An Analysis of Hard Drive Energy Consumption

Anthony Hylick, Ripduman Sohan, Andrew Rice, and Brian Jones
University of Cambridge, Computer Laboratory
Firstname.Lastname@cl.cam.ac.uk

October 24, 2008, This version corrects an error in the previous version of the paper regarding write energy consumption (§7.3).

Abstract

The increasing storage capacity and necessary redundancy of data centers and other large-scale IT facilities has drawn attention to the issue of reducing the power consumption of hard drives. This work comprehensively investigates the power consumption of hard drives to determine typical runtime power profiles. We have instrumented at a fine-grained level and present our findings which show that (i) the energy consumed by the electronics of a drive is just as important as the mechanical energy consumption; (ii) the energy required to access data is affected by physical location on a drive; and (iii) the size of data transfers has measurable effect on power consumption.

1. Introduction

Reducing power consumption in grid-powered computing contexts has garnered much attention over the past six years. Prior to this, power consumption was a concern primarily for mobile computing domains. The rising cost of energy and increased public awareness surrounding the environment and sustainability has prompted even more attention to reducing the power footprint of IT infrastructure as a whole.

Optimizing hard drive power consumption is an important part of this challenge. The amount of storage needed to fulfil current and growing computing needs seems set to continue to increase. Enabling this increased storage capacity is often accomplished through the addition of more hard drives, and these added drives will provide service at the expense of requiring additional power. The International Data Corporation (IDC) reports that storage capacity is increasing at a rate of almost 60% per year [1]. This increase in storage capacity is likely to increase the proportion of total IT power that is consumed by hard drives.

We focus our attention on obtaining an in-depth understanding of the power consumption trends of hard drives, instrumenting at a fine-grained level to present results that provide insight into the mechanical and electronic power consumption of hard drives at runtime. In particular, we show that *(i)* the power consumed by the electronic components of hard drives is just as significant as the consumption of the mechanical components and should not be considered as a secondary concern when approaching the issue of drive power management; *(ii)* data placement has an effect on the power consumption of data access but does not affect the power consumption of committing data to disk; and *(iii)* the size of transfers affects power consumption. Based on observed results, we make recommendations to aid hardware and software designers in optimizing data storage and access.

2. Motivation

Previous research has implicitly prioritized the mechanical power consumption of hard drives over the electronic components by attempting to reduce power consumption by spinning disks down. We have attempted to determine possibilities for power savings while hard drives are in their active and idle states (states in which drives are capable of servicing requests without substantial latency penalties). Our interest led us to thoroughly explore the runtime power consumption of hard drives; requiring a detailed understanding of both the mechanical and electronic component operation and power consumption of modern drives while in their different operating modes. The goal is to draw conclusions regarding drive power consumption in the general case for a wide range of commodity disks.

The remainder of the paper is organized as follows: In section 3 we summarize other power measurement efforts and highlight existing emphasis on spin-down policies; Section 4 provides a background of drive mechanics and operation; Sections 5 & 6 explain the measurement setup and testing methodology, respectively; Results and analyses are presented in Section 7; and Section 8 concludes.

3. Related Work

Our objective was to discover ways of reducing drive energy consumption not presented by previous research. A detailed understanding of the following areas has enabled us in our endeavor:

Detailed Measurement It has been our motivation to accurately acquire online power consumption measurements of hard drive activity for the discovery of opportunities to reduce drive power consumption outside of drive spin-down. Our measurement setup provides direct, online measurements of hard drive power consumption that few previous authors have described and employed. For example, Colarelli and Grunwald limited their measurement of seeks and state transitions of "large IDE" drives to calibrating their storage array simulator [4]. Similarly, Zedlewski et. al. took precise measurements of a laptop hard drive's power consumption for the purpose of creating their disk simulation environment, Dempsey [23]. Weddle et. al. constrained their facility to comparing seek, spin-up/down, and standby measurements to drive datasheets and those numbers used in other simulations [21].

With our custom hardware, we aim to uncover hidden power consumption trends and behaviors that are not exposed through coarser-grained measurement techniques. Our goal is to exploit these hidden trends for reduced energy consumption.

Spin-Down At the forefront of addressing the reduction in power consumption is the need to do so while maintaining the "always-on" or "always-available" state of computers and the internet. The same may be applied to reducing the power consumption of hard drives. However, observe that commodity hard drives do not have any active, lower-power operating modes like modern processors (i.e. voltage and frequency scaling).

Previous energy-saving research has been based around the standby paradigm; the notion that the only way to save power is to place drives in standby—an inactive state—for as long as possible. Since accessing a spun-down drive costs time and energy, devising the most effective and efficient spin-down strategy has been a common topic for prior work.

Wilkes provides an early adaptive prediction approach to determining when to spin the disk down [22]. His method calculates a delay value which is an adaptive wait time before spinning down the drive based upon the user's previous activity patterns. Greenawalt models hard drive power consumption and concludes that spinning down drives reduces power consumption for carefully selected timeout values [8]. Douglis et. al. show that spinning a drive down reduces energy consumption while comparing different spin-down algorithms [6]. Douglis et. al. also present some adaptive spin-down techniques that provide

more desirable and efficient spin-down transitions by reducing spin-ups [5]. Their adaptive approach monitors usage trends and adapts the policy as needed. Papathanasiou and Scott propose a strategy to group drive accesses together in order to maximize standby times and utilization once the drive is spun-up [15]. Helmbold et. al. use a machine learning approach to dynamically determine when to spin a drive down, arguing that more intelligent approaches to drive spin-down are needed to be effective in saving energy [11]. All of these approaches focus solely on placing drives in standby mode as a means to reduce power consumption.

