Datacenter Scale Evaluation of the Impact of Temperature on Hard Disk Drive Failures

SRIRAM SANKAR, Microsoft Corporation
MARK SHAW, Microsoft Corporation
KUSHAGRA VAID, Microsoft Corporation
SUDHANVA GURUMURTHI, University of Virginia

With the advent of cloud computing and online services, large enterprises rely heavily on their datacenters to serve end users. A large datacenter facility incurs increased maintenance costs in addition to service unavailability when there are increased failures. Among different server components, hard disk drives are known to contribute significantly to server failures; however, there is very little understanding on the major determinants of disk failures in datacenters. In this work, we focus on the inter-relationship between temperature, workload, and hard disk drive failures in a large scale datacenter. We present a dense storage case study from a population housing thousands of servers and ten thousands of disk drives, hosting a large scale online service at Microsoft. We specifically establish correlation between temperatures and failures observed at different location granularities: a) inside drive locations in a server chassis, b) across server locations in a rack and c) across multiple racks in a datacenter. We show that temperature exhibits a stronger correlation to failures compared to the correlation of disk utilization with drive failures. We establish that variations in temperature are not significant in datacenters and have little impact on failures. We also explore workload impacts on temperature and disk failures and show that the impact of workload is not significant. We then experimentally evaluate knobs that control disk drive temperature, including workload and chassis design knobs. We corroborate our findings from the real data study and show that workload knobs show minimal impact on temperature. Chassis knobs like disk placement and fan speeds have a larger impact on temperature. Finally, we also show the proposed cost benefit of temperature optimizations that increase hard disk drive reliability.

Categories and Subject Descriptors: **C.5** [Computer System Implementation] – Servers; **B.8** [Hardware - Performance and Reliability] General;

General Terms: Design, Experimentation, Reliability

Additional Key Words and Phrases: Datacenter, Hard disk drives, Temperature impact

ACM Reference Format:

Sankar, S., Shaw, M., Vaid, K., and Gurumurthi, S. 2012. Datacenter Scale Evaluation of the Impact of Temperature on Hard Disk Drive Failures. ACM Trans. on Storage.

1. INTRODUCTION

As large enterprises move to modular datacenters [Hamilton 2007] and efficient cooling practices become prevalent [Greenberg et al. 2006], we move closer to the limits of cost efficiency achievable in that domain. Capital and operational cost margins are first order constraints for large scale online services like Search and Cloud Computing. Traditional datacenter designs that guaranteed stable operating conditions are giving way to more flexible designs that cut cost at the expense of datacenter operating conditions. Hence understanding impact of datacenter

---

Author's addresses: Sankar, S., Shaw, M., Vaid, K.,, Microsoft Corporation, Redmond. Gurumurthi. S., Computer Science Department, University of Virginia.

DOI10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

operating conditions on server and datacenter reliability is of utmost importance in large scale datacenters.

Server components are typically composed of commodity electrical and mechanical parts, and hence they are prone to failures. Frequent failures reduce infrastructure availability and increase the cost of datacenter operations. In addition, the server design in itself could be a major catalyst for most of the server component failures. For instance, we found that a particular drive location in a dense storage configuration under a fairly constant workload was continuously exposed to high temperature conditions, even under nominal inlet temperature to the server. We found a higher number of drives in this location failing more often, thereby showing strong correlation to operating conditions. Understanding the reason behind such failures enabled us to address the design issues, thereby increasing the availability of machines for the particular online service. Availability of online services is a key differentiator in today's competitive market. Higher server reliability ensures that online services can have increased availability. Increasing the number of available servers also delays the need for provisioning new server deployments in datacenters. New server deployments have a longer delay cycle, and might cause a high impact launch to be delayed, thereby causing significant financial damage to the enterprise. Hence, having more servers that are readily available affects the financials of a large enterprise.

