Leveraging Disk Drive Acoustic Modes for Power Management

Doron Chen, George Goldberg, Roger Kahn, Ronen I. Kat, Kalman Meth
{cdoron,georgeg,rogerk,ronenkat,meth}@il.ibm.com
IBM Research - Haifa, Israel

***Abstract*—Reduction of disk drive power consumption is a challenging task, particularly since the most prevalent way of achieving it, powering down idle disks, has many undesirable side-effects. Some hard disk drives support *acoustic modes*, meaning they can be configured to reduce the acceleration and velocity of the disk head. This reduces instantaneous power consumption but sacrifices performance. As a result, input/output (I/O) operations run longer at reduced power. This is useful for power capping since it causes significant reduction in peak power consumption of the disks.**

We conducted experiments on several disk drives that support acoustic management. Most of these disk drives support only two modes - quiet and normal. We ran different I/O workloads, including SPC-1 to simulate a real-world online transaction processing workload. We found that the reduction in peak power can reach up to 23% when using quiet mode. We show that for some workloads this translates into a reduction of 12.5% in overall energy consumption. In other workloads we encountered the opposite phenomenon-an increase of more than 6% in the overall energy consumption.

I. Introduction

Achieving the power and cooling requirements of enterprise data centers has long become a burden to the growing needs of business requirements. The continuous demand for increased enterprise storage space fuels the growth of storage power consumption. IDC [27] estimates the storage capacity compound annual growth rate (CAGR) at 50.9%. The utilization of more and more RAID arrays leads to an increase in storage power consumption. Disk drives account for 13%-20% of the cost of powering and cooling a data center [28], [33]. As a result, finding the optimal balance between storage power and performance has become increasingly important [23], [32].

The term *acoustic management* refers to the ability to limit the allowable acoustics of a disk drive when performing a seek operation. The primary applications of acoustic management, or quiet disks, are in consumer devices and enterprise systems. Consumer devices, such as Internet appliances or consumer electronics, operate in relatively quiet environments; therefore, storage fitted for such devices should not increase the overall system acoustics. On the other hand, enterprise disk drives are densely packed in order to keep pace with new performance requirements. The additive nature of acoustics requires a reduction of the noise emitted from disks. Enterprise disks are fitted into high-end workstations, which are quieter than servers. Therefore, workstation storage noise should not exceed system noise.

We show how to utilize acoustic management for the power and energy management of disk drives. Specifically, we show how slowing down of the disk seek operations by using acoustic management reduces the *power consumption* of the disk. Intriguingly, a reduction of the power consumption may lead to an increase of the overall *energy consumption* in some cases. We examine the effect of using acoustic management on power consumption, energy consumption and performance for various scenarios and workloads.

Power is measured instantaneously, whereas *energy* is the overall power consumption over a given interval. Energy is closely related to carbon emissions and costs. Data centers usually have a given power budget defined by the data center's internal power distribution, its cooling capacity, and the energy company's ability to supply power. A power budget may change due to factors such as cooling limitations or problems at the energy company. When the power budget decreases, the data center is required to reduce, or cap, its power consumption.

The disk drive is the primary storage medium in today's storage systems. A typical disk drive operates on two electrical power sources: a $12V$ power source to operate the mechanical components, and a $5V$ power source for the electrical components. The mechanical components include the spindle motor and the voice-coil actuator (seek head). The electrical components include the onboard processor, the cache memory, the physical I/O channel, and the magnetic head. Details on how these components affect disk power consumption are presented in [30], [32].

When disk drives are powered on, the platters commence spinning, accelerating to the required operational rotational speed, and remain spinning at a constant speed. In enterprise storage systems, the platters are required to spin at a given constant speed in order to serve I/O requests in a timely manner. Therefore, most disk drives spin at a constant speed even when no I/O is being performed (i.e., *idle*). The power consumed in this state is called the *static power* of the disk. When an I/O operation is requested, the seek head of the disk drive must move, also consuming power. Moreover, transferring data to and from the disk consumes additional power. The power consumed by these I/O operations is called the *dynamic power* of the disk.

The power consumed by the disk drive is the sum of the dynamic and static power. Generally speaking, the static power is at least two-thirds of the total power. However, this

978-1-4244-7153-9/10/$26.00 2010 IEEE

breakdown depends on the workload.

The ATA/ATAPI-6 specification [3] defines automatic acoustic management (AAM) to be a standardized way of setting limits on the disk acoustics (noise) during seek operations. AAM defines quiet (slow) and performance (fast) seek modes. This is achieved by controlling the acceleration of the disk head during seek operations. Controlling the acceleration is done by reducing the energy provided to the seek actuator, which then reduces the dynamic power consumed by the disk drive. The AAM specification allows a range of acoustic modes between the values of 128 and 254. However, most vendors have implemented only two modes - a quiet mode (128) and a normal mode (254).

Seagate's *Just-In-Time (JIT)* is another implementation for controlling the acceleration and velocity of a disk drive head [2]. Using uniform acceleration and velocity when performing seek operations from one logical block address (LBA) to another, the seek head may arrive before the data is under the head. The JIT mechanism computes the time available until the requested LBA is under the disk head, and then adjusts the seek speed accordingly. The mechanism performs a faster seek if the LBA is close, and a slower seek if the data is nearly a revolution away. Thus, reducing the speed of the disk head does not incur a performance reduction, but in some cases allows reduction of the noise and power.

Our contribution: We performed a detailed analysis of normal and quiet acoustic modes, and showed that the quiet acoustic mode helps us control the storage's instantaneous power consumption, effectively enabling us to budget (or cap) the power consumption of the storage. Controlling the instantaneous power consumption is important when the data center has insufficient cooling or power. In such scenarios, the ability to perform capping may prevent the need to power off machines. In addition, we show that for some workloads, using quiet mode leads to energy savings, and we can trade performance for overall energy savings. In other instances, using quiet mode leads to an increase in the overall energy consumption for the same work. In these cases, the tradeoff is between an increase in energy consumption and a reduction in instantaneous power.

This paper is organized as follows: Section II discusses our research methods and methodology. In Section III, we review power capping using acoustic modes, followed by Section IV, in which we discuss cases in which the overall energy consumption can or cannot be reduced. In Section V, we investigate the performance considerations. Section VI presents the current state of the art, and finally, in Section VII, we summarize our work.

