SODA: Sensitivity Based Optimization of Disk Architecture

Yan Zhang
Qualcomm
San Diego, CA
zhangy@qualcomm.com

Sudhanva Gurumurthi
CS Dept., University of Virginia
Charlottesville, VA
gurumurthi@cs.virginia.edu

Mircea R Stan
ECE Dept., University of Virginia
Charlottesville, VA
mircea@virginia.edu

ABSTRACT

Storage plays a pivotal role in the performance of many applications. Optimizing disk architectures is a design-time as well as a run-time issue and requires balancing between performance, power and capacity. The design space is large and there are many “knobs” that can be used to optimize disk drive behavior. Here we present a sensitivity-based optimization for disk architectures (SODA) which leverages results from digital circuit design. Using detailed models of the electro-mechanical behavior of disk drives and a suite of realistic workloads, we show how SODA can aid in design and runtime optimization.

Categories and Subject Descriptors

B.4.2 [**Input/Output Devices**]

General Terms

Algorithms, Performance, Design.

Keywords

Disk drives, storage, power, performance, optimization.

1. INTRODUCTION

We are in the era of data-driven computing. Many applications deal with large datasets that need to be processed with low turnaround time. Several enterprise class applications, such as On-Line Transaction Processing (OLTP), On-Line Analytical Processing (OLAP), and web-services, are I/O intensive, thus require a high performance storage system. Optimizing disk drives involves trading off capacity, performance (in particular, the data rate), and power [Gur05]. The capacity is increased by using larger platters or more of them; but the larger platters increase the viscous heating (i.e., air friction due to the rotating platters) and adding more platters causes the power dissipation to increase [Sat90]. The data rate is increased by improvements in the linear density (which had been growing at the rate of 30% per year [HTech]) and the rotational speed of the platters (in Rotations per Minute, RPM). However, since the viscous dissipation has a cubic relation to RPM, increasing the rotational speed causes more heat to be generated. Improvements within this power-constrained design space are obtained through a combination of increase in the magnetic recording density and structural changes to the drive.

Disk drive behavioral optimization is a combination of a design-time effort, coupled with run time adaptation. For example, a data center can have specific energy constraints based on the electricity supply and capabilities of the cooling system in the building. Since disks are used in large numbers in server systems (e.g. RAID arrays), they are a significant source of power consumption and stress the cooling system. The applications that run on these systems also have a variety of characteristics and requirements. OLTP applications (eg. TPC-C [TPCC]) tend to transfer small chunks of data and do random I/O, thus the disk seeks need to be optimized in that case. On the other hand, for a video server, the disk needs to be optimized for a constant data transfer rate that is good enough to provide the desired playback speed.

It is important to understand the key *figures of merit* (i.e., the objectives and constraints in the optimization) and the *knobs* (the variables) that can be used in disk drive optimization. The figures of merit include performance (both throughput and latency), power, form factor, capacity, cost etc. Of the available knobs, some are usable at design time (static knobs) and others can potentially be varied at runtime (dynamic knobs). Static knobs include the number of platters and their size, the characteristics of the spindle motor (SPM) and the voice coil motor (VCM). Dynamic knobs include the voltages for the SPM and VCM, which can be used to trade off performance and power by slowing down or speeding up the platter rotation and the seek time.

In this paper we present the Sensitivity-based Optimization of Disk Architecture (SODA) framework. Compared to other optimization methods, sensitivity-based optimization (originally proposed for energy-delay optimization in circuit design [Hor02, Zyu03]), recognizes that the optimal trade-off between power and performance is not uniquely defined, but rather depends on the actual level of desired performance or acceptable power consumption. The sensitivity analysis approach provides a way to identify these optimal points by calculating the *ratio* of energy to delay *sensitivities* (partial derivatives) with respect to each knob that is used for the optimization, and making sure that *all* of those sensitivity ratios are equal [Hor02, Mar04, Zyu03]. Indeed, there are inherent similarities between circuit design optimization and those for disk drives. For example, the energy delay product that is used in circuit optimization is similar to energy (1/throughput) product for disk drives. There are also several other similarities such as between the *charging of a capacitor* and the *spinning up of a motor*, the electric charge energy stored on a capacitor and the magnetic field energy stored on a spinning motor, the leakage current in CMOS circuits and the DC motor current losses, Dynamic Voltage Scaling (DVS) and Dynamic RPM (DRPM) [Gur03], etc. Having such mappings facilitates modeling and understanding, as well as optimization, in one field (disk drives in this case) by reusing and applying the large body of knowledge developed in another (e.g. circuit design).