More recent work has focused on partitioning larger storage deployments into groups, allowing subsets to be placed in standby mode under lighter workload demands.

Colarelli and Grunwald take advantage of the large amounts of idle drive time that results from the reliability and redundancy of most high-fidelity storage environments [4]. The pool of available drives is divided into one set of active drives, which are on and servicing requests, and another set of passive drives, which are spun-down into standby mode. Pinheiro and Bianchini introduce Popular Data Concentration (PDC) where frequently accessed data is placed onto a subset of available network storage drives [16]. The remaining drives may be spun-down to conserve energy. Ganesh et. al. present a method of intelligently spinning down a large portion of drives in large storage deployments while managing performance trade-offs [7]. Narayanan et. al. show that enterprise storage environments are amenable to drive spin-down and describe an approach to spin-down a large portion of drives during "write-dominated" access periods for reduced energy consumption [14]. Pinheiro et. al. take advantage of the redundancy of large storage deployments through a method called "Diverted Access" [17]. This technique allows for redundant drives to be spun-down while maintaining high availability. Weddle et. al. introduce a power-aware RAID (PARAID) method for reducing the power consumption of commodity hardware via an adaptive algorithm designed to allow for drives to be powered down [21]. These attempts produce the effect of an "always-available" storage environment while reducing the number of drives that are actively servicing requests. Still, the underlying premise of these approaches is that the reduction of power consumption as it relates to hard drives is only achieved by placing the disks in inactive states.

Other research has attempted to exploit scheduling and intelligent caching to reduce the energy consumption of storage elements. Zhu et. al. introduce several power-aware cache management techniques that aim to reduce energy consumption by allowing drives to remain in lower power modes longer [25]. Li and Wang describe power-aware scheduling and cache management algorithms for

different RAID configurations [13]. These authors leveraged power-aware caches and scheduling knowledge to extend drive spin-down periods and most efficiently utilize a spun-up drive.

Son et. al. explore physical layout optimizations across drive arrays for increased performance and increased idle periods in large systems [20]. The increased idle periods resulting from the authors' optimizations allow for drives to be placed in standby mode for longer periods of time. For applications whose data access patterns are known and understood in detail, this work presents a viable option for reduced energy consumption. Again, the energy-savings proclaimed by this work are only realized after drives are placed in standby mode.

Throttling Zhu et. al. present Hibernator, an energy management system for drive arrays in large storage environments [24]. This system adaptively changes the speeds at which drives are spinning in order to save energy while maintaining performance guarantees. Gurumurthi et. al. explore a method for dynamically controlling the speed at which drives spin with DRPM [9]. Focused on large-scale storage deployments like data centers, this approach reduces the total energy consumption while preserving high availability. Modern commodity hard drives do not offer the multi-speed operating modes needed by this work, limiting the usefulness of these approaches.

4. Hard Drive Operation

In analyzing drive power consumption, it is important to have a detailed understanding of how drives operate and their internal make up.

Modern hard drives are composed of (i) magnetic platters, on which data is stored; (ii) a platter spindle motor, responsible for turning the spindle attached to the platters; (iii) read and write heads, which retrieve and commit data to the platters; (iv) read and write head actuator arms, which suspend the read and write heads above the platters; (v) a voice coil actuator, which positions the actuator arms over the correct location on the platters; (vi) and printed circuit board electronics, which consists of a buffer cache, a motor driver, and a main controller.

Power is supplied to hard drives via 12-Volt and 5-Volt supply lines. The 12-Volt supply is responsible for powering the motor to spin the hard drive's platters and supplying power to the read/write arm actuator, while the 5-Volt supply is responsible for powering the read and write heads, the buffer cache, the flash memory, and the main controller. Laptop drives do not make use of a 12-Volt supply line. The internal mechanics of laptop drives is the same as 3.5" drives, but they operate on a 5-Volt supply. This would make it more difficult to discern mechanical power con-

sumption from electronic power consumption. Our initial focus is on server-sized storage infrastructure.

Drive platters spin constantly, only stopping in standby mode. Read/write heads are only powered when reading and writing, and the arm actuator is only powered when seeking to and residing over locations on a platter. Printed circuit board electronics are always powered.

According to the ATA/ATAPI-5 specification and the Advanced Configuration and Power Interface (ACPI), four power management states are supported by modern hard drives: active, idle, standby (or spin-down), and sleep (requires a system reset to recover from this state) [3].

5. Measurement Environment

We have developed a custom measurement platform that provides direct, online measurements of individual PC components. We configured our platform to capture two, 12-bit inputs—the 5-Volt and 12-Volt supply lines of a connected hard drive—at 1100 Hz.

Measurement samples are generated by intercepting the 5 and 12-Volt supply lines with .02 Ohm resistors, leading to a small, but detectable, drop in potential across these resistors. We capture this potential difference using differential amplifiers and digitize the outputs. A microcontroller collects the digital outputs and sends them to a separate PC for logging. Our measurement instrumentation is powered separately and has a negligible effect on the actual power consumption of the drive under test. Pictures of the measurement board and measurement setup are provided in Figures 1 and 2, respectively.