II. Methodology

For our investigation of acoustic modes we measured the performance and power consumption of disk drives. We used a custom-made power measurement setup (see Figure 1) that is comprised of the following components: (i) An external TDK-Lambda laboratory power supply which supplies 5V and 12V to the disk drive. (ii) A measurement circuit made up of two 0.1Ω resistors connected in series to the 5V and 12V power sources (one resistor for each source). (iii) A National Instruments (NI) digital acquisition card (DAQ) model NI PCI-6230 that samples the differential voltage on the measurement circuit resistors. The DAQ card is connected to a dedicated workstation. (iv) NI Labview 8.6 measurement suite [4].

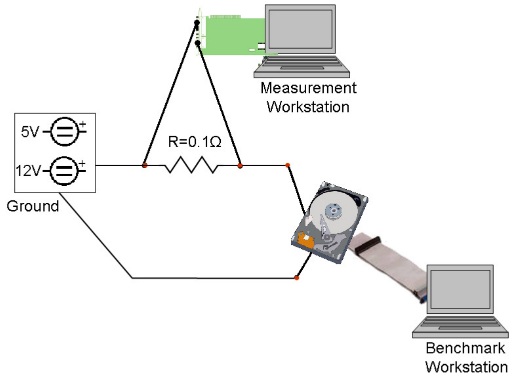


Figure 1. The custom-made setup for running performance and power consumption benchmarks.

We use a custom-made Labview application for measuring the power consumption of the disk drives. For measuring I/O workloads we sampled the 5V and 12V power source at 10K samples per second and then averaged the result per second. For a detailed analysis of the disk operations, we used 50K samples per second and analyzed the raw (not averaged) data. In those cases, we limited ourselves to short sampling intervals (due to the number of samples).

Vdbench [26], a Java-based open-source tool, was used for running I/O workloads. We ran random access micro-benchmarks using Vdbench to observe the effect of the differences in seek operations in normal and in quiet acoustics.

In addition, in order to understand the effect on real-world workloads, we ran an industry standard SPC-1-like workload [1]. The SPC-1 workload is a synthetic, yet sophisticated and fairly realistic, online transaction processing (OLTP) workload. The benchmark simulates real-world environments that are seen in enterprise storage systems [25].

III. Power Capping

Quiet acoustic mode reduces the dynamic portion of the disk power consumption. This allows us to reduce the power budget for the storage, and to conform to a given power usage profile. Common scenarios for reducing the power budget are cooling limitations (e.g., in thermal emergencies) or limitations enforced by the energy company. Data centers that have the ability to dynamically cap their power consumption can operate in reduced performance modes, without shutting down services. Enabling an enterprise data center to provision its power and cooling requirements without over-provisioning for the worst-case scenario is another important application of power capping. In these cases, a data center can reduce its power budget to accommodate temporary fluctuations in power and cooling, and still stay within its operating parameters.

We studied several disk drives that support acoustic management. Most of the results that we show are for a high capacity Seagate HUA721010KLA330 1TB 3.5" disk drive, but some results are for a high performance Western Digital WD30000BLFS 300GB 2.5" disk drive. These disks have unique characteristics that affect how they behave in normal and quiet modes. We identify the main differences in the section below.

A. Seek Analysis of Acoustic Modes

We analyzed the different seek behavior of normal and quiet modes by reviewing the power profile of a single seek operation. We sampled the power dissipation of a *long-range* seek, from one end of the platter to another, at a rate of 50K samples per second. Figure 2 shows the detailed power dissipation of two consecutive seek operations on the 1TB 3.5" disk drive. The diamond shapes represent the 12V instantaneous power dissipation (see the left Y-axis). Changes in 12V power consumption mainly reflect movement of the disk head. The square shapes represent the 5V instantaneous power dissipation (see the right Y-axis). Changes in 5V power consumption mainly reflect data-transfer. Figure 2a details the power dissipation for normal-mode seek, while Figure 2b describes the power dissipation for quiet mode. We observe the power dissipation for the different phases of a seek operation:

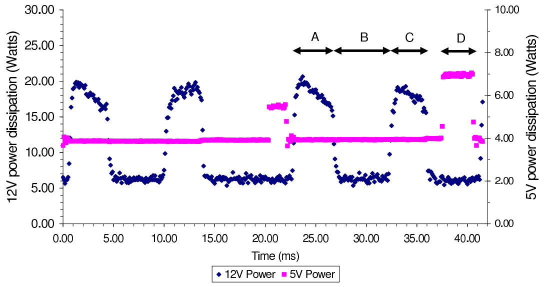
Acceleration:
: During this phase the seek head accelerates to its maximum speed. In quiet mode the acceleration is slower, so the power dissipated at any given time throughout this phase is less than in normal mode.

Coast:
: In this phase the disk head remains at its maximum speed (the maximum speed of quiet mode is slower than that for normal mode). The power dissipated at any time throughout this phase is about the same in both normal and quiet modes. This phase lasts longer in quiet mode, causing more overall energy to be consumed per seek.

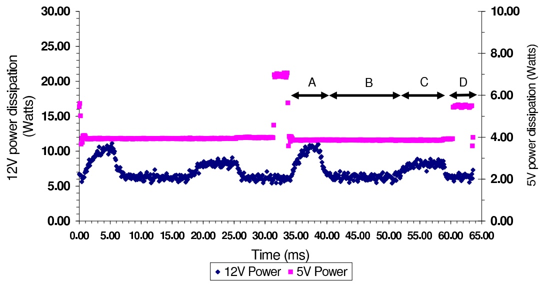
Deceleration:
: During this phase the disk head is slowed down by reversing the current direction of the voice coil motor (VCM). The power dissipation is generally similar to that of the acceleration phase. In quiet mode less power is dissipated at any given time.

Data transfer:
: During this phase, the disk head is at a complete standstill, and the data is being transferred to and from the disk. The 5V power dissipation increases, but the behavior is the same in both normal and quiet mode.

Although the power in quiet mode can be capped at 73% of the power dissipated in normal mode, the overall energy consumed by a long-range seek operation is greater in quiet mode. This is due to: *a*) the fact that the duration of the long-range seek is 53% longer in quiet mode than in normal mode; and *b*) only 12V power decreases during acceleration and deceleration, while the 5V power is almost constant throughout the seek. In the next section, we will address cases in which executing workloads in quiet mode can lead to energy savings.