DAC 2007, June 4–8, 2007, San Diego, California, USA

2. RELATED WORK

Optimizing disk drives has been widely studied from both the performance and power viewpoints. Disk drive level optimizations include disk arm scheduling [Wor94] and data layout optimizations [Rue91,Hsu05] to improve seek behavior, techniques to boost bandwidth [Pat88], and optimize disk caches [Hu96]. The power optimizations for laptop/desktop systems include spin-down based schemes [Dou95] to exploit idleness, and techniques to increase idleness via prefetching and caching [Pap04]. In the context of servers, multi-speed disk drives (called Dynamic RPM or DRPM) have been proposed [Gur03, Car03]. The Physical Effect Modeling approach [Vos00] captures the physical phenomena that occur within and between electro-mechanical devices and has been used to model the SPM and VCM [Dam01]. The SODA methodology that we present in this paper provides a framework to craft policies to achieve specific energy-performance tradeoffs, using these previously proposed power management techniques as the underlying control mechanisms, and the disk drive models that we use are similar to the physical effect approach. Performance modeling of disk drives has been studied extensively [Rue94] and has resulted in detailed I/O simulation tools, such as Disksim [DS], which we also use in this paper. There have also been studies on modeling the power consumption [Zed03] and temperature [Kim06] of storage systems. Sensitivity based optimization techniques have been proposed for energy delay optimizations in circuits [Hor02, Mar04, Zyu03]. Constraint-based optimization has been used in the past for generating schedules for dynamic voltage scaling in real-time systems [Zha05] to reduce energy consumption.

3. OVERVIEW

3.1 Hard disk drives

A Hard Disk Drive (HDD) is an electro-mechanical magnetic storage device, whose activities are controlled and coordinated by digital controllers and buffers. The three main power dissipaters in a HDD are the *spindle motor* (SPM), which is used to rotate the platters, the *voice-coil motor* (VCM) that moves the disk arms, and the on-board *electronics*. When the disk is spinning and not servicing any requests, it is said to be in an *idle* power mode, and most of the power consumed is by the SPM. When a transfer request comes, a physical seek is needed, the VCM has to be activated, and the disk transitions into the *seek* power mode. The actual transfer of bits between the magnetic media and the electronic buffers in the drive takes place when the drive is in the *active* mode, where the read/write channel (also called the data channel) is enabled and leads to additional power consumption.

Designing disk drives involves tradeoffs between capacity, performance, and power. The *number* of platters and their *size* determines the disk capacity, but also affects the generated heat due to viscous dissipation by a *linear* factor, and by nearly the *fifth* power, respectively [Sat90]. The data rate is increased by improvements in the linear density (expressed in Bits per Inch or BPI) and/or increases in the RPM. The latter causes the generated heat to increase by nearly a *cubic* factor. In order to ensure reliability, one of the requirements in disk drive design is to always keep the operating temperature below a particular threshold, known as the thermal envelope [Her97]. Given the high costs associated with cooling modern electronic systems [Vis00], it is important that disk drives do not further increase this burden.

3.2 Sensitivity based optimization

In order to introduce sensitivity based optimization for disk drives we briefly present the formalism behind the method of "true power optimization" [Hor02, Mar04]. This method is the culmination of a series of attempts in the low-power circuit design community to come up with an "ideal" figure of merit for power-aware design [Sta01, Zyu03]. The main result is that in reality there is no *single* optimal point in the design space, but rather an entire *series* of points that optimally trade-off power for performance. The way to identify these points is by calculating the energy to delay *ratio of sensitivities* with respect to each "design knob" that is available for optimization - then the points for which all those sensitivity ratios are equal are "optimal". There are two dimensions (figures of merit) in the design space, Energy and Delay, with the cost being the Energy that needs to be minimized for a given Delay constraint. For simplicity let's assume that there are only two knobs that can be used in the optimization, x and y (e.g. these can be the supply voltage for the spindle motor and for the voice coil motor). The optimization problem then is:

$$\min E(x, y) \quad \text{s.t.} \quad D(x, y) = D_0$$

The main result of the method of true power optimization states that the sensitivity ratios of change in energy (E) with respect to change in delay (D) for knob x has to be the same as for knob y:

$$\frac{\partial E}{\partial x}\Big/\frac{\partial D}{\partial x} = \frac{\partial E}{\partial y}\Big/\frac{\partial D}{\partial y}$$

Although the discussion above concerns energy and delay, similar results are obtained for any other pair of dimensions in the design space; the choice of knobs is also arbitrary, and the result extends to any larger number of knobs. Also interesting, the same result is obtained if the roles of energy and delay are reversed (i.e. minimize delay under a constant energy constraint). This suggests that some of the distinctions made in the community between the areas of *low-power* and *power-aware* design may be unnecessary.

