The Need for Temperature-Aware Storage Systems

Sudhanva Gurumurthi
Department of Computer Science
University of Virginia
Charlottesville, VA 22904, USA
Phone: (434) 982-2227
Fax: (434) 982-2214
e-mail: gurumurthi@cs.virginia.edu

ABSTRACT

Storage has become ubiquitous. Disk drives are commonplace in most laptops and desktops. In addition, they are used in large numbers in high-end server systems. Storage devices have also proliferated the consumer electronics market with their use in products like cameras, and portable music devices. This widespread usage of disks has been the result of tremendous growth in both the density and speed of these devices. Over the past two decades, we have been enjoying a 40 percent annual growth in the data rate of disks, due to innovations in the recording technology coupled with a scaling up of the drive RPM. Since raising the RPM increases the heat that is generated due to viscous dissipation by nearly a cubic factor, in order to design the drives for a constant thermal envelope, the platter sizes are also shrunk, as the latter has a fifth power impact on the temperature.

In this paper, it shall be shown that this thermal-constrained scaling is going to be very challenging to sustain even for very small platter sizes, causing a significant slowdown in the pace of performance growth in future drives. Using real workloads, the need for continued scaling of the data rate is motivated. Some simple techniques will be presented that can be employed to overcome these effects. Finally, it will also be shown that managing temperature in disks cannot necessarily be tackled merely via energy saving techniques.

Keywords: Storage System, Disk Drives, Power and Temperature Management.

NOMENCLATURE

BPI	Linear Density (Bits/Inch)
IDR	Internal Data Rate (MB/s)
RAID	Redundant Array of Independent Disks
RPM	Rotations Per Minute
TPI	Track Density (Tracks/Inch)
C_{max}	Raw Disk Capacity (Bits)
η	Stroke Efficiency
n_{surf}	Number of Disk Surfaces
r_i	Platter Inner Radius (Inches)
r_o	Platter Outer Radius (Inches)

INTRODUCTION

Storage plays an important role in computer systems. Many applications, such as transaction processing, web services, multimedia content dissemination, and scientific data processing, are data driven and their performance is critically dependent on the underlying storage system. Designing disk drives involves tradeoffs between capacity, speed, and power. Capacity of a disk drive can be increased by using larger platters and/or multiple of them. However the number of platters and their size affect the heat that is generated in the disk drive (due to viscous dissipation) by a linear factor and by the 4.6th power respectively [1]. The data rate of the disk drive can be increased by improvements in the linear density and/or increases in the RPM. The latter causes the viscous dissipation to increase by the 2.8th power.

In order to design disk drives for a constant thermal envelope, the conventional method has been to shrink the platter size, which reduces the viscous dissipation by nearly the fifth power, and exploit this slack to ramp up the RPM. This design methodology has enabled the data rate of disk drives to grow annually by 40 percent for nearly the past two decades. The primary motivation for designing disks to operate within the thermal envelope is reliability. High temperatures can cause a variety of reliability problems, from data corruption to a permanent failure of the device [2]. For example, a 15 °C rise in the ambient temperature can double the failure rate of a disk drive [3].

In this paper, it will be shown that this growth is going to be very challenging to sustain in future disk drives using the current thermal-constrained design methodology. The need for *average-case* design via *Dynamic Thermal Management (DTM)* techniques is motivated. Finally, it will be shown that temperature management of storage systems is *not* necessarily the same as energy management.

THERMAL ROADMAP FOR FUTURE DISK DRIVES

This section presents a "Roadmap" for future disk drives. The roadmap gives the projected data rate and capacity for disk drives based on expected trends in recording technology, the overheads associated with magnetic recording, and the relationships between platter size, RPM, and number of platters with respect to temperature. It is assumed that the disk drives have to be designed such that, even under worst-case duty cycle conditions (i.e., platters are spinning, and the arms are moving), the temperature does not exceed the thermal envelope.