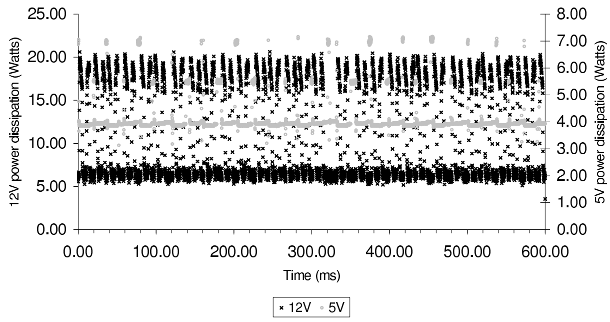


(a) Normal Mode

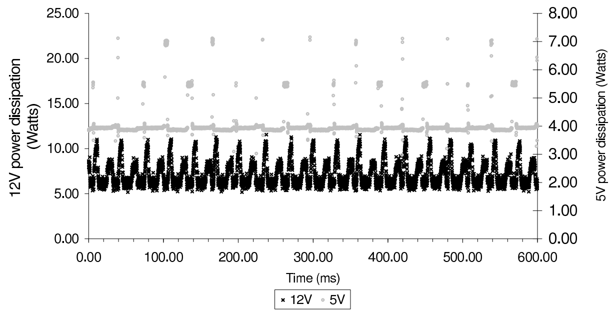


(b) Quiet Mode

Figure 2. Power profile of two successive seeks using normal and quiet acoustic modes. The phases of the seek operation are marked by: A) Acceleration, B) Coast, C) Deceleration, and D) Data Transfer.



(a) Normal Mode



(b) Quiet Mode

Figure 3. Comparison of instantaneous power dissipation in normal and quiet acoustic modes.

Analysis of the 12V and 5V power consumption of the long-range seek operation shows that overall energy consumed by a single seek operation is 17% greater for quiet mode than for normal mode. Due to the longer seek time in quiet mode, the 5V energy consumption increases by some 49% and the 12V energy consumption increases by about 3%. Figure 3 is a detailed power dissipation analysis for a consecutive execution of long-range seek operations. The 5V power dissipation is shown in gray and the 12V power dissipation is shown in black. The figure clearly shows that the reduction in peak power for quiet mode results from reductions in the 12V power dissipation only.

B. Trading Performance for Power Reduction

The power consumption of the disk is determined by the workload. In Figures 2 and 3, we have investigated the power dissipation of the consecutive seek operations. Now we turn to investigate the power reduction when running concurrent random-read workloads for various sizes of data transfers.

Figures 4 and 5 show the power consumption and performance of two types of disks for different sizes of data transfers. Figure 4 shows the power and performance of the 1TB 3.5” disk drive, and Figure 5 presents the data for the 300GB 2.5” disk drive. Figures 4a, 4b, 5a, and 5b show the power consumption for different I/O rates of normal and quiet modes. Figures 4c and 5c show the reduction in power consumption between normal and quiet modes for different I/O rates. We see that the power reduction, displayed here in percentages, increases with the requested I/O throughput.

Figures 4d, 4e, 5d, and 5e show the changing response time for different I/O rates of normal and quiet modes. Figures 4f and 5f show an increase in response time due to the use of quiet mode rather than normal mode. We see that the penalty is more than a 2× factor for a 3.5” disk for most workloads and between a factor of 1.4× and 1.8× for a 2.5” disk for most workloads.

The differences in latency between the 2.5” and 3.5” disks is due to the different seek behaviors. The 2.5” disk is rated at an average seek time of 4.2ms[1]. We measured an end-to-end average response time of 8.2ms in normal acoustic mode and 10.2ms in quiet acoustic mode. Thus, quiet mode adds 2ms on average. The 3.5” disk is rated at an average seek time of 8.2ms[2]. We measured an end-to-end average response time of 13ms in normal acoustic mode and 18ms in quiet acoustic mode. Therefore, quiet mode adds an average of 5ms for the 3.5” disk. For the 3.5” disk, quiet mode increases the seek latency by 60%, as compared to an increase of 48% for the 2.5” disk. Interestingly, in the normal acoustic mode, the 2.5” disk serves requests 58% faster than the 3.5” disk. In the quiet acoustic mode, the 2.5” disk serves requests 75% faster than the 3.5” disk.

Though the absolute power reduction is different for the two disk types, the relative power reduction (in percentages) is about the same, and is nearly linear to the I/O workload rate. At the highest I/O rate, the reduction in peak power is approximately 22%.

IV. Energy Saving or Energy Wasting

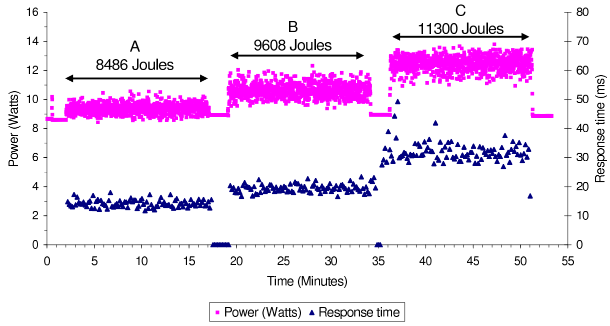
The previous section showed that quiet mode leads to a reduction in peak power consumption. Now we turn to investigate the energy consumption of various workloads when using quiet mode.

[1]This is the internal disk seek time and does not include end-to-end latency.
[2]Internal disk seek time.

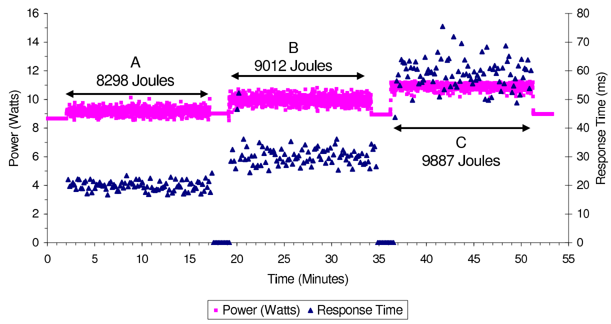
Table I
Energy savings for various SPC-1 workloads.

SPC-1 I/O rate	Normal mode	Quiet mode	Energy difference
10	8486 J	8298 J	−2.2%
25	9608 J	9012 J	−6.2%
50	11300 J	9887 J	−12.5%

One effect of running in quiet mode is the fact that moving the disk head takes longer - that is, the seek time increases. Since the disk power consumption has a static component, a seek operation that takes longer may consume more energy, depending on the balance between the saved energy of the slower acceleration and deceleration and the added energy for longer seek time.