6. Test Setup & Methodology

The test set is comprised of ten hard drives of varying age, capacity, and interface. The test set is chosen specifically to model the diversity of drives available in production environments. A detailed list of drives tested can be seen in Table 1.

Specialized microbenchmarks were created to test each of the drive components individually:

Arm Actuator The arm actuator, powered by the 12-Volt supply, is tested by specifying reads and writes in different locations on the drive and across the entire drive. The power consumption of seeking to specified locations and the power consumed keeping the heads in the correct location up to request completion is captured. To expose differences in power consumption depending on the origination and destination of seeks, measurements are also taken of seeking to different parts of the drive from different starting locations (i.e. inner to outer, center to outer, inner to center, etc.).
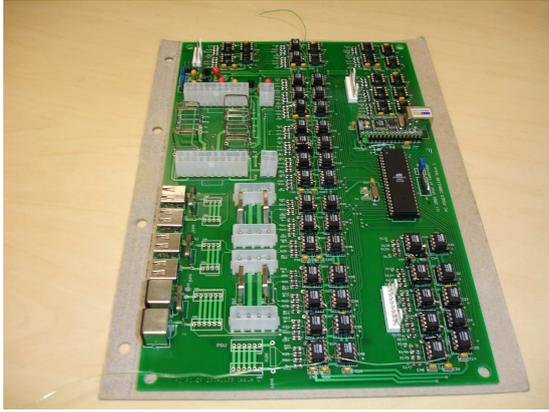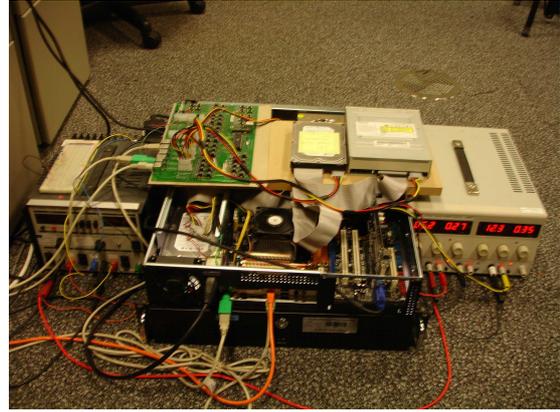
Figure 1. Custom Measurement Board

Figure 2. Measurement Setup

Table 1. Details of Drives Tested

	Make & Model	Capacity(GB)
1	IBM Deskstar IC35L060AVER07-0	61.5
2	Fujitsu MPE3084AE	8.45
3	Seagate ST380215A	80
4	Samsung HD501LJ	500
5	Seagate ST3250820A	250
6	Hitachi Deskstar HDS725050KLA360	500
7	Hitachi Deskstar HDP725025GLA380	250
8	Western Digital WD5000AACS-00ZUB0	500
9	Seagate ST380011A	80
10	Seagate ST373455LW	73

Read/Write Heads The read/write heads, powered by the 5-Volt supply, are isolated by capturing the 5-Volt consumption during data transfers. Varying data sizes were tested, and drive accesses were conducted in different physical locations on the drive and across the entire drive. Transfers were large enough as to eliminate the possible interference of on-drive caches.

Printed Circuit Board Electronics The printed circuit board electronics, powered also by the 5-Volt supply, are always active. Because the printed circuit board electronics are always on, capturing measurements when no request was being serviced allowed us to subtract printed circuit board consumption from read/write head consumption measurements.

Platter Spindle Motor The platter spindle motor, powered by the 12-Volt supply, is only deactivated when the drive is in standby mode. Measuring 12-Volt consumption in idle mode and in standby mode permitted us to differentiate arm actuator power from platter spindle motor power.

7. Results

In this section, we present the results of testing the hard drives listed in Table 1 along with brief analyses. The competitive nature of the hard drive industry meant we could find no literature on the low-level construction or operation of the drives, and hence we were unable to corroborate our speculations as to the causes of the observed trends. Nonetheless, we are confident that the trends themselves are reliable and feel it is useful to present our current reasoning in its speculative form.

These results are most usefully viewed in terms of the drive's operating modes: standby, idle, and reading/writing data. The microbenchmarks of the previous section are combined to explain the power consumption of these modes. Lastly, we investigate the mechanical operation of the spindle motor and the actuator arm.

7.1. Standby Mode

Standby mode measurements were conducted by issuing the ACPI standby command and recording power consumption over a 5-minute interval. Multiple runs were collected to expose any variation in the measurements. Figure 3 presents the energy consumption averages for all the drives tested. Because the mechanical components are off in standby mode, the drive electronics are responsible for the majority of the total energy consumption for this duration of time. The numbers show that even in standby mode (an inactive state where no work is being done), the drive electronics are still active. For the drives tested, the mechanical energy savings of entering standby versus remaining idle range from 92-99%. The electronic component energy savings span a wider range between 35% and 95%. Throughout our testing, we have confirmed that placing drives in standby mode significantly reduces the mechanical power consumption of hard drives, but we believe
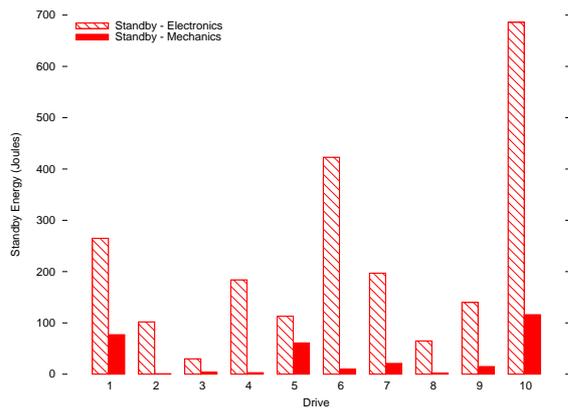
Figure 3. Summary of 5-Minute Standby Energy Consumptions