1.1 Motivation – Hard Disk Failures in Datacenters

Server component failures have indeed been recognized as important and prior works have studied individual component reliability, such as for hard disks [Schroeder et al. 2007] [Pinheiro et al. 2007] and memory [Schroeder et al. 2009]. Figure 1 presents actual data on the different kinds of failure types observed over a period of two years from typical large-scale datacenters housing more than 100,000 servers. We see clearly that hard disk drives account for 71% of the known failures, making it the most dominant failing part. This is in part due to the mechanical moving parts of the disk drives and also due to the extensive use of commodity SATA drives in large deployments. SATA disk drives are known for failing more often than SAS drives, but are also cheaper for storage capacity per dollar [HP 2003]. Given that hard disk drives are the most significant failing component and recent previous studies established no conclusive relationship between temperature and hard disk drive failures [Pinheiro et al. 2007], we evaluate whether temperature experienced at the hard disk drive has stronger correlation to failures in this work.

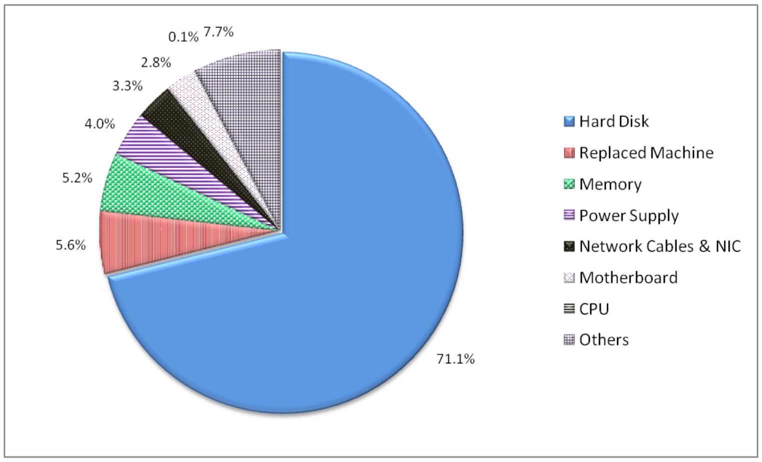


Figure 1. Breakdown of Hardware Component Errors in a Large Datacenter

1.2 Datacenter Total Cost of Ownership (TCO)

Total Cost of Ownership defines the overall cost that a large enterprise incurs to build and operate a large datacenter. In addition to capital expenditure costs, the TCO incorporates the operational costs of maintaining the datacenter, and hence is a holistic representation of the cost. We use the Total Cost of Ownership (TCO) model for a datacenter given by Hamilton [Hamilton et al 2008]. We use the TCO model with the following assumptions: We use a typical large scale datacenter with 10MW critical power capacity, and a Power Usage Effectiveness (PUE) of 1.25. PUE refers to the fraction of power consumed by the entire facility including cooling divided by the power consumed by IT equipment alone. A PUE closer to 1 denotes a very efficient datacenter facility (1.25 PUE is typical of traditional datacenters similar to the one used in our study). We use $10 per Watt for construction costs and a cost of $0.10 per Kilowatt-hour for utility power costs. We assume that the total cost of an individual server is $2000 and each server has a typical power draw of 200 Watts to calculate the Server Capital Expenditure. We assume a 3 year server amortization and a 15 year datacenter amortization for computing the amortization costs.

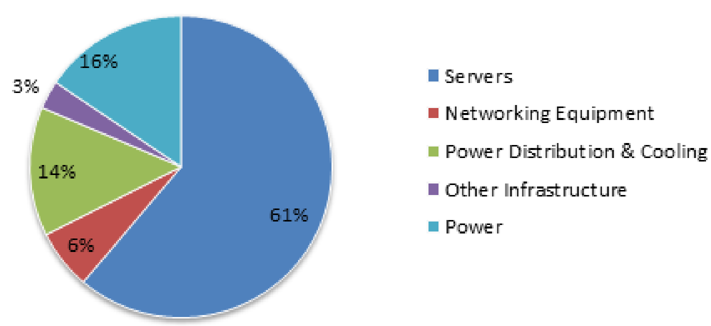


Figure 2: Datacenter 3 year Total Cost of Ownership

As can be seen from Figure 2, 61% of the total cost of ownership of a datacenter is contributed by the actual cost of the server. This is desirable, since we want to spend as much as possible in putting in more servers into the datacenter. However,

approximately 30% of the cost is contributed by the power related expenses. Datacenter designers like to reduce this cost and make tradeoffs accordingly. One straightforward methodology is to increase the temperature at which the datacenter operates, and hence reduce the amount of cooling overhead. However, this methodology comes with a consequence – it increases the failures in the datacenter, and hence makes it necessary to purchase more servers or repair the servers that failed. That increases the TCO of the datacenter, since we need to stock more hardware components in our supply, and also maintain a larger team of technicians to replace failing components. Hence, there is a clear tradeoff between reducing cooling costs through temperature control, and the reliability in a datacenter. Given that hard disk drives contribute to more than two-thirds of all hardware replacements, we explore this tradeoff with respect to hard disk drive failures and temperature in this work.