(a) Normal mode

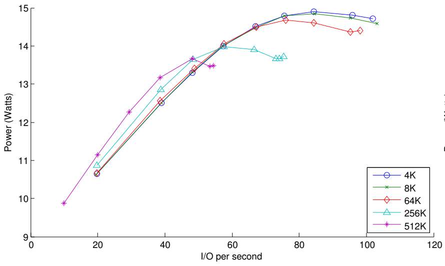


(b) Quiet mode

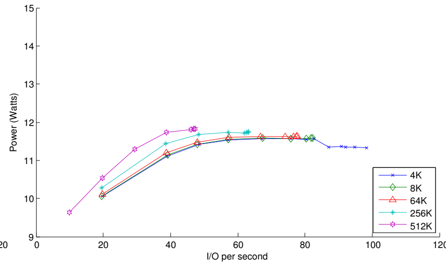
Figure 6. Execution of a SPC-1 trace using normal and quiet modes on the 3.5” 1TB disk.

We simulated a real-world online transaction processing workload by running an SPC-1 workload. SPC-1 workload is a concurrent workload composed of random reads, random writes, and sequential access across various parts of the disk drive. We generated an SPC-1 I/O trace and replayed the I/O trace in normal and quiet modes. Figure 6 describes the power consumption and response time of the trace replay in normal (Figure 6a) and in quiet (Figure 6b) modes. The run consists of three parts: *A*) SPC-1 I/O execution at 10 I/Os per second, *B*) SPC-1 I/O execution at 25 I/Os per second, and *C*) SPC-1 I/O execution at 50 I/Os per second. We measured the power consumption and computed the energy in Joules of each step of the run. Table I summarizes the energy difference of each workload. A negative energy difference value represents an energy saving in quiet mode, as compared to normal mode.

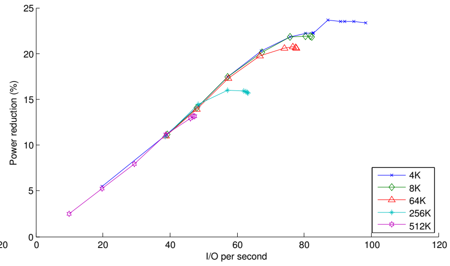
In this benchmark, we executed I/Os at the same rate for both normal and quiet modes. Figure 6 shows that for the



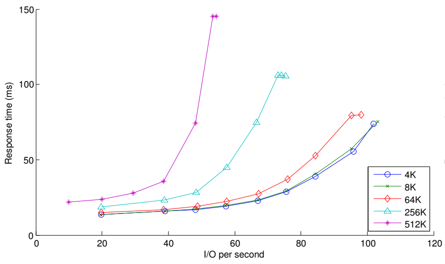
(a) Normal mode power consumption.



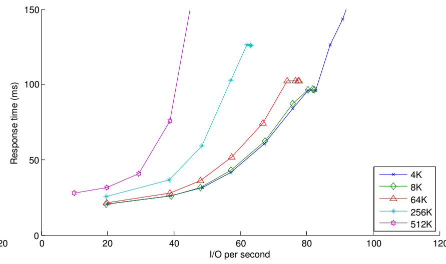
(b) Quiet mode power consumption.



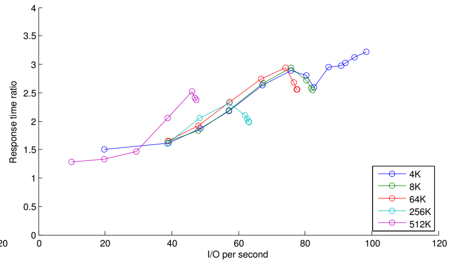
(c) Quiet vs. normal acoustics power reduction.



(d) Normal mode response time.

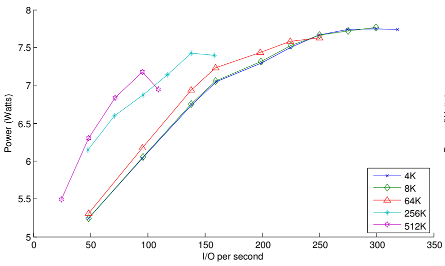


(e) Quiet mode response time.

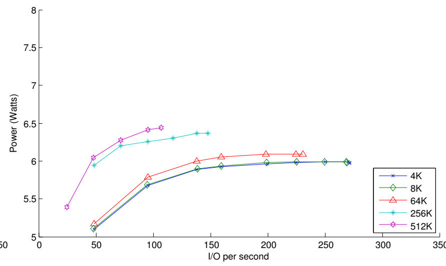


(f) Quiet vs. normal response time ratio.

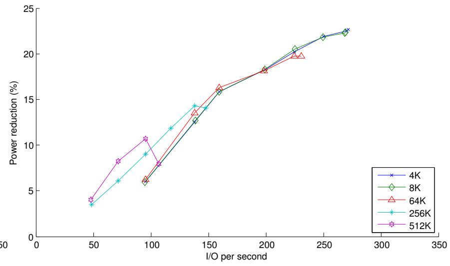
Figure 4. Acoustic modes power and performance with different data transfer sizes for a 3.5” 1TB disk drive.



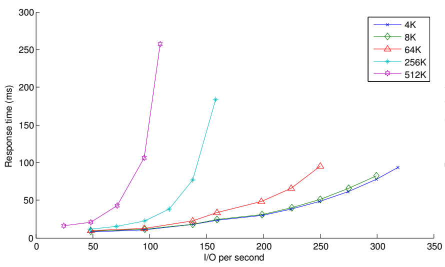
(a) Normal mode power consumption.



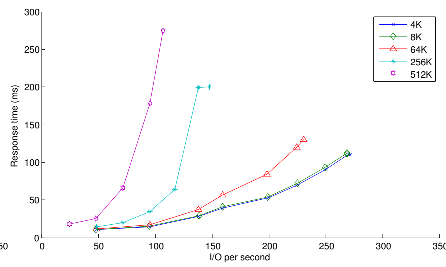
(b) Quiet mode power consumption.



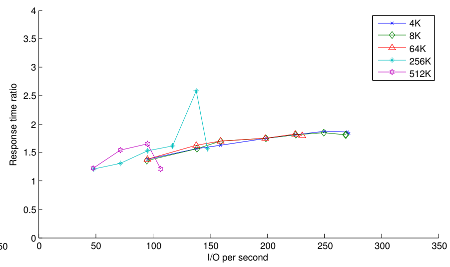
(c) Quiet vs. normal acoustics power reduction.



(d) Normal mode response time.



(e) Quiet mode response time.



(f) Quiet vs. normal response time ratio.