Figure 4. Summary of 5-Minute Idle Energy Consumptions

that there has not been enough focus on reducing the electronic power consumption for standby and other operating modes.

The electronics should strive to minimize power consumption while the drive is in standby mode. However, from the numbers that we collected, it seems that savings could be greater. The power consumptions for Figure 3 are of the order of hundreds of mW, while modern integrated circuit components in standby/sleep modes consume less than 10 mW.

7.2. Idle Mode

Idle mode power consumption is measured by capturing data when the drive is not servicing I/O requests. Figure 4 illustrates that the hard drive electronics are still a significant power consumer in idle mode as well. In idle mode, the drive is not servicing any requests, but the platters are spinning. The energy consumption values in Figure 4 show that mechanical components responsible for spinning the platters do not dominate the electronics as expected. Moreover, the electronic energy consumption ranged from 25% to nearly 75% of the total.

7.3. Read/Write Power Consumption

We conducted several tests to thoroughly test and understand the energy cost of data access and storage. Our first test measured the energy consumed while reading and writing the entire drive divided into 256 MB bins. This test was devised to get an average energy consumption for reading and writing a given LBN, determining the relationship between power consumption and physical drive location. Next, we measured the power consumption of reading and writing 1 GB of data with block sizes ranging from
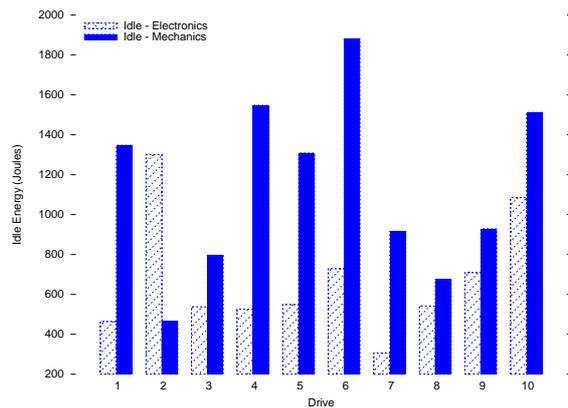
512 B to 8 KB. This test was structured to expose differences in power consumption resulting from different block size specifications.

Location dependence Figure 5 demonstrates that, generally, read and write energy consumptions were dependent on LBN, increasing as LBN increased. Data density (data per track length) increases at higher LBNs simply because the track length becomes smaller. However, as a result of the zoned bit recording employed by modern hard drives, the outer tracks contain more data as they have a larger circumference than the inner tracks. Because hard disk drives rotate at a constant angular velocity, more data can be read and written—and at faster rates—at the outer tracks (lower LBNs) than the inner tracks (higher LBNs). The smaller energy consumption values at the lower LBNs is a result of more sectors being seen by the Read/Write heads for a given angle of rotation as compared to the higher LBNs. In addition to the faster linear velocity of larger radii, the inner tracks cannot store as much data as the outer tracks, potentially leading to added seeking in order to service a single request.

To validate our thinking, We measured read and write bandwidths across all the drives. We observed that read and write bandwidths decreased as LBN increased (Figure 6). Bandwidth decreases as a result of transfers taking longer to complete at higher LBNs. This observation is due to there being less data per track and the fact that the linear velocity of the Read/Write heads decreases as the radius decreases.

All drives tested showed this same behavior of increasing read and write energy consumptions. Low LBNs correspond to tracks with larger radii near the outer parts of the platters, and high LBNs correspond to tracks of smaller radii near the inner edge of the platters. We fit a cubic polynomial to the data points of the read and write curves from
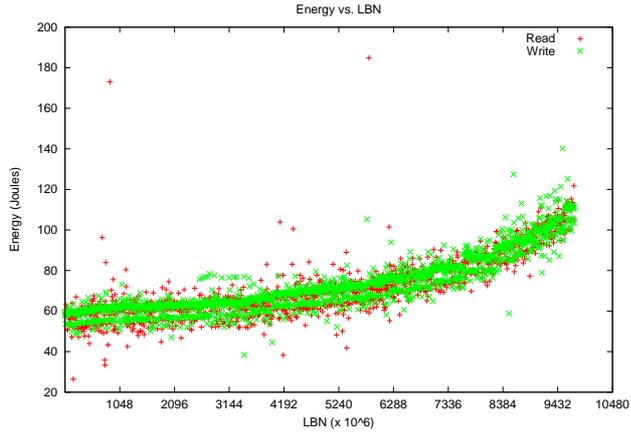
Figure 5. Hard Drive 6, Read and Write Energy Consumptions from low LBN to high LBN
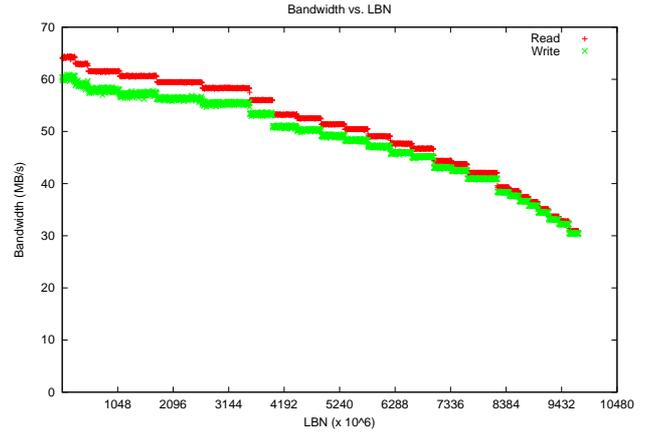