4. MODELING A HARD DISK DRIVE

Hard disk systems can be divided into electromechanical parts, including spindle motor (SPM), voice coil motor (VCM); and electrical parts, including data channel (I/Os), controllers, digital-to-analog converter, microprocessor, and RAM. The spindle motor (SPM) and voice coil motor (VCM) are DC motors in which the back-emf voltage (V_b) is proportional to the angular velocity (ω). Thus the voltage (V_a) applied to the motor is:

$$V_a = I_a \cdot R_a + V_b = I_a \cdot R_a + k_g \cdot \omega$$

where I_a is the armature current, R_a is the winding resistance of the armature and k_g is the motor voltage constant. The output torque T is proportional to the current; this provides a connection between the mechanical response and the electrical behavior. The torque is used to overcome the inertia and frictional drag:

$$T = k_t \cdot I_a = J \cdot \frac{d\omega}{dt} + b \cdot \omega^{\alpha}$$

where J is the inertia of the rotating parts, b is the viscous coefficient, and α is a coefficient that depends on the velocity. In this paper we consider $\alpha = 2$ for the SPM (high speed, turbulent flow [Sat90]) and $\alpha = 1$ for the VCM (low speed, viscous flow).

Combining the above and solving we can derive the steady angular velocity of the SPM:

$$\omega_{SPM} = \frac{\sqrt{\left(\frac{k_t k_g}{R_a}\right)^2 + \frac{4 k_t V_a b}{R_a}} - \frac{k_t k_g}{R_a}}{2b}$$

We can also get the mechanical response of the VCM:

$$\omega(t) = \omega_{VCM} \cdot (1 - e^{-\frac{t}{\tau}})$$

where τ is the time constant and ω_{VCM} is the maximum VCM angular velocity:

$$\tau = \frac{J}{b + \frac{k_g k_t}{R_a}} \quad , \quad \omega_{VCM} = K \cdot V_a \quad , \quad K = \frac{1}{k_g + \frac{b R_a}{k_t}}$$

We first derive a model for the disk *average behavior*. For the SPM power, assuming that the disk is always rotating at constant speed, the steady state power used to overcome the friction and windage losses can be expressed as [Sat90]:

$$P_{SPM} = n \cdot b \cdot \omega_{SPM}^{2.8} \quad , \quad b = \frac{\pi}{2} \cdot \rho \cdot C_d \cdot r^{4.6}$$

where n is number of platters, b_{SPM} is the viscous friction coefficient, ρ is density of air, Cd is the drag coefficient (0.005 for a flat platter [Hoe65]) and r is the platter radius.

For the VCM we use the seek operation model described in [Kim06]. A seek operation normally involves an *acceleration* phase, followed by a *coasting* phase of constant velocity, and then by a *deceleration* phase. The average distance *Dseek* is approximated as seeking across 1/3 of the data zone. For an average seek the VCM is accelerated from 0 to the maximum velocity *Vmax* and then decelerated to 0 with no coasting time. If the seek distance is less than the average seek distance *Dseek*, the VCM velocity will not reach the maximum velocity *Vmax*; if the seek distance is larger than *Dseek*, after the acceleration phase there will be a coasting phase before deceleration. The power for the VCM is given by $T \cdot \omega$, where T is the torque and ω is the angular velocity; the energy for an average seek is then:

$$E_{VCM} = \frac{n \cdot J \cdot \omega_{VCM}^2}{2} + \frac{n \cdot b \cdot \omega_{VCM}}{3}$$

where n is number of platters, ω_{VCM} is the maximum angular velocity, J is the inertia of the arm actuator (proportional to r^3) and b is the friction coefficient of the arm actuator (proportional to r^2). We assume that the arm length of the VCM is $2r$.