Capacity and Performance Models

In order to derive the roadmap, the relationships between capacity, performance, and temperature need to be established. The capacity models are based on two fundamental properties of the underlying recording technology, namely, the linear density (BPI) and the track density (TPI). Based on the geometry, the raw capacity (in terms of bits) of a platter are calculated. This raw capacity, C_{max}, is given by the equation:

$$C_{max} = \eta \times n_{surf} \times \pi(r_o^2 - r_i^2)(BPI \times TPI)$$

r_o and r_i denote the outer and inner radii of the disk platter (the spindle motor is housed within the inner circular aperture of the platter). n_{surf} is the number of disk surfaces, which is twice the number of physical platters. η is called the *stroke efficiency* and is the fraction of the total platter surface that is user-accessible. This capacity is derated to account for overheads such as Zoned Bit Recording and Servo sectors. Zoned Bit Recording is a coarse-grained way of accommodating variable sized tracks, where the tracks are grouped into zones, with each track in a zone having the same number of sectors. Servo is used to guide the movement of the arms and are encoded as a Gray code over the total number of tracks on a surface. The Error Correcting Code (ECC) overheads, due to change in Signal-to-Noise ratios caused by Terabit and higher areal densities [4], are also accounted for. The two performance metrics modeled are: seek time and data rate. When we need to read or write a block of data in a disk drive, and the block is not on the current track, the disk arms need to be physically moved to the desired track. The time that is incurred for this operation is called the *seek time*. The seek time depends on the inertial power of the voice-coil motor (VCM) that is used to move the arms and the radial length of the data band on the platter [5]. Physically, a seek involves an acceleration phase when the VCM is turned on, followed by a coast-phase of constant velocity where the VCM is off, and then a deceleration to stop the arms near the desired track, during which the VCM is again active but the current is reversed to generate the braking effect. The *(Internal) Data Rate (IDR)* is the actual speed at which data can be read from or written to the the platters. The IDR is proportional to the product of the BPI, RPM, and platter diameter and is calculated from the previous modeling steps. The models were validated against 13 real disk drives from four different manufacturers and from four different calendar years and it was found that the capacity and performance models are within 12% and 15% respectively of the real products for a given input BPI, TPI, platter diameter, and RPM. Complete numerical data of this validation is given in [6].

Thermal Model

The thermal model is an extension of the work done by Eibeck et al [7]. This model evaluates the temperature distribution of the drive by calculating the amount of heat that is generated by components such as the spindle motor (SPM) and the VCM, the conduction of heat along the solid components and the convection of heat to the air. It is assumed that the drive is completely enclosed and the only interaction with the external air is by the conduction of heat through the base and cover and the convection to the outside air. The outside air is assumed to be maintained at a constant temperature. In order to calculate the convective heat transfer coefficient, the model makes use of empirical correlations for known geometries. The heat of the internal drive air is calculated as the sum of the heat energy convected to it by each of the solid components and the viscous dissipation in the air itself minus the heat that is lost through the cover to the outside. The heat equations for the different components are solved using the finite difference method.

In order to do the thermal simulation, the model requires a set of input parameters. The first set of parameters relate to the disk geometry, such as the inner and outer radii of a platter, the drive enclosure dimensions, and the length of the disk arms. In order to develop parameterized models for the geometry, actual disk drives were taken apart and their geometry was studied in detail. This allowed us to determine how the components are internally laid out and create geometry models parameterized for the platter size and their quantity. The physical drive parameters such as the length of the disk arm, thickness of the platter, base, and cover etc., which are not considered by the capacity and performance models, were measured using Vernier calipers. The next set of parameters relates to the materials that compose the drive. The platters on most disk drives are typically made of an Aluminum-Magnesium alloy and the base/cover casting are Aluminum. As the exact alloy that is used tends to be proprietary information, it was assumed that the platters, together with the disk arm and spindle hub, are made of Aluminum. The power output of the VCM, which is dependent on the size of the data zone, was calculated using the data given in [8]. The external ambient temperature was assumed to be 28 °C.

This model was validated for a Hitachi Deskstar 7K500 disk drive, which is equipped with thermal sensors that report the temperature of the disk cover. The disk drive was placed in an isothermal oven and was connected to a workstation. The oven temperature was set to 68 °C and the disk was allowed to spin in idle mode (i.e., the SPM is active but the VCM is not) for a period of 6 hours in order to allow the device to stabilize at its steady-state temperature. The steady-state temperature was found to be 81 °C. This experiment was replicated in the simulator and the reported steady-temperature was 83.4 °C. This small difference between the actual and simulated readings is due to the lack of power models for the electronic components of the disk drive in the simulator and possible imprecision in the thermal sensor.