Figure 6. Hard Drive 6, Read and Write Bandwidths from low LBN to high LBN
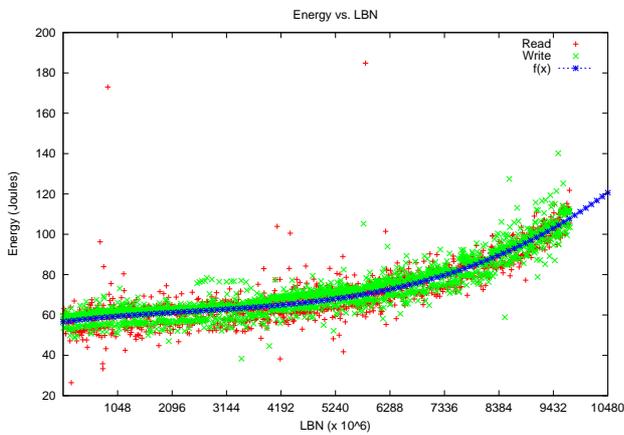

Figure 7. Hard Drive 6, Read & Write Energy vs. Logical Block Number

Figure 5 in Figure 7 with a root mean square (RMS) error of 6.44 Joules.

We saw a consistent trend in read energy consumption across all drives tested. The curve fit to the data points as LBN increases suggests that there is a cubic relationship between read and write energy consumption and LBN, i.e.

$$Energy \propto LBN^3 \qquad (1)$$

Figure 7 and Equation 1 help give some indication as to how a drive's energy consumption is related to the LBN that is accessed. Including this type of consumption model and coefficient values in the technical specifications of hard drives would be a way for drive manufacturers to aid developers in programming with power in mind. This type of knowledge could be used in reducing drive energy con-

sumption further by laying out data in a way that keeps frequently accessed data at low LBNs where the energy required to retrieve the data is lowest. For Drive 6, reading the lowest LBN versus the highest LBN results in a power savings of approximately 14 Watts.

We took great care in ensuring our tests suppressed both write and read caching. We are confident that neither the system cache nor the on-disk caches were a factor in the results we present. System developers would only need to understand the relationships we present with respect to logical block number, however, we could roughly transform this relationship with LBN into a relationship with respect to radius by analyzing our bandwidth and energy/power graphs and estimating the radius of the tracks and zones. 9 of 10 drives were described by this cubic relationship, and RMS errors were less than 8 Joules.

Seeing this read behavior, we then investigated the 5 and 12-Volt components that make up this curve. We decomposed Drive 6's read energy curve (Figure 5) into the 5-Volt and 12-Volt components and plotted them in Figure 8. Both supply lines show a similar shape to the combined curve. It is interesting to note that the 5-Volt energy consumption is just as significant in the total energy consumption as the 12-Volt supply. Here again, we see that both the 5-Volt and 12-Volt power consumptions are equally important, and neither can be ignored when addressing hard drive power consumption.

Finally, Figure 9 considers the relationship between read energy (Figure 5) and bandwidth (Figure 6). Our aim was to uncover any relationship between the performance and energy consumption of Drive 6.

The quadratic fit in Figure 9, RMS = 3.81 Joules, reinforces our previous conclusion of favoring data accesses at lower LBNs (outer edges of the platter) to lower energy
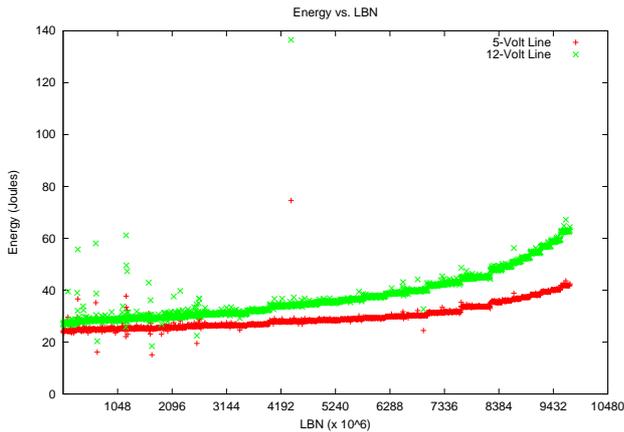
Figure 8. Hard Drive 6, Read Energy Consumption 5 V and 12 V
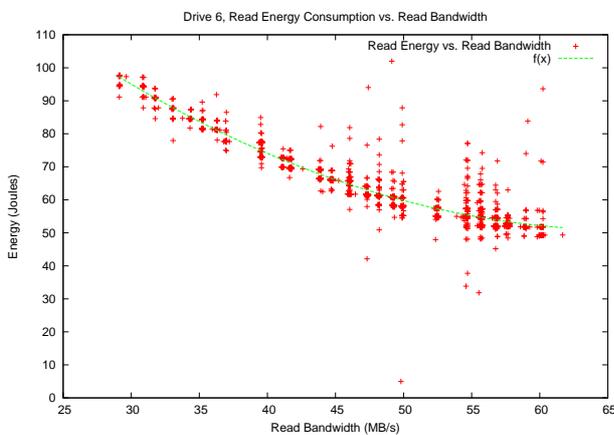