1.3 Major Contributions

In this work, we establish the different aspects of correlation between temperature and disk drive failures observed from the large datacenter case study. In addition to temperature impact at different granularities, our paper quantitatively evaluates the impact of variations in temperature as measured in a live production environment. We also explore whether workload variations cause temperature behavior to be impacted, and also if workload intensity has correlation to failures observed in the datacenter. We conduct experimental studies to validate our observations from real data.

In summary, our major contributions are:

1) We show strong correlation between temperature observed at different location granularities and failures observed. Specifically, we establish correlation between temperatures and failures observed at the following location granularities: a) inside drive locations in a server chassis, b) across server locations in a rack and c) across multiple racks in a datacenter.
2) Although average temperature shows a correlation to disk failures, we show that *variations* in temperature or workload changes do not show significant correlation to failures observed in drive locations.
3) We corroborate our findings from the datacenter study through an experimental evaluation and show that Chassis design knobs (disk placement, fan speeds) have a larger impact than tuning Workload knobs (intensity, different workload patterns), on disk temperature.
4) With the help of Arrhenius based temperature models and the datacenter cost model, we quantify the proposed benefits of temperature optimizations and increased hard disk drive reliability and show that datacenter temperature control has a significant cost advantage over increased fan speeds.

We believe that this work shall motivate new research in analyzing tradeoffs for datacenter optimizations, given the recent investment in large scale cloud computing.

Our paper is organized as follows: Section 2 discusses related work in this field, while Section 3 specifies our experimental infrastructure, including datacenter measurement infrastructure and workloads. Section 4 presents the data and observations from a large scale datacenter study on the impacts of temperature and

workload on disk drive failures. In Section 5, we discuss the experimental evaluation of the different temperature control knobs that we considered in our study. Section 6 presents the Arrhenius based reliability model and an application of the model, while in Section 7, we provide a cost analysis of the different optimizations. Section 8 discusses future work, and we conclude the paper in Section 9.

2. RELATED WORK

Server component failures and reliability are yet to be understood completely. Previous research works in this field have generated conflicting results, especially in relation to subjects like the impact of temperature on disk drive failures. With respect to large scale installations, Gray et al [2005] observed failure rates ranging from 3.3-6% in two large web properties at Microsoft. Schwartz et al [2006] report failure rates of 2-6% in the drive population at the Internet Archive. Elerath et al [2004] report that end-user failure rates can be as much as ten times higher than what the drive manufacturer might expect in their study on server class disk drives. Schroeder et al [2007] find that in the field, annual disk replacement rates typically exceed 1%, with 2-4% common and up to 13% observed on some systems. The authors also present interesting per-component failure percentages for three different types of systems that they considered. They also report a significant overestimation of mean time to failure by manufacturers. Schroeder et al [2006] in their study of failures in petascale computers, review sources of failure information for compute clusters and storage systems, and project corresponding failure rates. There are research works that explore the tradeoffs between workload characteristics and temperature with the help of simulation [Kim et al. 2006], but do not consider reliability impacts. Our paper considers the inter-relationship between workload, temperature and disk drive reliability.

One of the most closely related works to this study is by Pinheiro et al [2007], which identified correlation between disk errors and SMART attributes from a large population of serial and parallel ATA drives. This paper also concluded that temperature and activity levels had less correlation to disk failures and was a surprising result when compared to previous studies [Cole et al. 2000][Yang et al. 1999]. Recently, El-Sayed et al [2012] show that temperature correlation to failures are weaker than expected in a diverse population of disks, and point out there might be other factors that are more dominant than temperature, whereas we try to eliminate the impact of diverse factors by selecting a more controlled environment. Yang et al [1999] establishes that a 25 C delta in temperature derates the MTTF by a factor of 2 in their study on Quantum hard disk drives. Cole et al [2000] from Seagate, present thermal de-rating models showing that MTTF could degrade by close to 50% when going from operating temperatures of 30C to 42C. Our results agree with the observations made by Cole. Our measured failure rates also exceed the AFR rates that manufacturers mention in their datasheets [Seagate ES 2011]. Also interestingly, Vishwanath et al [2010] report no correlation between failures and location of servers within a rack. We find in our case study that temperature does have a strong correlation to failures (within chassis, racks and across racks). We propose that temperature impacts for datacenter scale environments should be factored in knowing the server configuration and datacenter inlet temperature range.