The Roadmap

In order to setup the roadmap, a predictive technology model was used for the BPI and TPI. This was done by leveraging data that has been published about the past trends in the recording technology [9] and projecting it for an 11-year period starting from the year 2002. The growth of BPI is expected to slow down due to factors such as the difficulty in lowering the fly height of the head, problems in using a recording medium that is more coercive (since it is challenging to design the write head for such media), and the magnetic grain size constraints imposed by the superparamagnetic limit. The TPI growth is expected to slow down as well since narrower tracks are more susceptible to media noise, and more closely packed tracks lead to intertrack interference [10]. The ratio of the BPI to TPI (called the Bit-Aspect Ratio) was calculated by studying a set of proposals for Terabit areal density disk drives [4, 11, 12]. When generating the roadmap, the goal is to determine the RPM for a given platter size and also the BPI and TPI values for a given calendar year such that we sustain the annual IDR growth of 40 percent. The roadmap for a 1-platter disk drive is shown in Fig. 1.

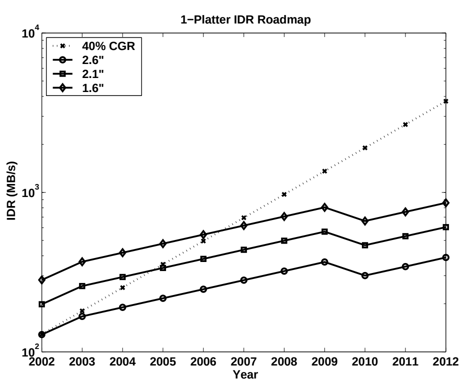


Figure 1: Disk Drive Roadmap. Each solid curve gives the maximum attainable IDR within the thermal envelope. The dotted line indicates the 40 percent target growth rate in IDR.

The graph shows the data rates for future disk drives with three different platter sizes housed within a 3.5-inch form-factor drive enclosure. The dotted straight line in the graph corresponds to the 40 percent IDR growth rate. A brief analysis of this roadmap shall now be presented. Consider the 2.6-inch platter size. Along the 40 percent growth rate curve, the IDR requirements increase nearly 29 times from the year 2002 to 2012. A portion of the required increase is provided by the growth in the linear density alone. Any demands beyond that has to be provided by an increase in the RPM. In order to determine the points on the roadmap where significant changes in the RPM are required, it is useful to sub-divide the timeline into three regions, namely, the years before 2004, where the BPI and TPI growth rates are 30 percent and 50 percent respectively [9], the years from 2004 to 2009, which are in the sub-terabit areal densities. In this region, the BPI and TPI growth rates slow down but ECC requirements are still moderate. In order to compensate for the slowdown in recording density, the RPM would need to be scaled up more aggressively so that we can meet the data rate requirements. The region from 2010 to 2012 corresponds to the Terabit areal density region where there is a steep growth in the ECC requirements, whereby requiring even more aggressive RPM scaling.

From the thermal viewpoint, it is found that the viscous dissipation increases from 0.91 W in 2002 to 1.13 W in 2003 for the 2.6-inch disk drive. In the second region of the roadmap, due to more aggressive scaling of the RPM (and its nearly cubic power impact), the viscous dissipation increases from 2 W in 2004 to over 35.55 W in 2009, causing a significant rise in temperature, well above the thermal envelope. Overall, we observe that this scaling of the RPM is not possible to sustain within the thermal envelope from the 2007 timeframe onwards and the feasible IDR falls off the 40 percent curve even for the 1.6-inch platter size. The viscous dissipation increases even further from the year 2010 onwards and reaches a value of 499.73 W in 2012, causing the internal air temperature to reach as high as 602.98 °C! In order to prevent these situations from happening the RPM would need to be capped, thereby the IDR would fall below the 40 percent curve. This is projected to happen in the 2007 timeframe (as shown in the figure).

One trivial way of preventing this fall-off is by incorporating progressively more powerful cooling systems. Indeed, it was found that for every 5 °C of extra reduction in the ambient air temperature, we can buy nearly one year along the roadmap in the short-term (although the terabit transition would require a more aggressive reduction in the temperature). However, increasing the cooling budget is not a scalable solution due to the high cost, both in terms of the fixed infrastructure (chilled water plants, air distribution system etc.) and also operational costs. Moreover, other components that are in close proximity to the disk drives (eg. processors and memory modules inside a rack-mount unit) can preheat the ambient air. Given that the focus on data center design today is to minimize the costs of cooling [13], it is important to design the storage system to be able to deliver the required performance/capacity in an energy efficient manner.