Combining the above, we get the total average energy for the disk running for a time t_0 (with one seek occuring during that time):

$$E = P_{SPM} \cdot t_0 + E_{VCM} + P_c \cdot t_0$$

where Pc is the power consumption for the electronic part of the disk system (estimated at 40% of the total idle power [Sri95]).

In order to perform the optimization, we also need a model for the *performance* of the disk drive. We consider the average performance as given by the *throughput* of the disk drive:

$$TP = \frac{B}{t_{rot} + t_{seek} + t_{trans}}$$

where B is the average number of bits per transfer, t_{rot} is the rotational latency π/ω_{SPM}, and t_{trans} is the average time to transfer:

$$t_{trans} = \frac{B}{BPI \cdot \omega_{spm} \cdot \frac{3}{4} r}$$

where BPI is the data density in bits per inch and r is the radius of the platter (here we assume that the average seek occurs in the middle track of the data zone i.e. at 3/4 of the platter radius).

5. EXPERIMENTAL SETUP

For the experimental results we use a set of commercial workload traces. The details of these workloads and the configuration of the storage system on which they were collected are similar to [Kim06]. The Openmail trace was obtained from [OPM] and the OLTP and Search-Engine were downloaded from the University of Massachusetts Trace Repository [UMT]. The TPC-C and TPC-H traces were collected on a 2-way and 8-way server systems, respectively, with the IBM DB2 database running on Linux.

5.1 Adapting the model

The disk model described in Section 4 was based on the average case; in order to get results specific for each workload, the input parameters of the model need to be scaled based on the workload characteristics. For example, in section 4 we assumed that all seeks show average-case behavior and that the average angular seek distance is 1/12 (assuming the disk arm at $2r$). We also assumed that the rotational latency is π/ω_{SPM} (i.e. the spindle motor always rotates half circle in order to reach the desired transfer position); and we assumed that, on average, data transfers occur around the middle track of the data zone, or at 3/4 of platter radius. However, for various workloads, physical seeks will vary from single cylinder seeks to full-stroke seeks; also the rotational latency and the data transfer cylinder will change depending upon how data is laid out on disk. In order to perform workload-specific optimization, we use data from the workload profiles in order to modify the model to a workload-specific case.

The parameters that are used for adapting the hard disk model are obtained by running the workload traces on the Disksim storage system simulator [DS], which models the performance aspects of disk drives, caches, and interconnects in a fairly detailed manner. The traces also provide seek time, data transfer time, idle time, rotational latency, number of transfer blocks, number of total physical seeks, number of zero distance seeks, single cylinder seek time, average seek time and full strobe seek time.

The workload specific spindle power is calculated similar to section 4 but ω_{SPM} is replaced with the actual speed ω_{SPM}. The energy for the VCM is modified to be calculated as:

$$E_{VCM} = p_{single} \cdot E_{VCM\_single} + p_{avg} \cdot E_{VCM\_avg}$$

where p_{single} and p_{avg} are the percentage of single track and average seeks in the workload (here the percentage of full stroke seeks is considered zero). The workload-specific seek time, rotational latency, transfer time and throughout are derived as:

$$t_{actual\_seek} = p_{single} \cdot t_{actual\_single} + p_{avg} \cdot t_{actual\_avg}$$

$$t_{actualL\_rot} = \frac{rotational\_angle}{\omega_{SPM\_actual}}$$

$$t_{actual\_transfer} = \frac{B}{BPI_{actual} \cdot r_{actual} \cdot \omega_{SPM\_actual} \cdot seek\_pos}$$

$$TP = \frac{B}{t_{actual\_rot} + t_{actual\_seek} + t_{actual\_transfer}}$$

6. RESULTS

6.1 Design Space Exploration

The SODA methodology consists of the combination of the HDD physical model and the sensitivity-based optimization. In the first experiment using SODA we investigate the effect of varying two dynamic knobs, namely, the voltages of the SPM and VCM, for all combinations of three different platter sizes: 1.8", 2.6", 3.3"; and 1, 2, and 3 platters/disk; effectively resulting in 9 distinct disk drive organizations with capacities ranging from 76 GB to 769 GB. For each configuration we apply the methodology to find the optimal trade-off between energy and performance. We conducted the experiment for all five workloads, but, due to space limitations, we show results only for Openmail and Financial; the results for the other workloads are similar. Each curve in Figure 1 corresponds to one of the nine disk drive configurations and represents the Pareto-optimal (i.e. equal sensitivity ratios, except when a knob has reached its maximum).