Figure 5. Acoustic modes power and performance with different data transfer sizes for a 2.5” 300GB disk drive.

low I/O rates, 10 and 25 I/Os per second, the response time increases slightly. In these cases, when running in normal mode, the disk is in fact idle in between some I/O operations. In quiet mode, the seeks take longer, and the disk has less or no idle time. In this case, we exchange wasted disk idle time with a longer and slower seek. Running at 50 I/Os per second results in little or no idle time, even in normal mode. The I/O requests are generated at the same rate both in normal and quiet modes. However, in quiet mode, the disk serves these requests at a slower rate, which may cause a longer queue of I/Os to form based on the fact that the response time doubles.

Since the queue in quiet mode is longer, the disk can perform better queuing optimizations that reduce seek distances, as compared to in normal mode. This leads to a reduction in the overall energy cost of the seek operations.

Next, we turn our attention to workloads for which the use of quiet mode leads to an *increase* in the overall (total) energy consumption.

We generated a trace of $30,000$ random-read I/Os. We executed the trace using $1, 2$, and 4 I/O threads in both normal and quiet modes. The I/Os were executed one after the other, without delay. We measured the power consumption and computed the energy consumption of each step.

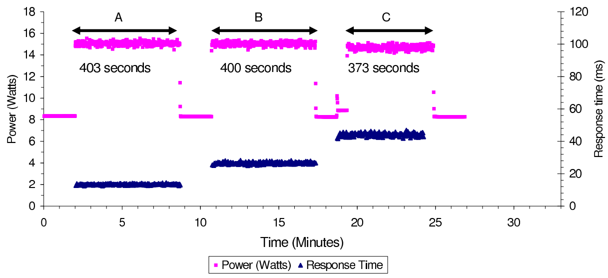
Figure 7 shows the results of replaying the trace in normal

Table II
ENERGY INCREASE FOR RANDOM READ WORKLOADS.

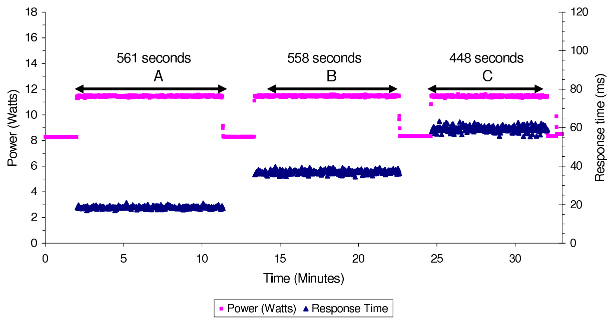
Number of concurrent requests	Normal mode	Quiet mode	Energy difference
1	6063 J	6419 J	+5.9%
2	6001 J	6390 J	+6.5%
4	5234 J	5122 J	−2.2%

mode (Figure 7a) and in quiet mode (Figure 7b) with: *A*) one I/O thread, *B*) two concurrent I/O threads, and *C*) four concurrent I/O threads. The energy consumption is summarized in Table II. A positive energy difference value means that running in quiet mode consumes more energy than running in normal mode. A negative value reflects energy savings.

Reducing queuing to little or none (i.e., when using no more than two I/O threads), energy consumption is notably increased in quiet mode. This is mainly due to longer seek times, which in turn leads to a longer run time. An increase in the number of I/O threads leads to energy reduction. In those cases, a queue is formed and the queuing optimization of seek distances is performed. This reduces the seek coast phase, and allows us to benefit from power reduction during the acceleration and deceleration phases.



(a) Normal mode



(b) Quiet mode

Figure 7. Execution of a random read I/O trace using normal and quiet modes on the 3.5" 1TB disk.

V. APPLICATION PERFORMANCE

We now analyze application performance in different acoustic modes. Here we compare real benchmark runs of the same length using different acoustic modes. Applications' I/O workloads differ from one another by the number of I/O threads and by the access patterns. First, we show the behavior of a random-read workload with different numbers of I/O threads. Each I/O thread performs 4KB random-read operations at maximal rate. Figure 8 shows thread performance and power reduction. The vertical bars show the required number of threads needed to reach the given I/O rate using normal and quiet modes. The curve shows the reduction in power consumption when using quiet mode rather than normal mode.

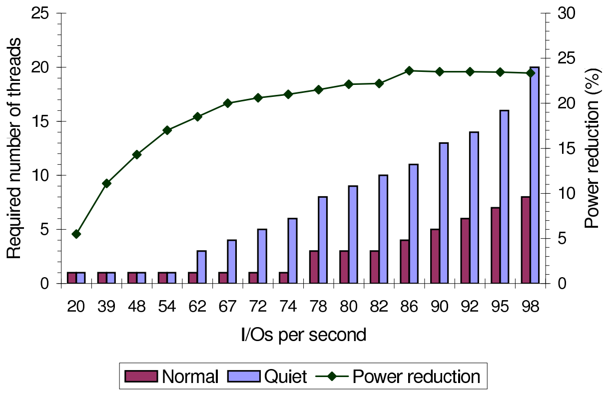


Figure 8. The power reduction when using quiet acoustics in different I/O throughput versus the required thread concurrency for the 3.5" 1TB disk.

Figure 8 shows that single-threaded applications experience a reduction of 27% in I/O throughput when using quiet mode (comparing 74 I/Os per second, the maximal single-threaded throughput at normal mode, with 54 I/Os per second, the maximal single-threaded throughput at quiet mode). In contrast, multi-threaded applications can reach a throughput of almost 100 I/Os per second. Each I/O thread will experience a reduced I/O rate and a longer response time (not shown in the graph) as compared to normal mode. Note that both single-threaded and multi-threaded applications can enjoy nearly maximal peak power reduction when switching to quiet mode.

Next, we ran an SPC-1 workload using different I/O rates in both normal and quiet modes. Here we ran the actual SPC-1 workload, not a replay of the SPC-1 workload. Figure 9 shows the response times and power reduction when running SPC-1 at various I/O rates in normal and quiet modes. The vertical bars show the response times in normal and quiet modes, and the curve shows the reduction in power consumption when using quiet mode instead of normal mode. When using quiet mode, we managed to get up to only 55 I/Os per second. Therefore, for rates beyond 55 I/Os per second, we show only the results for normal mode, and the power reduction shown is the difference between the normal mode power and the quiet mode power at 55 I/Os per second. We see that up to 20 I/Os per second, the response times of normal and quiet modes are similar. However, from 25 I/Os per second upwards, the latency difference clearly increases. In quiet mode, the seek time is greater than in normal mode, which is reflected in an increasingly longer queue as requests are piled up. Recall that in quiet mode, the disk serves requests more slowly. Requests are generated at a given (fixed) rate and are queued until they are served by the disk.