THE NEED FOR FASTER DISKS

In this section, the need for continued growth in disk drive data rates will be motivated and the thermal ramifications of this growth will be presented. The performance of various I/O intensive server applications were studied. The original storage system on which the traces were collected were re-created in the performance simulator (which is Disksim [14]), both in terms of the actual disk drives used and also higher level organizational parameters such as RAID. Disksim is widely used in storage systems research. The RPM of the disks were increased in increments of 5000 RPM (all other things remaining constant) and the impact on the response time was studied. The question that we seek to answer is whether using higher speed disk drives is able to provide significant performance benefit for real applications.

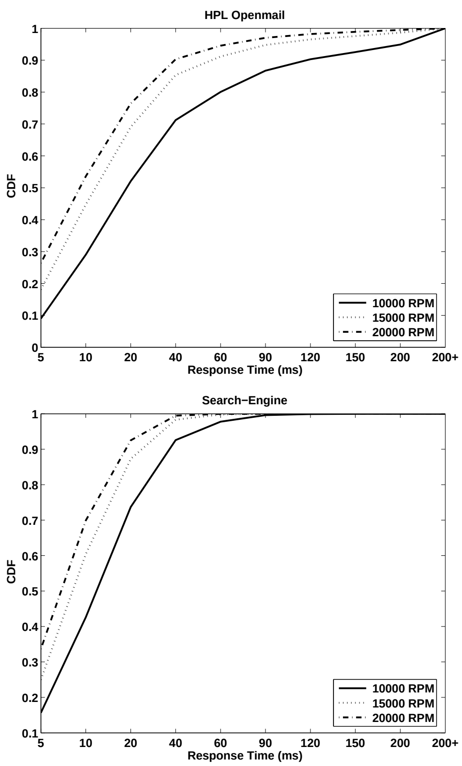


Figure 2: Disk Drive Roadmap. Each solid curve gives the maximum attainable IDR within the thermal envelope. The dotted line indicates the 40 percent target growth rate in IDR over time.

Although several applications were analyzed in this study, in the interest of clarity, the results are presented only for two, namely, *Openmail* and *Search-Engine*. Openmail is a trace of the I/O requests to a mail-server at the Atlanta Response Center. Search-Engine is a trace of the requests to an Internet search engine server. The results for these two applications are shown in Fig. 2. The data shows the Cumulative Density Function (CDF) of the response times. It is observed that a 10K RPM rise from the baselines provides significant performance benefit. The CDF curves shift to the left whereby indicating that most of the I/O requests benefited from the higher RPM. This motivates the need for continued growth in the data rate of disk drives.

In order to analyze the dynamic thermal behavior of the storage system under actual workload execution, we have built a detailed thermal-performance simulator [15]. The simulator consists of two components, namely, the performance model and the thermal model. The thermal state is maintained and updated as the performance simulation progresses. The state information is maintained separately for each simulated disk. The thermal state of a disk is updated on any event in the performance simulation that can affect the thermal behavior. For example, a seek operation requires an arm movement, which causes the VCM to expend power and therefore generate heat. Since the performance model (Disksim) is an event-driven simulator and the thermal model is based on explicit time-varying behavior, the performance model is made to sample the latter a high frequency, in order to get accuracy. Here, the performance model invokes the thermal simulator a certain number of times between any two I/O events and communicates its state. Since disk seeks have a significant impact on the disk drive temperature, this component is modeled accurately. For this, a Bang-Bang Triangular Seek [16] model is used, where both the acceleration and deceleration times are equal. Between these two phases, the disk may coast for a certain period of time depending upon the maximum velocity at which the arms are designed to travel and also the physical distance that needs to be traversed. For a given seek operation, the power that is expended is the highest when the distance to be traversed is such that the disk accelerates to the maximum velocity and immediately decelerates. Seeks of shorter or longer distance (where the arm coasts along the platter surface with the momentum that it has built up) generate less heat.

The thermal behavior of the two applications is shown in Fig. 3. The graphs show the temperature of one of the disks, namely, disk 0, during the simulation of the workloads. In each graph, we plot the temperature when the disks are operating in their baseline configurations, in addition to the higher speed ones shown in Fig. 2. In order to study the thermal behavior at the steady-state, the workload is executed after the disks are allowed to “warmup” for 150 minutes from their initial, off (cold) state. For these 150 minutes, it is assumed that the disks are rotating but there is no movement of the arm. Once the disks reach their steady-state idle temperature, the workload begins execution.