This design space exploration study using SODA provides several insights into both the design as well as the dynamic behavior of disk drives. First, as expected, the curves corresponding to the highest and lowest capacity disk drives are the farthest and closest to the origin, respectively. However, between these two extremes, the design space shows some interesting trade-offs. For instance, a 2.6" disk drive with 3 platters provides roughly the same energy-performance trade-offs as a 3.3" disk with a single platter, but the capacity of the former is significantly higher (477 GB) than the latter (256 GB). Therefore, given a power and performance target, the designer can make use of such information to target the same disk drive to multiple segments of the market.

Another way to utilize SODA is to choose the most desirable configuration (from a power-performance point of view) that satisfies a given capacity target. For example, we can observe two configurations that provide approximately the same capacity (a 1.8" disk drive with 2 platters and a 2.6" drive with 1 platter). However, the curve for the 1.8" disk drive is closer to the origin, thereby being the more desirable design choice.

A key observation from Figure 1 is that the disks with the larger platters can operate over a larger dynamic range in the E vs. 1/TP design space. Each point in this dynamic range corresponds to a particular setting of the SPM and VCM voltages. This suggests that for a disk drive that can work over a range of voltages (thus speeds) it may be better to choose drives that use larger, rather than smaller, platters. There has been research in recent years on designing such multi-speed disk drives [Gur03, Car03]; SODA can provide guidance for optimizing such a drive at an early stage of the design process and in an application-aware manner.