Figure 9. Hard Drive 6, Read Energy Consumption vs. Read Bandwidth

consumption. In addition, this relationship shows that performance is not sacrificed in accessing data at lower LBNs but enhanced. Contrary to many other power-performance optimizations, energy consumption and performance of hard drives seem to have a symbiotic relationship. It is also important to note that this quadratic relationship between energy and read bandwidth suggests that the power consumption per bit is not constant, unlike commonly believed (e.g., equation 1 in [12]). A 1/x fit to this data yielded an RMS of 12.08 Joules. The longer a transfer takes, the bandwidth is lower and the energy consumption is greater. What sounds like a simple inverse relationship is actually a quadratic relationship.

These findings are important because they provide new information that hardware manufacturers and kernel developers could leverage in attempts to reduce drive energy
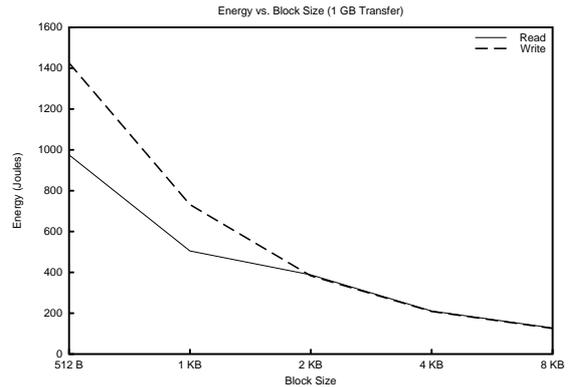

Figure 10. Hard Drive 7, Energy Consumption as Block Size Increases

consumption without sacrificing performance: (i) reading and writing low LBNs for better performance and decreased energy consumption, (ii) writing frequently read data at low LBNs and less frequently read data at higher LBNs, and (iii) using our model to understand the rate at which read and write energy increases with respect to LBN.

Effect of Transfer Size The data to this point considered block sizes of 256 MB. In general, reading consumes more energy than writing for large ($\geq$ 2 KB) block sizes, but the opposite is true for smaller ($<$2 KB) block sizes. Figure 10 shows this happening on Drive 7 as 1 GB is read and written with the block size increasing from 512 B to 8 KB. Transfers are made up of both a fixed and variable cost. The fixed cost is due to the pre-processing and setup of the transfer, while the variable cost is a result of the size of the actual data transfer. Small-block-size transfers are dominated by the fixed cost, and large-block-size transfers are dominated by the variable cost. After observing similar behavior on 7 of 10 drives, it seems that the pre-processing and setup for writes takes longer than that of reads, generally, and this becomes apparent as block size decreases. The longer fixed cost for writes at lower block sizes results in the increased energy consumption, however, once the variable cost becomes the dominant factor, writes consume less energy than reads.

With large-block-size reads consuming more energy than writes, the fact that reads occur at least 4-5 times more than writes [19] becomes more important when attempting to reduce the power consumption of data storage. In addition, digital data is growing rapidly as more digital data is created in a year than there is capacity to store it [2]. The implication of these observations is that more consideration should be given to current data access patterns to obviate unnecessary increases in energy consumption for the future.
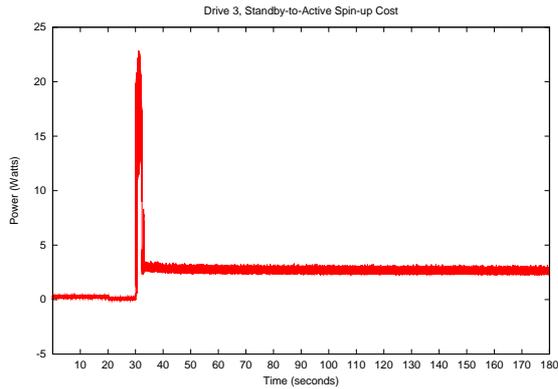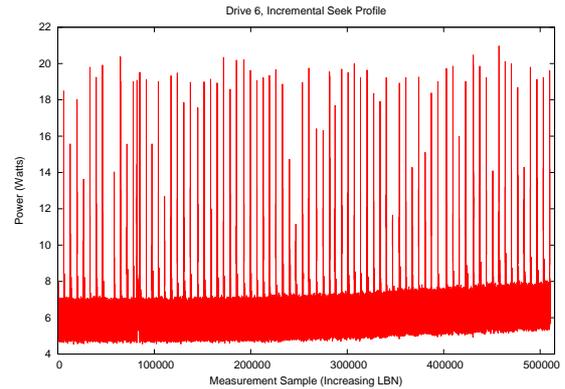
Figure 11. Spin-up Cost of Drive 3