For both the applications, it can be observed that a 5000 RPM increase from the baseline can be accommodated within the thermal envelope without increasing the cooling requirements. This is due to the nature of the seek times, which are mostly small, thereby causing very little heat to be generated. However, when the speed is increased by another 5000 RPM, the disk temperature crosses above the thermal envelope. The disk that is used for Search-Engine application experiences a more significant excursion above than that of Openmail. This is because the storage system of Search-Engine uses 4-platter disks, which causes a greater amount of heat dissipation at the higher RPM than the 1-platter disks used in Openmail. Since the cooling budget

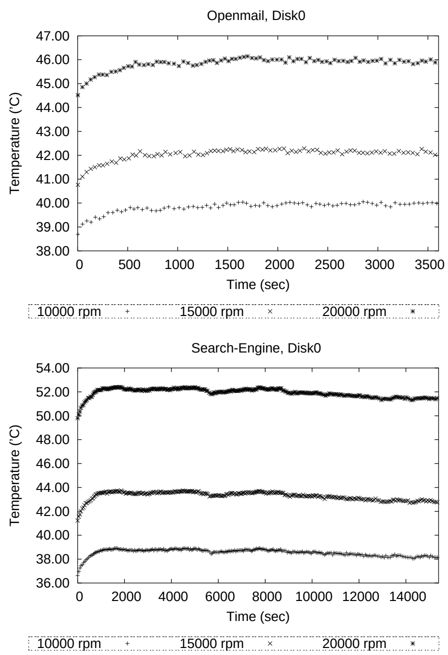


Figure 3: Thermal profiles of the applications.

has been established to meet the thermal envelope specifications at 10K RPM, the excess heat that is generated is not sufficiently extracted out by the cooling system, causing the disk drive temperature to rise.

TECHNIQUES FOR TEMPERATURE MANAGEMENT

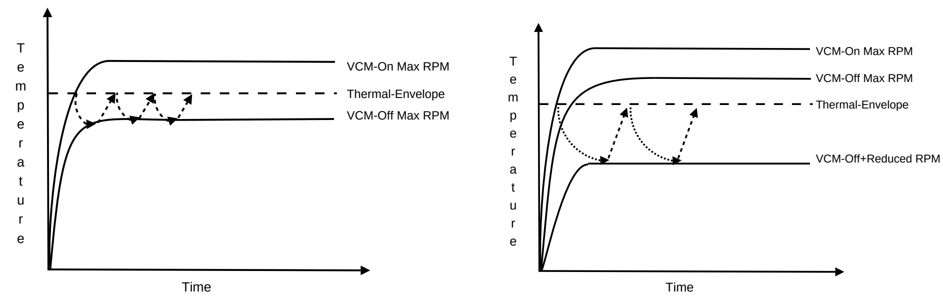
As the results indicate, it is going to be difficult to design disk drives that can deliver performance along the IDR curve that we have enjoyed for nearly the past two decades, and still be able to meet the thermal budget. Moreover, in enterprise class systems, there is no viable alternative to hard disk drives. Flash-based hybrid disk drives [17], which are starting to appear in the market, are not attractive for servers. The reliability of flash memory, which is primarily linked to the number of write/erase cycles, would cause the device to fail within a short span of time under the relatively heavy I/O loads that are common in servers. Moreover, the latency of the underlying hard disk drive is still an issue for read traffic. It has also been shown that applications benefit from continued growth in the data rate. Increasing the cooling budget to meet these requirements is not cost-effective, especially for the highly commoditized disk drive market. The implication of all these factors is that we need to move away from designing for the worst-case thermal envelope and focus more on *average-case* behavior. However, we need to prevent the temperature from exceeding the thermal envelope at run-time. This can be accomplished by dynamically modulating the activities on the disk drive via *Dynamic Thermal Management (DTM)*.

DTM Mechanisms and Policies

Two approaches to implement DTM are illustrated in Fig. 4. The technique shown in Fig. 4(a) is the simplest and requires very little modification to the existing disk drive designs. This graph shows the temperature response of a hypothetical disk drive. The dashed horizontal line is the thermal envelope. For this disk, in the absence of any seek activity (i.e., VCM is off), the temperature remains below the thermal envelope. This is shown by the lower of the two solid curves. However, if the VCM is active, the temperature could rise above the threshold. In order to prevent this situation from happening, we would need to invoke the DTM countermeasure. This DTM policy can work as follows. When the temperature of the disk drive is close to the thermal envelope, we stop issuing requests to it. This causes all seek activity to stop whereby allowing the disk to cool down. After the temperature has been sufficiently reduced, it is allowed to resume servicing requests again.