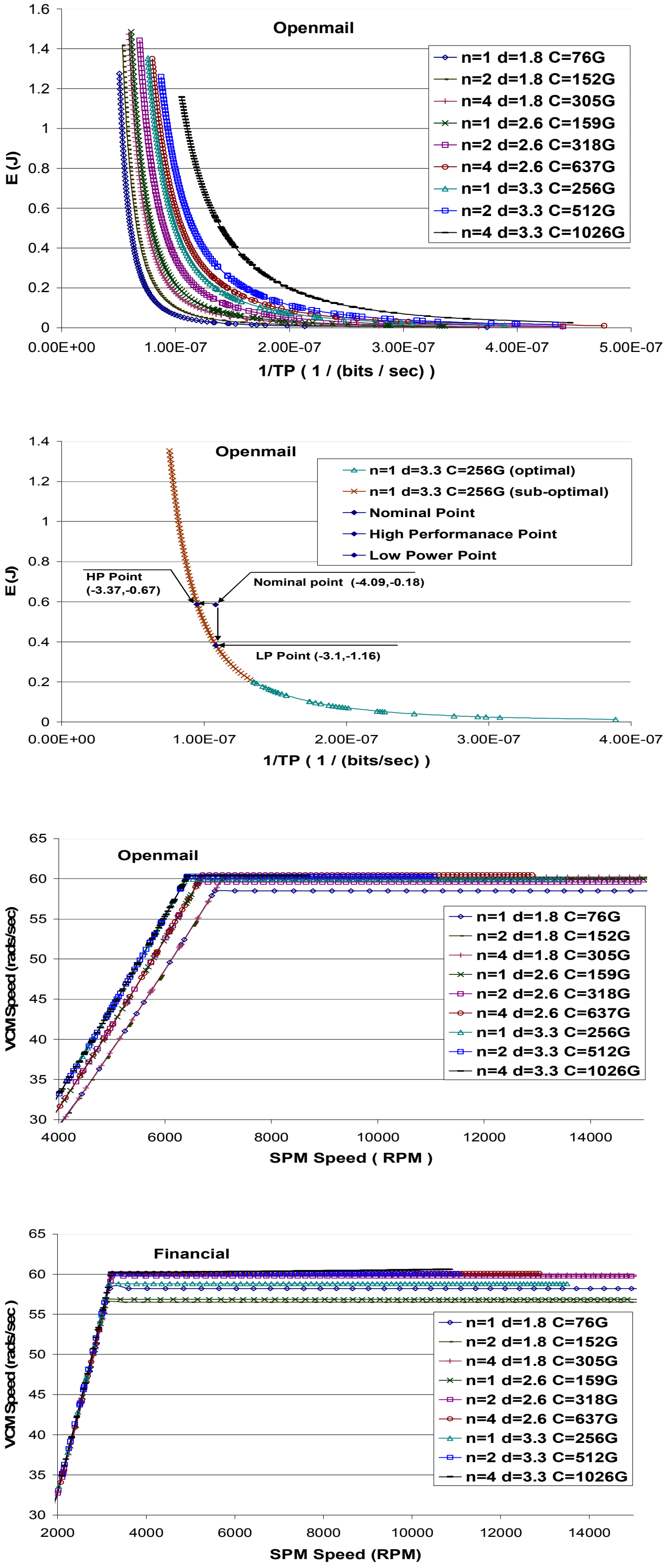

Figure 1 – Design space exploration - Energy vs. 1/ Throughput (top) SPM Speed vs. VCM Speed (bottom)

Also from Figure 1 we can see how SODA can be used to optimize for either high-performance or for low power. In the second graph it can be seen that the original nominal point for the design (the one used in the original workload) is not on the optimal curve (it has non-equal sensitivity ratios). This means that one can get the same performance but lower power by projecting on the *x* axis (LP point in the graph), or get higher performance with a lower power consumption by projecting on the *y* axis (HP point in the graph). Incidentally, for this case even the Pareto-optimal curve is suboptimal since the VCM has already reached its maximum speed as can be seen from the bottom graphs.

Interestingly, from the bottom graphs in Figure 1 we observe significantly different trends (for optimal SPM and VCM) for Openmail and Financial. This is because the Financial workload has significantly longer idle time compared to Openmail (111.45 ms vs. 12.5 ms); thus the energy/request in Financial tends to be much higher than for Openmail. The bulk of this energy is consumed by the SPM which reduces the range of values that can be used optimally. On the other hand, since the VCM consumes far less power than the SPM, the performance target can be recovered by performing large modulations of the VCM speed, as shown in the y-axis of the graphs at the bottom. However this modulation reaches saturation due to physical limits on the VCM voltage, and is not enough to fully compensate for the restricted RPM range; therefore the optimal range of values for performance (the *x*-axis of the graphs) tends towards a higher latency for Financial. A higher RPM could be used in order to attain a higher performance, but this would break the optimality in the design.

6.2 The Impact of Seek Time

In this section we use SODA to study the performance-power tradeoffs for seek operations. From a performance viewpoint, seeks impede the flow to and from the platters, thereby diminishing the effective data rate of the disk. Disk seek operations also exercise the VCM and therefore dissipate power. In order to isolate the impact of seeks we use SODA with the same workloads as in the previous section, but with seek times as input, all the way from single-cylinder seeks to full stroke seeks. The results of this analysis are given in Figure 2. Each curve in this graph corresponds to a particular value of the seek time.

As Figure 2 indicates, since shorter seek times benefit performance and consume less power in the VCM, their sensitivity curves are closer to the origin. When we look at the speed characteristics for the Openmail and Financial workloads, we observe that as the seek time increases from 0.5 ms (single cylinder seek), the optimal curves for higher seek times result in a higher RPM range. This is because longer duration seeks hurt performance - to compensate for this, the optimization algorithm increases the disk RPM, which improves the rotational latency and the transfer time, and hence the curves shift to the right. This trend continues till the seek time reaches 4.48ms, which is the average seek time for the 3.3”-platter disk drives. At that point, the disk arm has reached its terminal velocity. Any further increases in the seek time will induce coasting of the head - this coast time causes extra power to be consumed by the SPM. In order to optimize for energy during these coasting periods, the optimization algorithm needs to scale back the RPM and hence the curves for seek times greater than 4.48ms shift leftward.