Note that the SPC-1 workload is not a purely random-read workload, but is a combination of random reads, random writes and sequential access across various parts of the

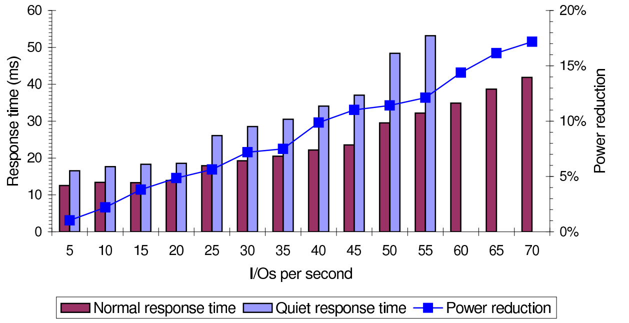
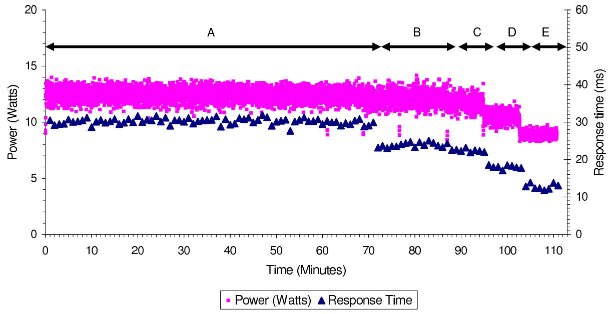
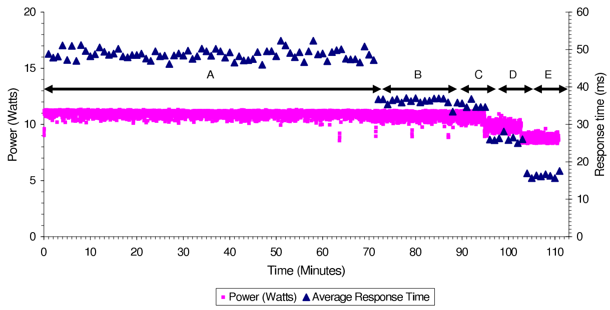


Figure 9. Response time and power reduction for SPC-1 workload using different I/O rates for a 3.5" 1TB disk.

disk drive. The difference between the SPC-1 workload and random-read workloads is evident when comparing the power reduction for the same I/O rates, as shown in figures 8 and 9.



(a) Normal mode



(b) Quiet mode

Figure 10. SPC-1 benchmark run running on a 3.5" 1TB disk using normal and quiet acoustic modes.

Figure 10 shows the power consumption and the response times of an SPC-1 benchmark sequence on the 3.5" 1TB disk using normal (Figure 10a) and quiet (Figure 10b) modes. Note that this is not a replay, but the actual workload generated by SPC-1. Our SPC-1 benchmark sequence run has five[3] steps: *A*) a warmup and maximal performance load of 50 SPC-1 I/Os per second; *B*) 95% and 90% of the maximal performance, which translates to 45 I/Os per second; *C*) 80% of the maximal performance, which translates to 40 I/Os per second; *D*) 25 I/Os per second (50% of maximum performance); and *E*) 10% of the maximal performance, which translates to 5 I/Os per second.

We clearly see a reduction in instantaneous power and a degradation in response times when using quiet mode in intervals A and B. In intervals C and D, we see a reduction in power for both normal and quiet modes, but the instantaneous power in quiet mode is still somewhat lower. In interval E we see that the instantaneous power is similar for both normal and quiet modes. This is mainly due to the low workload rate, which enables little savings in quiet mode.

[3]The SPC-1 sequence has six steps. Step B is, in fact, composed of two steps. One step for 95% of maximal performance and another step for 90%. However, in our case, both steps results in the same 45 I/Os per second rate.

VI. Related Work

Power management for storage has received much attention in recent years. One key approach is spinning down disks during idle periods, and spinning up whenever an I/O request arrives. The main problem with this approach is the considerable cost in latency, as we need to wait for the disk to spin back to operational speed. Thus, this method of power savings is very dependent on whether the specific application can tolerate high latency when spinning up disks in order to access the data. Another concern is that spinning the disks up and down too often may result in an actual increase in energy consumption. This is due to the fact that the spin-up energy cost is considerably more than that required to keep the drive spinning for a significant length of time. Therefore, spinning a disk down is not effective unless the lengths of the idle periods exceed a certain threshold [14]. This is generally not the case in enterprise storage systems. Li et al. [18] deal with the question of when to instruct the disk to enter a low power mode. Douglis et al. [10] describe a method for dynamically varying the spin-down threshold on mobile computers, by adapting to the users' access patterns and priorities.

A large body of research deals with multiple speed disks, also known as dynamic RPM (DRPM). While acoustic modes affect the disk-head speed, DRPM deals with the disk's rotational speed. In notable contrast to acoustic modes, there are currently no available disks that support DRPM. The works of Carrera et al. [7], Gurumurthi et al. [12], Li et al. [18], Pinheiro and Bianchini [22], and Zhu et al. [33] show that adapting the disk rotational speed to the required performance level can reduce power consumption.

Another way to reduce power consumption is to use a massive array of idle disks (MAID)-which is an array of disks in which the majority of the disks are idle and therefore can be powered down. MAID employs two approaches: (a) caching-based, described by Colarelli and Grunwald [8], which uses cache disks that are always active, and (b) migration-based, described by Pinheiro and Bianchini [22], which places data with similar access patterns and frequencies on the same disk, minimizing the number of active disks. As MAID is based on disks spinning down and up, when accessing data placed on a powered-down disk, MAID leads to significant increases in response time. Another problem with MAID is that certain workloads may lead to increased energy consumption, due to frequent spin down and up of disks. To date, MAID has not been adopted for mainstream enterprise storage, but is an alternative for tape drives and offline storage. Another type of system is the *near-online* system, which can handle the longer response time of seconds, when needed.

Storer et al. [24] propose a novel way to use disks, rather than tapes, as a long term storage medium, while reducing the energy cost inherent in disks. They propose going beyond what MAID does and adding a small NVRAM at each node, which would store crucial data and allow for deferred writes and other housekeeping, all while the disk is powered off.

Otoo et al. [21] translated file allocation strategy into disk allocation strategy. This method enables the spinning down of disks that are not likely to be used soon.

Strunk et Al. [9] propose provisioning a storage system using utility functions as a method of tradeoff between cost and benefit. Our leveraging of acoustic modes for power savings could be used in such utility functions, which could then determine the cost/benefit of a storage system.