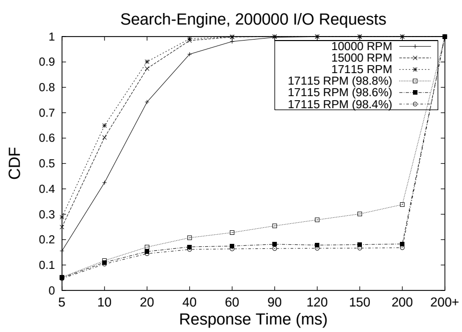
The results for this strategy are shown in Fig. 5 for the Search-Engine workload. In order to explain this experiment, two terms need to be defined, namely, *thermal emergency* and *thermal safety*. The thermal emergency is the temperature at which the DTM mechanism is invoked. After this point, no requests are served and they are queued up until the temperature falls to the *thermal safety* value, at which time normal operation resumes. In the graph, the results are shown for three values for the thermal safety, namely, 98.8, 98.6, and 98.4 percent respectively. Due to long simulation times, the results are shown only for the first 200,000 I/O requests. Figure 5(a) shows the CDF of the response times using this technique for various values of the thermal safety. The response times are plotted for the baseline configuration which uses 10000 RPM disks, a configuration that uses disks whose speeds are 5000 RPM higher than the baseline but not needing DTM, and also disks of 17115 RPM that require DTM but plotted in order to show the difference between the idealized case where DTM is not required and the actual behavior under throttling. It can be observed that even though in the idealized case, the 17115 RPM disk drive can provide good performance, the application of DTM severely degrades performance wherein a majority of the requests have response times in excess of 200 ms (compared to 40-60 ms even for the 10000 RPM disk). This is because of delays incurred due to throttling very frequently (as shown in Fig. 5(b)).

An alternative approach to (indirectly) implement seek-based temperature management is by laying out the data on the disk drives in such a way that most of the time the physical seek does not expend much power. There are potentially two ways of achieving this: (1) make the seek shorter than the hottest seek profile (discussed earlier) (2) make the seek longer by introducing coasting periods (during which time the VCM is turned off). Figure 6 gives an indication of which approach could be more promising. In this graph, the temperature of the disk drives are plotted for various seek-time values (given on the x-axis) for a given value of the Inter-Seek Time (IST). (The Inter-Seek Time is the time

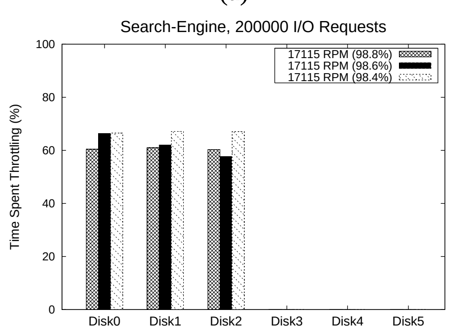


(a) (b)

Figure 4: Dynamic Thermal Management techniques. In (a), temperature control is achieved by modulating movement of the disk arm. In (b), RPM modulation is employed.



(a)



(b)

Figure 5: Impact of the Delay DTM technique for the Search-Engine workload

between the end of a seek operation and the beginning of another). The "summit" of each of the curves corresponds to the hottest seek where the arm accelerates to the maximum velocity and immediately decelerates towards the target track. By observing the relative slopes of the curves to the left and to the right of the peak, it is clear that disk seeks that are shorter than the hot value keep the disk temperature lower than those in the other region. Achieving such a seek profile is synergistic with performance-centric data layout optimizations and is also used by disk arm scheduling algorithms like Shortest Positioning Time First (SPTF) that are implemented inside the disk drive to reduce the seek time.

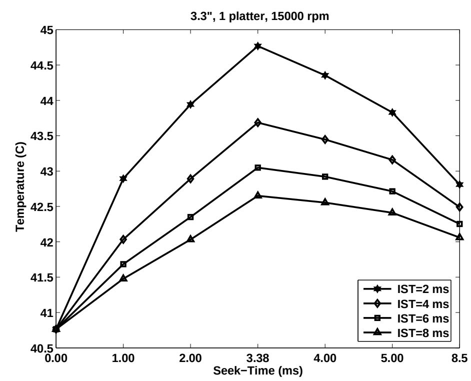
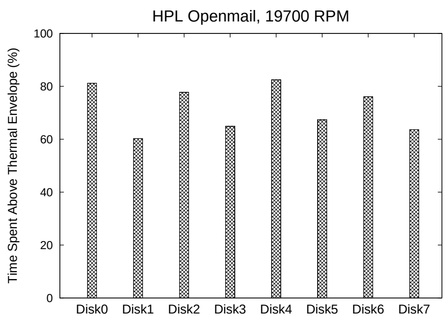


Figure 6: Relationship between seek-time and disk temperature for a 1-platter 15K RPM disk drive.