3. EXPERIMENTAL INFRASTRUCTURE

3.1 Temperature Measurement Infrastructure

We perform our data measurements on a population measuring thousands of servers and ten thousands of hard disk drives. All the servers in this study are identical, with dual CPUs and an additional storage enclosure containing up to 40 SATA drives in a RAID 1+0 configuration. In our server chassis, we are able to fit 5 disk drive columns (3.5" SATA HDD) across the length of the server. The traditional datacenter racks have a cold aisle from which cold air is pulled across the server and exhausted out in the hot aisle [Hoelzle et al. 2009]. Hence the air gets preheated by the time it reaches the interior hard disk drives and leads to higher temperatures for those hard disk drives.

The temperature measurements are collected by SMART counters (counters monitored as part of every disk drive's self-monitoring facility) from a sensor included in the HDD enclosure in every hard disk drive. The SMART counters for temperature are logged every 20 minutes by the array controller at a local controller log along with several other SMART counters. Every day this local server log is shipped to a central server and archived. Since the population is in a live production environment and there are various data that is collected, the duration of sampling is limited to 20 minutes on account of data storage limitations.

3.2 Server Test Configuration

For evaluating the impact of server chassis design parameters including placement of disk drives and fan speeds, we use a dense storage disk enclosure [HP 2011] along with a standard enterprise server. This dense storage enclosure is setup to mimic the actual production setup as close as possible. However, this does not directly reflect any production storage configurations for proprietary reasons. The test server also has a controller that can log instantaneous temperature at each disk drive.

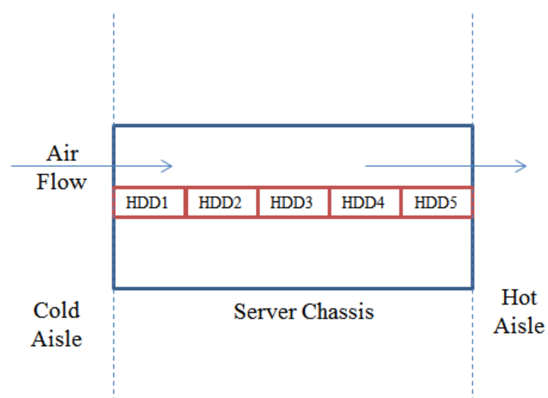


Figure 3: Dense Storage configuration and layout

The storage enclosure has 5 columns of hard disk drives arranged from right to left as shown in Figure 3. For this enterprise configuration, there are 34 disk drives present in an enclosure that can hold up to 35 disk drives. This presents an opportunity on which disk bay to leave empty, and we explore this tradeoff in later

sections. The logical drive is made up of all the 34 disk drives in a RAID 1+0 configuration that is typical of an enterprise RAID setup.

3.3 Workloads

For analyzing workload behavior and its resultant impact on varying temperatures, we use real datacenter storage workloads obtained from trace characterization. Enterprise storage systems are designed to support multiple concurrent users in order to amortize the cost of data access over a large number of users. Hence enterprise workloads are typically composed of random IO operations, with high inter-arrival rates. Table 1 shows the four workloads that we consider being representative of large scale datacenter workloads.

Workload	**Rd:Wr Ratio**	**Random %**	**Dominant Block Size**	**Average Inter-arrival (ms)**
Email	1.4	83%	8K	1.48
UserContent	7.0	91%	4K	22.22
Exchange	2.0	66%	32K	0.71
Messenger	9.6	99%	8K	0.30

Table 1: Datacenter Workload Characteristics – random access with short inter-arrival times.