Huang et al. [15] propose a technique to reduce seek time and seek power by dynamically replicating data within disks, according to the access patterns of the disk head. This method reduces the distance that the disk head must travel.

Energy-efficient storage relies on tools and frameworks that allow us to measure or estimate the storage energy consumption. Allalouf et al. [5] present a method of providing workload-aware power estimation for storage systems. Their model translates frontend I/O operations into the number of backend disk operations, including both seeks and data transfers. It then uses interpolation to estimate the power consumption of the storage system based on the number of disk operations.

Kim et al. [16] present an infrastructure for studying the performance and thermal behavior of storage systems. By carefully modeling the details of a seek, they can account for the heat generated in the acceleration/deceleration phases, and the coast phase in between acceleration and deceleration, when the disk can possibly cool down. They have conducted detailed micro-benchmark studies which understand how I/O characteristics can affect the temperature of disk drives. Their model is based on the observation that the two governing performance parameters affecting the thermal model are the seek activity and any RPM changes of the disk platter (if a multi-speed/DRPM disk is used).

Gurumurthi et al. [13] suggest a new design for storage systems, which would provide considerable power savings while maintaining high performance. Rather than using a large number of disks to improve performance, they suggest using a smaller number of disks with intra-disk parallelism. In a conventional disk drive, only a single I/O request can be serviced at a time. Although the arm and spindle assemblies are physically independent electro-mechanical systems, they are used in a tightly coupled manner due to the way that disk accesses are performed. Intra-disk parallelism would decouple the two electro-mechanical systems so that seek time and rotational latency would overlap.

Sankar et al. [23] present a sensitivity-based power optimization technique, that dynamically tunes disk settings, adapting them to the varying workload. Their approach systematically balances the dynamic "knobs" in the disk in order to meet certain storage-system performance constraints while maximizing energy savings. Their work distinguishes itself from others' by simultaneously playing with several dynamic knobs, rather than with just a single knob. One of the knobs that they tune is the seek head speed.

The results we present are based on measurements performed on actual physical disk drives and storage systems. This is in contrast to Sankar et al.'s experiments in [23], which were carried out on DiskSim [6], a simulator of storage systems.

Power budgeting for storage has received relatively little attention, as methods such as DRPM are theoretical at this point, and the alternative is spinning down disk drives, which is not practical for most applications. In contrast, there has been plenty of work on power capping and power management of servers. Regulating CPU frequency and voltage is a common way to manage the power of servers. Wu et al. [29] deal with power management of processors. They use dynamic voltage and frequency scaling to manage power-performance tradeoffs. While these methods are easily applied to servers, they cannot be applied to disk drives. Lefurgy et al. [17] describe a power capping technique for high-density servers. Their technique implements a feedback controller that uses system-level power measurements to periodically select the highest performance state, while keeping the system within a fixed power constraint. They use processor clock modulation as the actuator in their power controller. Minerick et al. [20] present a real-time feedback mechanism for maintaining a target average power for a laptop.

Zeng et al. [31] propose a server energy management strategy at the operating system level. They suggest managing energy as a first-class operating system resource that cuts across all existing system resources, such as CPU, disk, memory, and the network, in a unified manner. Lu et al. [19] introduce a power management method that classifies requests by their tasks, using the information in the operating system kernel. For each task, the system computes device utilization, and when it is found to be low, it is shut down.

Femal and Freeh [11] show a non-uniform, automatic power allocation framework for server clusters, based on forecasted workload. Their mechanism manages the power of the entire cluster by allocating power budgets for each server.

VII. Conclusions and Future Work

In this paper we have explored the effects of acoustic management on performance and power consumption. Acoustic management is applicable for both energy saving and power capping (or power budgeting). Quiet acoustic modes change the way the disk performs seek operations, so that power reduction, when no seeks are performed is not possible–when the disk is idle or during sequential access. Moreover, as only disk head movement is affected, the power for the electronics and platter spinning remains the same. As a result, power reduction is only moderate. For random-read workloads the power reduction was at most 23%, depending on the actual I/O workload.

Quiet mode incurs a response-time penalty. This is clearly shown in Figures 4f and 5f. This precludes using quiet mode for mission-critical applications that are sensitive to I/O response time. Single-threaded applications that require high throughput will suffer a 25% reduction in I/O throughput. Moreover, they will consume more energy in quiet mode than in normal mode, but will benefit from a lower peak power consumption. However, multi-threaded applications with a mixed workload of read and writes operations, both random and sequential, will be able to sustain the same I/O throughput, but with higher response time. Such applications may need to use a larger number of threads when using quiet mode, in order to sustain the same I/O throughput as in normal mode.

We have found that in some cases, seek operations consume more overall energy (despite consuming less instantaneous power) in quiet mode than in normal mode. We have also shown workloads for which quiet mode leads to energy savings. The SPC-1 workload tests clearly demonstrate that OLTP applications are good candidates for energy savings, in cases in which they can endure a degradation in response time.

Our work has potential to further impact storage controllers when the following topics that are beyond the scope of this paper are included: First, a temperature model can be developed to evaluate the impact of acoustic modes on the cooling requirements of a data center. Second, expanded benchmark testing and a more generic algorithm for selecting the proper acoustic mode could increase its impact.

Acknowledgments: We thank Julian Satran for numerous discussions and innovative suggestions. We thank Avishay Traeger for his help with running the benchmarks and for his many insights.

References

[1] "Storage performance council," http://www.storageperformance.org/.

[2] "Seagate's sound barrier technology (SBT)," 2000, http://www.seagate.com/docs/pdf/whitepaper/sound_barrier.pdf.

[3] "INCITS 361-2002 (1410D): AT attachment - 6 with packet interface (ATA/ATAPI - 6)," 2002.

[4] "LabVIEW release notes," 2009, http://www.ni.com/pdf/manuals/371778e.pdf.

[5] M. Allalouf, Y. Arbitman, M. Factor, R. Kat, K. Meth, and D. Naor, "Storage modeling for power estimation," in *Proceedings of SYSTOR 2009: The Israeli Experimental Systems Conference*, 2009.

[6] J. S. Bucy, J. Schindler, S. W. Schlosser, G. R. Ganger, and Contributors, "The DiskSim simulation environment - version 4.0 reference manual," May 2008.

[7] E. V. Carrera, E. Pinheiro, and R. Bianchini, "Conserving disk energy in network servers," in *Proceedings of the 17th Annual International Conference on Supercomputing*, June 2003, pp. 86–97.

[8] D. Colarelli and D. Grunwald, "Massive arrays of idle disks for storage archives," in *Proceedings of the 2002 ACM/IEEE conference on High Performance Networking and Computing*, November 2002, pp. 1–11.