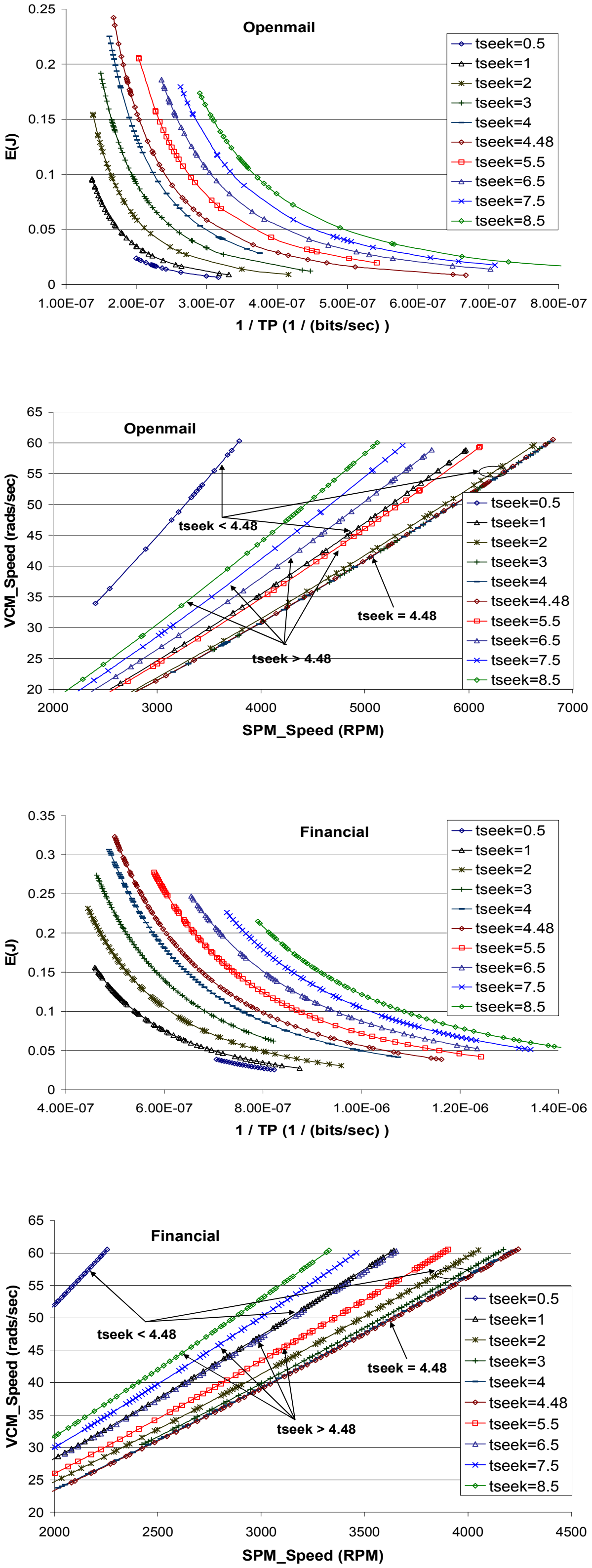

Figure 2 – Impact of Seek Time on energy and performance

This result shows another interesting similarity between disk and circuit power optimization. Circuit static power is consumed irrespective of any switching activity (mainly leakage), while circuit dynamic power is due to switching activity and is therefore a function of the usage of the circuit by some workload. Similar to the circuit static and dynamic power are the SPM and VCM power respectively. The SPM is always operational and draws power irrespective of whether the disk is in idle, seek, or active modes. The VCM is active only when disk seeks are needed, which is workload dependent. Like in modern CMOS circuits, the power consumed by disk drives is also dominated by the static part (SPM). Also similar to circuits, just as dynamic voltage scaling can be used to reduce leakage power in circuits, lowering the SPM voltage (which reduces its RPM) can mitigate its power.

In summary, minimizing seek times is important, both from the performance and power viewpoints. The results indicate that there is room for designing powerful VCMs to improve performance, since the VCM power is significantly lower than that of the SPM.

7. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented a detailed and parameterized model for disk drives for two key figures of merit, namely, performance and energy consumption. We have shown two scenarios where this model can be used for sensitivity-based optimization of disk drives, one at design-time and the other at run-time. A key advantage of the SODA framework is the ability to rapidly explore large design spaces efficiently, which can be especially useful during the early stages of development of a new architecture. In the future we will integrate the SODA model with a detailed storage architecture simulator, such as Disksim and then explore new storage architectures that can deliver performance for data-intensive applications in an energy-efficient manner.

8. ACKNOWLEDGMENTS

This work was supported in part by NSF grants CCR-0105626, CCR-0133634, CNS-0627527, CNS-0551630, and a grant from MARCO Interconnect Focus Center. We would like to thank Shahrukh Tarapore for his valuable feedback on this paper.