A denser storage solution typically acts backend storage for applications that require a lot of data storage, like Email and OLTP applications, since denser solutions makes it possible to pack more storage in lesser space. Hence for testing such a high density storage solution, we use storage profiles of Email backend server (Email), a large scale file system server at Microsoft (UserContent), Exchange server (Exchange) and an OLTP backend profile (Messenger) that represents user meta-data for a large online service. The trace characterization framework is based on ETW (Event Tracing for Windows) [Park et al. 2007] and it captures the disk IO events at the OS level. This ensures that if we design a system that is configured similarly, regenerating IOs as captured during the trace will be truly representative of a datacenter workload. We use publicly available disk IO generator like IOMeter [IOMeter 2011] to replay the workload for our experiments.

As can be seen from Table 1, all workloads have short inter-arrival times. UserContent is a file server workload with minimal storage requests, and has a larger inter-arrival time of 22.2 milliseconds between IO requests. Note that the other applications including Email, Exchange and Messenger (OLTP) workloads have less than 2 milliseconds between each IO request. Also, note that all these workloads are mostly random (66%-99%). Random IO requests require disks to seek to particular locations on the disk drive, and hence consume more power and hence could result in possible increase in temperature. Since most of these workloads are random, the disk drives are continuously performing seek activity and also has no time to shut down or save power, since inter-arrival times are relatively short. In addition, typically for enterprise workloads, seek activity is composed mainly of short-distance seeks [Kim et al. 2006] and hence there is minimal impact on temperature. We use this observation in the later sections to motivate our experimental evaluation to select different knobs that have impact on temperature.

4. REAL DATACENTER CASE STUDY

In this section, we present a case study with data collected from a live datacenter facility. We analyze the major determinants of hard disk failures, and explore temperature correlation in depth. The hard disk drive failures that are considered here denote actual hardware replacements as viewed from the datacenter management perspective. Detailed failure analysis that can identify sub-component errors or false positives (similar to manufacturer lab analysis) is not typical in such high security environments. Throughout the paper we define failures as events leading to system downtime that was fixed by a replacement of the component in the datacenter floor except when specified.

4.1 Hard Disk Drive Failure Determinants

There are a number of factors that can influence hard disk drive failures, including age of the disk drive, utilization on the hard disk drive (general wear and tear due to use), temperature of operation, and vibration.

4.1.1 Age of the Disk Drive

Several previous studies have established different failure rates with respect to the age of the disk drive population [Pinheiro et al. 2007]. A typical failure curve across age resembles a Weibull bathtub shaped curve with a large number of infant mortality, stable mid-life curve and steady increase in failures again at older age. In our study, most of the disk drives are of similar age since all the servers were deployed around similar timeframe when the datacenter became operational, and are past the infant mortality stage. Hence the age factor does not become a major determinant for our study. This is extremely beneficial since this helps isolate the impact of other factors on failure rates in datacenters.

4.1.2 Vibration and SMART Monitors

There could be significant vibration due to dense storage; however modern hard disk drives balance internal vibration through vibration compensation techniques in the servo mechanism of the hard disk drives [Guo et al. 2003]. We currently do not have metrics that expose the level of induced vibration, and measuring the impact of vibration is one of our projects that are currently underway. We do collect several SMART data from the disk drive population, including Reallocated Sector count, Seek errors, Spin up time, ECC errors, Temperature etc. Though we see SMART counters being indicative of some failures, a predictive methodology is hard to obtain. For one of our large populations, such a methodology would have been able to account for less than 20% of all disk failures. We do not present the details here in interest of space. Previous conclusions made by Pinheiro et al. [2007] also suggest that SMART counters do not provide a confident way of predicting hard disk drive failures.

4.1.3 Utilization vs Temperature

The remaining two significant failure determinants are disk utilization and temperature. We need to isolate the impact of these two metrics that are location

dependent. One of the primary factors that can cause more wear on the hard disk drive is the disk utilization (we use utilization as a proxy for workload duty cycle), which denotes the amount of activity on the hard disk drive. According to the volume and data layout, certain disks might be more stressed than other disks (for instance, a data volume in SQL might have higher level of activity than a Backup volume). We conducted a preliminary investigation to determine which of these two metrics is highly correlated to hard disk drive failures.

