechanisms improve the power dynamic range of storage devices. Our experiments confirm that storage power consumption can be controlled with device power control and IO shaping. Because cloud operators control hardware and the lower levels of storage stacks, operators can use similar power models, as derived through our experiments, as a foundation for power-adaptive storage systems, using SLOs and power budgets as inputs. Here we discuss the trade-offs inherent in such control and the implications for power adaptive storage system design.

Power-aware IO redirection. As we observe in §3.2.2, HDD standby and spin-up take seconds, which adds to the latency of data access, posing a risk of violating latency guarantees to applications when putting HDDs in standby mode to save power. SSDs have a much shorter wake-up time, as low as sub-millisecond. If workloads can be classified and IO requests directed to active devices in a power-aware manner, the standby period of the inactive storage devices can be maximized without QoS impact (cf. SRCMap [36]). Similarly, in tiered storage, the longer standby/spin-up latencies of HDDs may be masked by temporarily absorbing writes with SSDs. Power-aware caching and prefetching may mask read latencies for data stored on standby devices (cf. EXCES [35]). Operators can leverage power when it is abundant to reorganize data into appropriate tiers.

Power-capping and IO shaping. Our measurements, summarized in §3.3, show that power capping and IO shaping can be used to reduce storage power but with a relatively large impact on throughput. To walk this trade-off intelligently, storage system developers can use our power-throughput models to determine appropriate power caps and IO shaping on storage devices under performance constraints, to minimize throughput impact when reducing power usage. In tiered storage, weaker SLOs for slower tiers may allow operators to apply power-adaptive mechanisms more aggressively on those tiers. Operators can match the power-performance models with performance guarantees to determine when and how to apply these mechanisms to different tiers. For latency, a similar model can be drawn from the measurement results.

Leveraging asymmetric IO. Given the different performance trends in read versus write workloads when the device is power capped, segregating write traffic to a small set of disks, while power capping the remainder, is a possibility. Leveraging this form of asymmetric IO can reduce power consumption while minimizing the influence on storage system QoS. In tiered storage, directing writes to lower storage tiers when power is constrained can also be applied to reduce storage power consumption.

4.1 Broader implications.

Transitioning to power-adaptive storage. There are serious consequences to incorrectly controlling power, such as bringing down power infrastructure in the data center or violating agreements with the grid. As such, it is necessary to carefully roll out any power-adaptive storage system. A power-adaptive storage system could be designed for incremental deployment at the sub-rack granularity, i.e., below the lowest tier of the data center power hierarchy [19]. Local failures of the storage system to control power can safely be identified before a failure threatens to exceed the power budget of rack-level breakers. For the same reason, small-scale test deployments should be distributed among power domains so that coordinated failures of deployments to reduce power do not overwhelm a single domain. As confidence in small-scale deployments is achieved, the size of the power-adaptive storage system can be gradually increased.

Implications on broader data center power management. Interaction with other power control mechanisms for other system components should be assessed. For instance, if the power consumption of other components is reduced, how does that affect the power consumption of storage? Will it change the preferred mechanism for reducing storage power? For example, CPU throttling to reduce CPU power usage may in turn reduce request rates to storage. In this case, IO redirection together with putting devices on standby may be preferred over IO shaping, because lower IO request rates may mean devices can remain in standby mode for longer. Further, the order in which power control techniques are applied to different components (CPU, storage devices, networking devices, etc.) can impact which techniques are most effective, opening up an area of further study.

5 Related Work

There is a large body of work from the last decade measuring the power and energy characteristics of storage devices [5, 28, 30]. Prior work investigates power and energy characteristics of HDDs, including the effect of spin-down on power use [15, 16]. Measurement studies investigate the impact of SSDs' internal architecture on power and energy [7, 41, 42]. Grupp et al. characterized the power of flash operations on SLC and MLC devices [10], but did not look at workload-level impacts on device power usage. Our study draws inspiration from this prior work, including the design of our measurement system [5, 42]. We add to these findings by investigating modern storage devices, including the high-capacity NVMe devices used in data centers today, and by focusing on the mechanisms for adapting device power usage.

Other prior work models SSD power consumption, using SSD power measurements to parameterize or to validate models [7, 21, 23]. Such work typically does not report power measurements in detail. As observed by others [5], simulations often do not accurately model device behavior; hence we carry out a measurement study in this work.

More recent work measures whole-system power while handling storage workloads, investigating the impact on power of device type [12]; IO interfaces, submission and completion mechanisms [13, 31]; and IO schedulers [38]. Investigations into system power are complementary to our work. Understanding device power is necessary for large-scale data center storage systems where a significant percentage of power is drawn by storage devices.