[9] S. J. D., T. Eno, F. Christos, and G. G. R., "Using utility to provision storage systems," in *FAST '08: Proceedings of the 6nd USENIX Conference on File and Storage Technologies*. Berkeley, CA, USA: USENIX Association, 2008, pp. 313–328.

[10] F. Douglis, P. Krishnan, and B. Bershad, "Adaptive disk spin-down policies for mobile computers," in *Proceedings of the 2nd USENIX Symposium on Mobile and Location-Independent Computing*, April 1995, pp. 121–137.

[11] M. E. Femal and V. W. Freeh, "Boosting data center performance through non-uniform power allocation," in *ICAC '05: Proceedings of the Second International Conference on Automatic Computing*. Washington, DC, USA: IEEE Computer Society, 2005, pp. 250–261.

[12] S. Gurumurthi, A. Sivasubramaniam, M. Kandemir, and H. Franke, "DRPM: Dynamic speed control for power management in server class disks," in *Proceedings of the 30th Annual International Symposium on Computer Architecture*, June 2003, pp. 169–181.

[13] S. Gurumurthi, M. R. Stan, and S. Sankar, "Using intradisk parallelism to build energy efficient storage systems," *IEEE Micro Special Issue on Top Picks from the Computer Architecture Conferences of 2008*, pp. 50–61, 2009.

[14] S. Gurumurthi, J. Zhang, A. Sivasubramaniam, M. Kandemir, H. Franke, N. Vijaykrishnan, and M. J. Irwin, "Interplay of energy and performance for disk arrays running transaction processing workloads," in *Proceedings of the International Symposium on Performance Analysis of Systems and Software*, March 2003, pp. 123–132.

[15] H. Huang, W. Hung, and K. G. Shin, "FS2: dynamic data replication in free disk space for improving disk performance and energy consumption," in *Proceedings of the Symposium on Operating Systems Principles (SOSP)*, October 2005, pp. 263–276.

[16] Y. Kim, S. Gurumurthi, and A. Sivasubramaniam, "Understanding the performance-temperature interactions in disk i/o of server workloads." in *HPCA*. IEEE Computer Society, 2006, pp. 176–186, http://www.cs.virginia.edu/g̃urumurthi/papers/hpca06.pdf. [Online]. Available: http://dblp.uni-trier.de/db/conf/hpca/hpca2006.html#KimGS06

[17] C. Lefurgy, X. Wang, and M. Ware, "Power capping: a prelude to power shifting," *Cluster Computing*, vol. 11, no. 2, pp. 183–195, 2008.

[18] X. Li, Z. Li, F. David, P. Zhou, Y. Zhou, S. Adve, and S. Kumar, "Performance directed energy management for main memory and disks," in *Proceedings of the International Conference on Architectural Support for Programming Languages and Operating Systems*. ACM Press, 2004, pp. 271–283.

[19] Y.-H. Lu, L. Benini, and G. D. Micheli, "Operating-system directed power reduction," in *Proceedings of the 2000 International Symposium on Low Power Electronics and Design, 2000*, 2000, pp. 37–42.

[20] R. J. Minerick, V. W. Freeh, and P. M. Kogge, "Dynamic power management using feedback," in *COLP '02: Proceedings of Workshop on Compilers and Operating Systems for Low Power*, Charlottesville, Va, September 2002.

[21] E. Otoo, A. Pinar, D. Rotem, and S. C. Tsao, "A file allocation strategy for energy-efficient disk storage systems," *Lawrence Berkeley National Laboratory. LBNL Paper LBNL-637E*, July 2008.

[22] E. Pinheiro and R. Bianchini, "Energy conservation techniques for disk array-based servers," in *Proceedings of the 18th Annual International Conference on Supercomputing*, June 2004, pp. 68–78.

[23] S. Sankar, S. Gurumurthi, and M. R. Stan, "Sensitivity based power management of enterprise storage systems," in *Proceedings of International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, Baltimore, MD*, September 2008, pp. 865–870, http://www.cs.virginia.edu/ ss2wn/papers/mascots08.pdf.

[24] M. W. Storer, K. M. Greenan, E. L. Miller, and K. Voruganti, "Pergamum: replacing tape with energy efficient, reliable, disk-based archival storage," in *FAST'08: Proceedings of the 6th USENIX Conference on File and Storage Technologies*. Berkeley, CA, USA: USENIX Association, 2008, pp. 1–16.

[25] A. Traeger, N. Joukov, C. P. Wright, and E. Zadok, "A nine year study of file system and storage benchmarking," *ACM Transactions on Storage (TOS)*, vol. 4, no. 2, pp. 25–80, May 2008.

[26] H. Vandenbergh, "Vdbench 5.00 users guide," 2008, http://garr.dl.sourceforge.net/project/vdbench/vdbench/Vdbenchch.pdf.

[27] R. Villars, "Three keys for storage success: Content, architecture and getting personal. IDC," July 2007.

[28] S. W. Worth, "SNIA green storage tutorial," 2007, http://www.snia.org/forums/green/programs/SWWorth_Green_Storage.pdf.

[29] Q. Wu, P. Juang, M. Martonosi, L.-S. Peh, and D. W. Clark, "Formal control techniques for power-performance management," *IEEE Micro*, vol. 25, no. 5, pp. 52–62, 2005.

[30] J. Zedlewski, S. Sobti, N. Garg, F. Zheng, A. Krishnamurthy, and R. Y. Wang, "Modeling hard-disk power consumption," in *FAST '03: Proceedings of the 2nd USENIX Conference on File and Storage Technologies*. Berkeley, CA, USA: USENIX Association, 2003, pp. 217–230.

[31] H. Zeng, C. S. Ellis, A. R. Lebeck, and A. Vahdat, "ECOSystem: managing energy as a first class operating system resource," *SIGPLAN Notices*, vol. 37, no. 10, pp. 123–132, 2002.

[32] Y. Zhang, S. Gurumurthi, and M. R. Stan, "SODA: Sensitivity based optimization of disk architecture," in *Proceedings of the 44th annual Design Automation Conference, San Diego, CA*, June 2007, pp. 865–870, http://www.cs.virginia.edu/ gurumurthi/papers/dac07.pdf.

[33] Q. Zhu, Z. Chen, L. Tan, Y. Zhou, K. Keeton, and J. Wilkes, "Hibernator: helping disk arrays sleep through the winter," in *Proceedings of the Symposium on Operating Systems Principles (SOSP)*, October 2005, pp. 177–190.