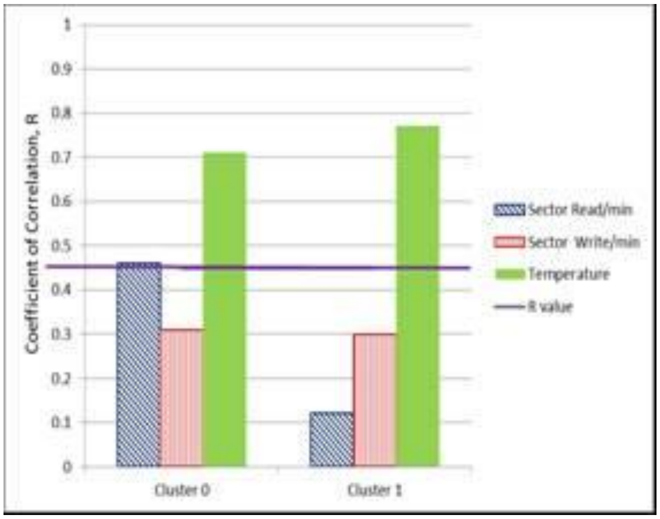


Figure 4 Temperature shows better correlation to HDD failures than Workload Utilization

Figure 4 presents the results of the analysis on a total of 10000 hard disk drives spread across two clusters. We correlated the 'sectors read/ minute' and 'sectors write/ minute' experienced by the disk drive in a particular location as seen by the controller over its entire lifetime, to the failures observed in that location over a year. On the other hand, we also correlated the temperature observed in those disk locations to the number of failures. We plot the resulting coefficient of correlation in Figure 4. As can be seen from the figure, the read and write activity on the disk drives correlate minimally with the failures. However, drive temperature inside the chassis shows stronger correlation to disk failures in the particular location within the chassis (R value for temperature is above the critical R value line, at df=30 for a two-tailed test at level of significance = 0.01). Hence for the remainder of the paper, we concentrate on disk drive temperature and do an in-depth temperature measurement and correlation analysis across disk drive locations inside chassis, location of a server within a rack and locations of racks in a datacenter.

4.2 Correlation of Disk Failures with Average Temperature

We present a case study where specific datacenter design parameters and a dense storage chassis design resulted in higher number of disk failures, under high operating temperature. The case study was conducted in a raised-floor datacenter, containing tens of thousands of hard disk drives in a dense storage server and failure data was collected for a period of 1 year.

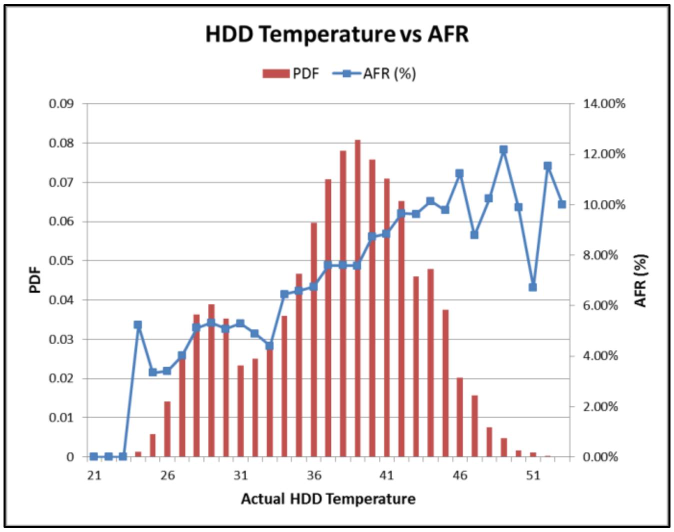


Figure 5 Failure rates at different hard disk drive temperatures

The result of our study is surprising since earlier studies [Pinheiro et al. 2007] establish that disk drive failures do not increase with increase in temperature in the field. Figure 5 shows the actual HDD temperature in increments of one degree and the corresponding AFR for our entire population. We see clearly that with increase in HDD temperature, the AFR rate increases. There are some data points at the end of the spectrum that have smaller number of samples and hence a higher skew. For the major part of the distribution (shown by PDF columns), we see that AFR steadily increases as HDD temperature increases. Interestingly, we found that the certain disk locations in the heavy storage enclosure were exposed to high temperature for a longer duration even under nominal inlet operation temperatures. We also observed a significant difference between the inlet temperatures measured at different locations in the datacenter. In the next section, we present our analysis and observations categorized by location granularity. We divide our correlation analysis into three distinct temperature impact zones: Drive locations inside the server chassis; Server locations within a rack and multiple rack locations across the datacenter. There are different factors that come into play for each of these temperature zones. We shall discuss each in more detail in the following sections.