6 Conclusion

Through a thorough measurement study, we characterize the power control dynamic range of modern data center storage devices. We find that device power states and IO shaping can halve idle power and achieve a power control dynamic range of up to 59.4% of a device's maximum operating power. We observe the throughput and latency trade-offs when applying these mechanisms, we build a power-throughput model across storage devices, and we discuss implications on power adaptive storage system design.

Acknowledgments

We are grateful to Jonggyu Park and the anonymous reviewers, whose feedback substantially improved this work. This work was supported in part by NSF grants NGSDI-2104548 and 2148209, and the UW Center for the Future of Cloud Infrastructure.

References

[1] Bilge Acun, Benjamin Lee, Fiodar Kazhamiaka, Kiwan Maeng, Udit Gupta, Manoj Chakkaravarthy, David Brooks, and Carole-Jean Wu. 2023. Carbon Explorer: A Holistic Framework for Designing Carbon Aware Datacenters. In *Proceedings of the 28th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS'23)*. Association for Computing Machinery (ACM), New York, NY, USA.

[2] U.S. Energy Information Administration. 2024. U.S. Energy Information Administration (EIA). https://www.eia.gov/opendata/index.php

[3] Hrishikesh Amur, James Cipar, Varun Gupta, Gregory R Ganger, Michael A Kozuch, and Karsten Schwan. 2010. Robust and flexible power-proportional storage. In *Proceedings of the 1st ACM Symposium on Cloud Computing 2010 (SoCC'10)*. Association for Computing Machinery (ACM), New York, NY, USA.

[4] Jens Axboe. 2017. fio - Flexible I/O tester rev. 3.36. https://fio.readthedocs.io/en/latest/fio_doc.html

[5] Matias Bjørling, Philippe Bonnet, Luc Bouganim, and Björn Þór Jónsson. 2010. uFLIP: Understanding the energy consumption of flash devices. *Bulletin of the Technical Committee on Data Engineering* 33, 4 (2010), 48–54.

[6] Hung-Ching Chang, Erik Kruus, Thomas J Barnes, Abhishek R Agrawal, and Kirk W Cameron. 2012. Storage Power Optimizations for Client Devices and Data Centers. *Intel Technology Journal* 16, 3 (2012).

[7] Seokhei Cho, Changhyun Park, Youjip Won, Sooyong Kang, Jaehyuk Cha, Sungroh Yoon, and Jongmoo Choi. 2015. Design tradeoffs of SSDs: From energy consumption's perspective. *ACM Transactions on Storage (TOS)* 11, 2 (2015), 1–24.

[8] Intel Corporation. 2011. Designing Energy Efficient SATA Devices: Overview and Implementation Recommendations. https://www.intel.com/content/dam/doc/reference-guide/sata-devices-implementation-recommendations.pdf

[9] Jeff Dean and Amin Vahdat. 2023. Exciting Directions for ML Models and the Implications for Computing Hardware. https://www.youtube.com/watch?v=EFe7-WZMMhc

[10] Laura M Grupp, Adrian M Caulfield, Joel Coburn, Steven Swanson, Eitan Yaakobi, Paul H Siegel, and Jack K Wolf. 2009. Characterizing flash memory: Anomalies, observations, and applications. In *Proceedings of the 42nd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO'09)*. Association for Computing Machinery (ACM), New York, NY, USA.

[11] Jonmichael Hands, Dennis Worley, and Lakhveer Kaur. 2022. Technology power features - NVM express. https://nvmexpress.org/resource/technology-power-features/

[12] Bryan Harris and Nihat Altiparmak. 2020. Ultra-Low Latency SSDs' Impact on Overall Energy Efficiency. In *Proceedings of the 12th USENIX Workshop on Hot Topics in Storage and File Systems (HotStorage'20)*. USENIX Association, Berkeley, CA, USA.

[13] Bryan Harris and Nihat Altiparmak. 2022. When poll is more energy efficient than interrupt. In *Proceedings of the 14th ACM Workshop on Hot Topics in Storage and File Systems (HotStorage '22)*. Association for Computing Machinery (ACM), New York, NY, USA.

[14] Tianhang Huang, Tianyi Zhai, Xin Zhang, and Xiaomeng Di. 2021. Electric Power-grid Friendly Characteristic Data Center Energy Consumption Optimization Method. In *Journal of Physics: Conference Series*, Vol. 2095. IOP Publishing, 012026.

[15] Anthony Hylick, Andrew Rice, Brian Jones, and Ripduman Sohan. 2007. Hard drive power consumption uncovered. *SIGMETRICS Perform. Evaluation Rev.* 35, 3 (2007), 54–55.