REFERENCES

[Car03] E.V. Carrera et al. Conserving Disk Energy in Network Servers. Int. Conference on Supercomputing (ICS), Jun 2003

[Dam01] D. Dammers et al. Motor Modeling Based on Physical Effect Models. IEEE/ACM International Workshop on Behavioral Modeling and Simulation (BMAS), Oct 2001

[Dou95] F. Douglis, P. Krishnan. Adaptive Disk Spin-Down Policies for Mobile Computers. Computing Systems, 8(4), pp 381–413, 1995

[DS] G.R. Ganger et al. The DiskSim Simulation Environment Version 2.0 Reference Manual, Dec 1999

[Gur03] S. Gurumurthi et al. DRPM: Dynamic Speed Control for Power Management in Server Class Disks. ISCA, Jun 2003

[Gur05] S. Gurumurthi et al. Disk Drive Roadmap from the Thermal Perspective: A Case for Dynamic Thermal Management. ISCA, pp 38–49, Jun 2005

[Hor02] M. A. Horowitz et al. Methods for true power minimization. ICCAD pp 35-42, 2002

[Her97] G. Herbst. IBM's Drive Temperature Indicator Processor (Drive-TIP) Helps Ensure High Drive Reliability. IBM Whitepaper, Oct 1997

[HTech] Hitachi Global Storage Technologies - HDD Technology Overview Charts.http://www.hitachigst.com/hdd/

[Hsu05] W.W. Hsu et al. The Automatic Improvement of Locality in Storage Systems. ACM Transactions on Computer Systems, 23(4), pp 424-473, Nov 2005

[Hu96] Y. Hu, Q. Yang. DCD Disk Caching Disk: A New Approach for Boosting I/O Performance. ISCA, pp 169–178, May 1996

[Kim06] Y.Kim, S. Gurumurthi, A. Sivasubramaniam. Understanding the Performance-Temperature Interactions in Disk I/O of ServerWorkloads. HPCA, pp 179–189, Feb 2006

[OPM] The Openmail Trace. http://tesla.hpl.hp.com/private software/.

[Pat88] D. Patterson, G. Gibson, R. Katz. A Case for Redundant Arrays of Inexpensive Disks (RAID). ACM SIGMOD Conference on the Management of Data, pp 109–116, Jun 1988

[Pap04] A. Papathanasiou, M. Scott. Energy Efficient Prefetching and Caching. USENIX Annual Technical Conf, Jun 2004

[Rue91] C. Ruemmler, J. Wilkes. Disk Shuffling. Technical Report HPL-91-156, HP Laboratories, Oct 1991

[Rue94] C. Ruemmler, J. Wilkes, An introduction to disk drive modeling, IEEE Computer, 27(3), pp 17-28, Mar 1994

[Sat90] I.Sato et al.. Characteristics of Heat Transfer in Small Disk Enclosures at High Rotation Speeds. IEEE Transactions on Components, Packaging, and Manufacturing Technology, 13(4), pp 1006–1011, Dec 1990

[Sri95] M. Sri-Jayantha. Trends in Mobile Storage Design. International Symposium on Low Power Electronics, Oct 1995

[Sta01] M. R. Stan, Low Power CMOS with Sub-Volt Supply Voltages, IEEE Transactions on VLSI Systems, Apr. 2001

[TPCC] TPC-C Benchmark V5. http://www.tpc.org/tpcc/

[UMT] UMass Trace Repository. http://traces.cs.umass.edu

[Vis00] R. Viswanath, V.Wakharkar, A.Watwe, V. Lebonheur. Thermal Performance Challenges from Silicon to Systems. Intel Technology Journal, Q3 2000

[Vos00] L.M. Voßkamper et al. Combining Models of Physical Effects for Describing Complex Electromechanical Devices. IEEE/ACM International Workshop on Behavioral Modeling and Simulation (BMAS), Oct 2000

[Wor94] B. Worthington et al. On-Line Extraction of SCSI Disk Drive Parameters. ACM SIGMETRICS Conf. on Measurement and Modeling of Computer Sys, pp 146–156, May 1995

[Zed03] J. Zedlewski et al. Modeling Hard-Disk Power Consumption. In Conference on File and Storage Technology (FAST), Mar 2003

[Zha05] Y. Zhang et al. Optimal Procrastinating Voltage Scheduling for Hard Real-Time Systems. DAC, Jun 2005

[Zyu03] V. Zyuban, P. N. Strenski. Balancing hardware intensity in microprocessor pipelines. IBM Journal of Research and Development, Volume 47, Number 5/6, 2003