Figure 12. Seek Profile of Drive 6

7.4. Hard Drive Mechanics

Platter Spindle Motor By investigating how the 12-Volt consumption was affected during our tests, we gained insight into the mechanical operation of hard drives. The power consumption of the hard drives' mechanical parts in keeping the platters spinning was fairly constant and consistent across drives. The increased consumption during spin-up transitions varied across the drives. Costs varied between 7.2 Watts and 24 Watts. This transition consumes the greatest power in hard drives, and prior authors have expressed the need to avoid this cost, for the sake of power, performance, and device lifetime, while maximizing power efficiency and performance. The reduction in energy consumption at the expense of increased energy and latency from the subsequent spin-up has been highlighted in previous work [5, 10, 11]. Research has also addressed the potential reduction in device lifetime resulting from repeated spin-down/spin-up transitions [18, 21, 24]. The mechanical power cost of spinning a drive up is large, but short-lived.

Figure 11 shows the spin-up cost of Drive 3. This spin-up transition takes about 2.5 seconds and consumes approximately 37 Joules (5-Volt and 12-Volt combined). Spinning up this drive from idle to active mode takes approximately 1 second and consumes less than 1 Joule (5-Volt and 12-Volt combined). When data is not present in memory and must be retrieved from disk, latency is naturally orders of magnitude larger than typical memory access times. The difference between a spin-up from standby (2.5 seconds) and from idle (1 second) is only 1.5 seconds. And, by being in standby mode versus idling, this particular drive saves the 37 Joules needed to recover the drive from standby after only about 9 seconds in standby.

Our investigation has confirmed previous conclusions that placing a drive in standby is a means of reducing drive energy consumption, barring the drive is spun down for a long enough period of time.

Actuator Arm We structured two additional tests to profile the power consumption cost of seeking to different locations on the platters (zones) from different starting points. The first test grouped logical blocks into 5 GB bins. Our measurements captured the power consumption of seeks originating at bin 0 to all other bins until the end of the drive was reached. Figure 12 shows the results for this test. This was a 500 GB drive, and the seek power costs of seeking from bin 0 to all other bins were fairly constant between 18 and 21 Watts. 75% of the seek power costs were in this band, while the remaining quarter were less than 18 Watts. Drawing attention to the upward trend at higher LBNs, we saw that the 12-Volt line consumes more power to keep the actuator arm at higher LBNs. Positioning the actuator arm at higher LBNs (physically located towards the inner part of the platters) requires more current to pass through the voice coil. Figure 13 shows more specifically that the power cost of seeking from low to high LBNs is about 20 Watts at max and, just like the spin-up cost, is very short lived. The energy consumption for this outer to inner seek spike (LBN 0 to LBN_MAX) is only about 0.60 Joules. Regardless of destination, the power cost of seeking from the beginning of the disk is fairly constant.

The second test was setup to see how much power was consumed from seeks originating from the middle tracks of a platter. We measured the power needed to seek from the center logical blocks to higher LBNs and to lower LBNs in 1 GB block groupings. From a performance standpoint, data placement offers improvement, but from a power perspective, placement does not offer any significant reduction in power consumption as far as seeks are concerned (the seek results were similarly constant like those in Figure 12). Placement could possibly be explored for sustained or frequent data access over extended periods of time for energy savings. For example, since less energy is consumed for reading lower LBNs versus higher ones, frequently read data that is placed at a low LBN will save energy over accessing the same data placed at a higher LBN. More power
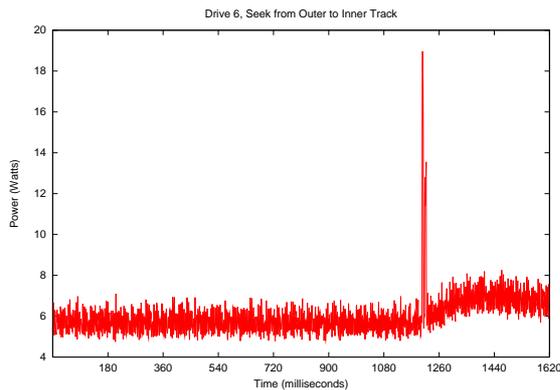
Figure 13. Drive 6, Power Cost of Outer Track to Inner Track Seek

was saved seeking from the center to lower LBNs than from the center to higher LBNs, but the savings were minimal.

8. Conclusion

We set out to uncover hidden power consumption behaviors of modern hard drives in efforts of discovering new possibilities to reduce drive energy consumption. The standby and idle mode energy consumption of the drive electronics could use a great deal of attention. This is an issue that will probably need to be addressed at the microarchitecture level by hard drive manufacturers. However, the disparity of read and write energy consumption at different LBNs and the difference in energy consumptions of same-size transfers specified with different block sizes are both issues that can be exploited by software developers. Investigating the energy consumption of reading and writing has led us to the possibility of reducing energy consumption by writing data that will be accessed frequently at lower LBNs and writing and reading data with the largest possible block sizes. The energy consumed during seeking is minimal, but restricting disk accesses between low and central LBNs could aggregate into savings over extended usage periods.

Storage is only going to increase into the future, which motivates us to reduce the consumption of hard drives in active as well as passive modes. We have presented several opportunities to make improvements in energy consumption, and the cumulative savings will substantially impact computing long-term.