[16] Anthony Hylick, Ripduman Sohan, Andrew Rice, and Brian Jones. 2008. An Analysis of Hard Drive Energy Consumption. In *2008 IEEE International Symposium on Modeling, Analysis and Simulation of Computers and Telecommunication Systems*. IEEE Computer Society, Washington, DC, USA.

[17] Intel. 2024. AHCI specification. https://www.intel.com/content/www/us/en/io/serial-ata/ahci.html

[18] Shaohong Li, Xi Wang, Faria Kalim, Xiao Zhang, Sangeetha Abdu Jyothi, Karan Grover, Vasileios Kontorinis, Nina Narodytska, Owolabi Legunsen, Sreekumar Kodakara, et al. 2020. Thunderbolt: Throughput-Optimized, Quality-of-Service-Aware Power Capping at Scale. In *14th USENIX Symposium on Operating Systems Design and Implementation (OSDI'20)*. USENIX Association, Berkeley, CA, USA.

[19] Sulav Malla, Qingyuan Deng, Zoh Ebrahimzadeh, Joe Gasperetti, Sajal Jain, Parimala Kondety, Thiara Ortiz, and Debra Vieira. 2020. Coordinated priority-aware charging of distributed batteries in oversubscribed data centers. In *Proceedings of the 53rd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO'20)*. Association for Computing Machinery (ACM), New York, NY, USA.

[20] Varun Mehra and Raiden Hasegawa. 2023. Supporting power grids with demand response at Google data centers. https://cloud.google.com/blog/products/infrastructure/using-demand-response-to-reduce-data-center-power-consumption

[21] Vidyabhushan Mohan, Sudhanva Gurumurthi, and Mircea R Stan. 2010. FlashPower: A detailed power model for NAND flash memory. In *2010 Design, Automation & Test in Europe Conference & Exhibition (DATE'10)*. IEEE Computer Society, Washington, DC, USA.

[22] Satadru Pan, Theano Stavrinos, Yunqiao Zhang, Atul Sikaria, Pavel Zakharov, Abhinav Sharma, Shiva Shankar, Mike Shuey, Richard Wareing, Monika Gangapuram, Guanglei Cao, Christian Preseau, Pratap Singh, Kestutis Patiejunas, JR Tipton, Ethan Katz-Bassett, and Wyatt Lloyd. 2021. Facebook's Tectonic Filesystem: Efficiency from Exascale. In *Proceedings of the 19th USENIX Conference on File and Storage Technologies (FAST'21)*. USENIX Association, Berkeley, CA, USA.

[23] Jinha Park, Sungjoo Yoo, Sunggu Lee, and Chanik Park. 2009. Power modeling of solid state disk for dynamic power management policy design in embedded systems. In *Software Technologies for Embedded and Ubiquitous Systems: 7th IFIP WG 10.2 International Workshop*. Springer, 24–35.

[24] Samsung. 2012. Samsung SM843 Pro. https://www.compuram.biz/documents/datasheet/SM843_Datasheet.pdf

[25] Samsung. 2017. Samsung SSD 960 EVO M.2. https://download.semiconductor.samsung.com/resources/data-sheet/Samsung_SSD_960_EVO_Data_Sheet_Rev_1_2.pdf

[26] Samsung. 2017. Samsung V-NAND SSD 860 EVO. https://download.semiconductor.samsung.com/resources/data-sheet/Samsung_SSD_860_EVO_Data_Sheet_Rev1.pdf

[27] Samsung. 2024. Samsung PM1743. https://semiconductor.samsung.com/us/ssd/enterprise-ssd/pm1743/

[28] Euiseong Seo, Seon-Yeong Park, and Bhuvan Urgaonkar. 2008. Empirical Analysis on Energy Efficiency of Flash-based SSDs.. In *Proceedings of the 2008 USENIX Workshop on Power-Aware Computing and Systems (HotPower'08)*. USENIX Association, Berkeley, CA, USA.

[29] Arman Shehabi, Sarah Smith, Dale Sartor, Richard Brown, Magnus Herrlin, Jonathan Koomey, Eric Masanet, Nathaniel Horner, Inês Azevedo, and William Lintner. 2016. United states data center energy usage report. (2016).

[30] Seungyong Shin and Dongkun Shin. 2010. Power Consumption Characterization of Flash Memory SSD. In *ICEIC: International Conference on Electronics, Informations and Communications*. 14–18.