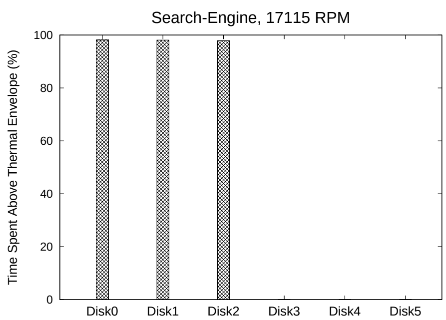
Another possibility is to modulate the RPM of the disk drive. Since the RPM has nearly a cubic impact on the viscous dissipation, the ability to modulate this component can be a very effective technique to control temperature. In this context, there are two possible techniques, which are of progressively increasing sophistication. In the first approach, we can use a disk drive that implements a low RPM idle mode (i.e., it cannot service any I/O requests) at the lower speed. Indeed, such hard disk drives do exist in the market today, such as the Hitachi Deskstar 7K400 [18]. Implementing DTM with such a disk is illustrated in Fig. 4(b). Here, when the temperature reaches the trip-point, seeking is stopped and the RPM is lowered in order to cool it down. Then, when I/O can be resumed, the RPM is ramped up to full-speed before servicing the pending requests.

A more promising approach can be to use a multi-speed disk drive that can actually perform I/O at the lower speeds as well. At the lower speeds, the disk is still available for I/O, although the rotational latencies and transfer times would be longer. Constructing a Dynamic RPM (DRPM) [19] disk drive requires a head assembly that can maintain its fly-height over the range of RPMs to be supported and also a variable-speed data channel to support the different bit rates. The key to the effectiveness of DRPM disk drives is the transition times between the various RPMs, in addition to the actual RPM values that they could support. It is worthwhile to study the design space of such disk drives to ascertain the extent of the benefit and the potential of DRPM as a long-term solution to the temperature problem for magnetic disk drives.

Is Energy Management = Temperature Management?



(a)



(b)

Figure 7: Percentage of workload time spent above the thermal envelope while servicing requests.

There has been extensive research on power management of disk drives. One popular technique that is used for multi-disk enterprise-class storage systems is to aggregate the data into a minimal set of drives at run-time and turn off the remaining ones [20]. However, funneling all the I/O traffic through just a subset of the disks can significantly increase the intensity of activity on them, leading to possible thermal violations. This is illustrated in Fig. 7.

The graphs show the percentage of time that the disk drives operate above the thermal envelope when servicing the I/O requests of these applications. The RPMs have been chosen such that they are candidates for temperature control via seek modulation (i.e., their temperature can be brought below the thermal envelope by not moving the disk arms). The Openmail system uses RAID and therefore its I/O requests are uniformly spread out over all the disk drives, which is the exact opposite to what is typically done for energy management. Due to this, the temperature of the disks are also lower, allowing them to operate below the envelope for a longer duration of time. On the other hand, for Search-Engine, almost all the I/O traffic is serviced by Disks 0-2. Although this can be used to save energy by spinning down disks 3-5, the former set of disks experience more severe excursions above the thermal envelope (analogous to a "hot-spot").

On the other hand, if the disks employ DRPM, one could potentially attain both energy savings and reduce the amount of heat that is generated, while still doing I/O, albeit at a degraded throughput. One could design appropriate control policies for doing the speed control for achieving these goals. This warrants serious consideration of DRPM as a feature, for at least enterprise class disk drives, to meet the performance demands and the thermal budgets in future storage systems.

CONCLUSIONS

This paper has shown that it is going to be challenging to design disk drives in the future that both provide growth along the established performance curve and also adhere to the thermal design envelope. Using real applications, the need for continued improvements in disk drive speeds was motivated and the need for designing for the average, rather than worst-case, operational conditions has been proposed. In order to achieve this, some form of Dynamic Thermal Management (DTM) would need to be provisioned. This paper has also shown that temperature management warrants special consideration and might not be achievable by merely applying techniques intended for reducing the energy consumption. However, it might be possible to meet both these optimization goals via DRPM-based techniques.