Solid-state drives will have different power consumption profiles, but the current price of these drives makes them an uneconomical option for large storage deployments. Future work will include in-depth analysis of the power consumption of solid-state drives and laptop hard drives as well as reporting on the implementation of

suggestions made here and quantifying the impact on a systems-level.

Acknowledgements

References

[1] IDC serves up 10 storage predictions for 2008, 2007. `http://www.cio.co.uk/news/index.cfm?articleid=2455`.

[2] Study: digital universe and its impact bigger than we thought, 2008. `http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=9067639`.

[3] Advanced Configuration and Power Interface (ACPI) specification, 1999. `http://www.acpi.info`.

[4] D. Colarelli and D. Grunwald. Massive arrays of idle disks for storage archives. pages 1–11. In Proceedings of the 2002 ACM/IEEE International Conference on Supercomputing, 2002.

[5] F. Douglis, P. Krishnan, and B. Bershad. Adaptive disk spin-down policies for mobile computers. pages 121–137. In Proceedings of the 2nd Symposium on Mobile and Location-Independent Computing, 1995.

[6] F. Douglis, P. Krishnan, and B. Marsh. Thwarting the power-hungry disk. pages 292–306. USENIX Winter Conference, 1994.

[7] L. Ganesh, H. Weatherspoon, M. Balakrishnan, and K. Birman. Optimizing power consumption in large scale storage systems. In Proceedings of the 11th Workshop on Hot Topics in Operating Systems (HotOS XI), 2007.

[8] P. Greenawalt. Modeling power management for hard disks. pages 62–66. In Proceedings of the 1994 IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), 1994.

[9] S. Gurumurthi, A. Sivasubramaniam, M. Kandemir, and H. Franke. DRPM: dynamic speed control for power management in server class disks. *ACM SIGARCH Computer Architecture News*, 31(2):169–181, May 2003.

[10] T. Heath, E. Pinheiro, and R. Bianchini. Application-supported device management for energy and performance. In Proceedings of the Workshop on Power-Aware Computing Systems, 2002.

[11] D. P. Helmbold, D. D. E. Long, and B. Sherrod. A dynamic disk spin-down technique for mobile computing. pages

130–142. In Proceedings of the Second Annual ACM International Conference on Mobile Computing and Networking, 1996.

[12] N. Joukov and J. Sipek. GreenFS: making enterprise computers greener by protecting them better. pages 69–80. In Proceedings of EuroSys 2008, 2008.

[13] D. Li and J. Wang. EERAID: energy efficient redundant and inexpensive disk array. page 29. In Proceedings of the 11th workshop on ACM SIGOPS European Workshop, 2004.

[14] D. Narayanan, A. Donnelly, and A. Rowstron. Write off-loading: practical power management for enterprise storage. In Proceedings of the 6th USENIX Conference on File and Storage Technologies (FAST), 2008.

[15] A. E. Papathanasiou and M. L. Scott. Increasing disk burstiness for energy efficiency. Technical Report TR792, University of Rochester. pages 1–31, 2002.

[16] E. Pinheiro and R. Bianchini. Energy conservation techniques for disk array-based servers. pages 68–78. In Proceedings of the 18th International Conference on Supercomputing, 2004.

[17] E. Pinheiro, R. Bianchini, and C. Dubnicki. Exploiting redundancy to conserve energy in storage systems. *ACM SIGMETRICS Performance Evaluation Review*, 31(1):15–26, Jun 2006.

[18] E. Pinheiro, W.-D. Weber, and L. A. Barrosa. Failure trends in a large disk drive population. In Proceedings of the 5th USENIX Conference on File and Storage Technologies, 2007.

[19] D. Roselli, J. R. Lorch, and T. E. Anderson. A comparison of file system workloads. pages 41–54. In Proceedings of the 2000 USENIX Annual Technical Conference, 2000.

[20] S. W. Son, G. Chen, and M. Kandemir. Disk layout optimization for reducing energy consumption. pages 274–283. In Proceedings of the 19th Annual International Conference on Supercomputing, 2005.

[21] C. Weddle, M. Oldham, J. Qian, A.-I. A. Wang, P. Reiher, and G. Kuenning. PARAID: a gear-shifting power-aware raid. *ACM Transactions on Storage*, 3(3):33, Oct 2007.

[22] J. Wilkes. Predictive power conservation. Technical Report HPL-CSP-92-5, Hewlett-Packard Laboratories. page 1, February 1992.

[23] J. Zedlewski, S. Sobti, N. Garg, F. Zheng, A. Krishnamurthy, and R. Wang. Modeling hard-disk power consumption. pages 217–230. In Proceedings of the 2nd USENIX Conference on File and Storage Technologies, 2003.

[24] Q. Zhu, Z. Chen, L. Tan, Y. Zhou, K. Keeton, and J. Wilkes. Hibernator: helping disk arrays sleep through the winter. *ACM SIGOPS Operating Systems Review*.

[25] Q. Zhu, F. M. David, C. F. Devaraj, Z. Li, Y. Zhou, and P. Cao. Reducing energy consumption of disk storage using power-aware cache management. pages 118–129. In Proceedings of the 10th International Symposium on High Performance Computer Architecture (HPCA), 2004.