[31] Sidharth Sundar, William Simpson, Jacob Higdon, Caeden Whitaker, Bryan Harris, and Nihat Altiparmak. 2023. Energy Implications of IO Interface Design Choices. In *Proceedings of the 15th ACM Workshop on Hot Topics in Storage and File Systems (HotStorage '23)*. Association for Computing Machinery (ACM), New York, NY, USA.

[32] Supermicro. 2024. Supermicro X13 for Storage. https://www.supermicro.com/en/products/x13

[33] Billy Tallis. 2020. The Samsung 980 PRO PCIe 4.0 SSD Review: A Spirit of Hope. https://www.anandtech.com/show/16087/the-samsung-980-

pro-pcie-4-ssd-review/8

[34] Eno Thereska, Austin Donnelly, and Dushyanth Narayanan. 2011. Sierra: practical power-proportionality for data center storage. In *Proceedings of the 6th European Conference on Computer Systems (EuroSys'11)*. Association for Computing Machinery (ACM), New York, NY, USA.

[35] Luis Useche, Jorge Guerra, Medha Bhadkamkar, Mauricio Alarcon, and Raju Rangaswami. 2008. EXCES: External caching in energy saving storage systems. In *Proceedings of the 2008 IEEE 14th International Symposium on High Performance Computer Architecture (HPCA'08)*. IEEE Computer Society, Washington, DC, USA.

[36] Akshat Verma, Ricardo Koller, Luis Useche, and Raju Rangaswami. 2010. SRCMap: Energy Proportional Storage Using Dynamic Consolidation.. In *Proceedings of the 8th USENIX Conference on File and Storage Technologies (FAST'10)*. USENIX Association, Berkeley, CA, USA.

[37] Charles Weddle, Mathew Oldham, Jin Qian, An-I Andy Wang, Peter Reiher, and Geoff Kuenning. 2007. PARAID: A gear-shifting power-aware RAID. *ACM Transactions on Storage (TOS)* 3, 3 (2007), 13–es.

[38] Caeden Whitaker, Sidharth Sundar, Bryan Harris, and Nihat Altiparmak. 2023. Do we still need IO schedulers for low-latency disks?. In *Proceedings of the 15th ACM Workshop on Hot Topics in Storage and File Systems (HotStorage '23)*. Association for Computing Machinery (ACM), New York, NY, USA.

[39] Qiang Wu, Qingyuan Deng, Lakshmi Ganesh, Chang-Hong Hsu, Yun Jin, Sanjeev Kumar, Bin Li, Justin Meza, and Yee Jiun Song. 2016. Dynamo: Facebook's data center-wide power management system. *ACM SIGARCH Computer Architecture News* 44, 3 (2016), 469–480.

[40] Jiali Xing, Bilge Acun, Aditya Sundarrajan, David Brooks, Manoj Chakkaravarthy, Nikky Avila, Carole-Jean Wu, and Benjamin C Lee. 2023. Carbon Responder: Coordinating Demand Response for the Datacenter Fleet. *arXiv preprint arXiv:2311.08589* (2023).

[41] Balgeun Yoo, Youjip Won, Seokhei Cho, Sooyong Kang, Jongmoo Choi, and Sungroh Yoon. 2011. SSD Characterization: From Energy Consumption's Perspective. In *Proceedings of the 3rd USENIX Workshop on Hot Topics in Storage and File Systems (HotStorage'11)*. USENIX Association, Berkeley, CA, USA.

[42] Jie Zhang, Mustafa Shihab, and Myoungsoo Jung. 2014. Power, Energy, and Thermal Considerations in SSD-Based I/O Acceleration. In *Proceedings of the 6th USENIX Workshop on Hot Topics in Storage and File Systems (HotStorage'14)*. USENIX Association, Berkeley, CA, USA.

[43] Mark Zhao, Niket Agarwal, Aarti Basant, Buğra Gedik, Satadru Pan, Mustafa Ozdal, Rakesh Komuravelli, Jerry Pan, Tianshu Bao, Haowei Lu, et al. 2022. Understanding data storage and ingestion for large-scale deep recommendation model training: Industrial product. In *Proceedings of the 49th Annual International Symposium on Computer Architecture (ISCA'22)*. Association for Computing Machinery (ACM), New York, NY, USA.

[44] Mark Zhao, Satadru Pan, Niket Agarwal, Zhaoduo Wen, David Xu, Anand Natarajan, Pavan Kumar, Ritesh Tijoriwala, Karan Asher, Hao Wu, et al. 2023. Tectonic-Shift: A Composite Storage Fabric for Large-Scale ML Training. In *2023 USENIX Annual Technical Conference (USENIX ATC'23)*. USENIX Association, Berkeley, CA, USA.