References

[1] I.Sato, K. Otani, M. Mizukami, S. Oguchi, K. Hoshiya, and K.-I. Shimokura, "Characteristics of Heat Transfer in Small Disk Enclosures at High Rotation Speeds," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, vol. 13, pp. 1006–1011, December 1990.

[2] G. Herbst, "IBM's Drive Temperature Indicator Processor (Drive-TIP) Helps Ensure High Drive Reliability," in *IBM Whitepaper*, October 1997.

[3] D. Anderson, J. Dykes, and E. Riedel, "More Than An Interface - SCSI vs. ATA," in *Proceedings of the Annual Conference on File and Storage Technology (FAST)*, March 2003.

[4] R. Wood, "The Feasibility of Magnetic Recording at 1 Terabit per Square Inch," *IEEE Transactions on Magnetics*, vol. 36, pp. 36–42, January 2000.

[5] E. Grochowski and R. Halem, "Technological Impact of Magnetic Hard Disk Drives on Storage Systems," *IBM Systems Journal*, vol. 42, no. 2, pp. 338–346, 2003.

[6] S. Gurumurthi, A. Sivasubramaniam, and V. Natarajan, "Disk Drive Roadmap from the Thermal Perspective: A Case for Dynamic Thermal Management," in *Proceedings of the International Symposium on Computer Architecture (ISCA)*, pp. 38–49, June 2005.

[7] P. Eibeck and D. Cohen, "Modeling Thermal Characteristics of a Fixed Disk Drive," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, vol. 11, pp. 566–570, December 1988.

[8] M. Sri-Jayantha, "Trends in Mobile Storage Design," in *Proceedings of the International Symposium on Low Power Electronics*, pp. 54–57, October 1995.

[9] "Hitachi Global Storage Technologies - HDD Technology Overview Charts." http://www.hitachigst.com/hdd/technolo/overview/storagetechchart.html.

[10] J. Chen and J. Moon, "Detection Signal-to-Noise Ratio versus Bit Cell Aspect Ratio at High Areal Densities," *IEEE Transactions on Magnetics*, vol. 37, pp. 1157–1167, May 2001.

[11] M. Mallary, A. Torabi, and M. Benakli, "One Terabit Per Square Inch Perpendicular Recording Conceptual Design," *IEEE Transactions on Magnetics*, vol. 38, pp. 1719–1724, July 2002.

[12] G. Tarnopolsky, "Hard Disk Drive Capacity at High Magnetic Areal Density," *IEEE Transactions on Magnetics*, vol. 40, pp. 301–306, January 2004.

[13] C.D. Patel, "A Vision of Energy Aware Computing from Chips to Data Centers," in *The International Symposium on Micro-Mechanical Engineering (ISMME)*, December 2003.

[14] G. Ganger, B. Worthington, and Y. Patt, *The DiskSim Simulation Environment Version 2.0 Reference Manual*. http://www.ece.cmu.edu/ ganger/disksim/.

[15] Y.Kim, S. Gurumurthi, and A. Sivasubramaniam, "Understanding the Performance-Temperature Interactions in Disk I/O of Server Workloads," in *Proceedings of the International Symposium on High Performance Computer Architecture (HPCA)*, pp. 179–189, February 2006.

[16] H. Ho, "Fast Servo Bang-Bang Seek Control," *IEEE Transactions on Magnetics*, vol. 33, pp. 4522–4527, November 1997.

[17] "Samsung Electronics Develops Solid State Disk Using NAND Flash Technology," May 2005. http://www.samsung.com/PressCenter/PressRelease/PressRelease.asp?seq=20050523_0000123980.

[18] "Hitachi Power and Acoustic Management - Quietly Cool," in *Hitachi Whitepaper*, March 2004. http://www.hitachigst.com/tech/techlib.nsf/productfamilies/White_Papers.

[19] S. Gurumurthi, A. Sivasubramaniam, M. Kandemir, and H. Franke, "DRPM: Dynamic Speed Control for Power Management in Server Class Disks," in *Proceedings of the International Symposium on Computer Architecture (ISCA)*, pp. 169–179, June 2003.

[20] E. Pinheiro and R. Bianchini, "Energy Conservation Techniques for Disk Array-Based Servers," in *Proceedings of the International Conference on Supercomputing (ICS)*, June 2004.