4.2.1 Correlation inside the Server Chassis

Server design is an important factor in determining the availability of machines in a datacenter. Depending on the placement of the hard disk drives, there could be significant variation in drive temperature. This is especially true in the case of dense storage, since cold air flows from the front of the storage enclosure to the back. Given that the workload running on the disk drives are similar (no significant duty cycle variations), we can establish the correlation if there are more failures for drives which experienced higher operating temperatures. We present the layout of a dense storage device in Figure 3 that was used in our case study. There are five hard disk drives columns where HDDs are arranged one behind the other from the front of the enclosure to the back. Hence the air gets preheated by the time it reaches the interior hard disk drives and leads to high temperatures for those drives. This results in an increase in number of failures observed in that location.

Figure 6 (a) shows the average temperature observed in each hard disk drive column (1 through 5) across all the machines under this study. Note that the temperatures increase from 27 C in the front-most hard disk drive (HDD1) to 39 C in the fourth hard disk drive column (HDD4). This is just the average temperature measurement, and there were hard disk drives that were at temperatures greater than 45 C in hotter parts of the datacenter as shown in Figure 4. The last drive (HDD5) closer to the hot aisle has a reduced temperature due to heat dissipation at the outlet. The corresponding total failures observed across the entire server population over a period of 1 year are denoted by the AFR line. Note that we present Annual Failure Percent (which is a measured population based value and should not be considered as the Annualized Failure Rate, which is a calculated metric that manufacturers provide) for our population that is on continuous mode of operation throughout the year (For a discussion on different annual failure rates, please see Elerath et al. [2004]). The failure rates measured here are hence not reflective of manufacturer quoted rates, and should be considered only as number of failures out of the population under deployment. Out of the hard disk drives that were in the front-most part of the server chassis (HDD1), only 4% failed, whereas, for the fourth hard disk drive (HDD4) around 6% of the total disks failed. This is almost 1.5X the number of failures compared to the front of the chassis. This result shows a strong correlation between temperatures observed through the SMART logs collected at the machines and the observed failures reported in this datacenter. In fact, the correlation coefficient measured across the entire population for (*average temperature for drive locations inside the chassis, number of failures*) pair is **R = 0.79**, which is significantly high. Our experience with this dataset does point out that lower temperature locations do have lower failures, and as system designers it is a strong motivation for reducing temperature impact inside a chassis design. Fan speed and airflow management helps reduce such temperature impact.

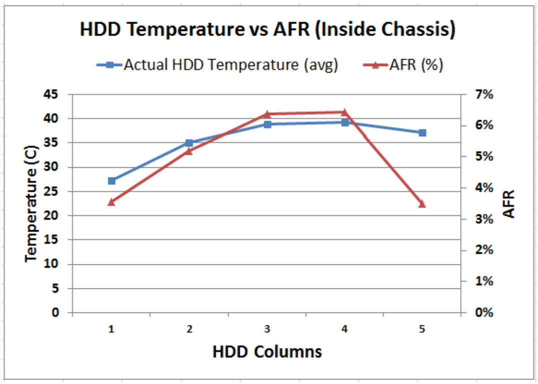
Observation: ***There is a significant correlation (r = 0.79) between actual hard drive temperature inside a server chassis design and the number of drive failures. Hence chassis design should incorporate temperature reduction optimizations.***

4.2.2 Correlation across Servers in a rack

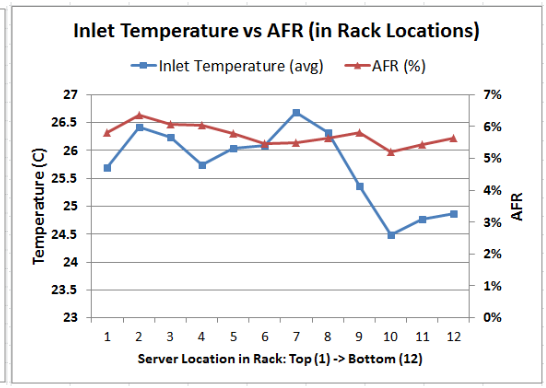
A datacenter rack consists of multiple server chassis arranged on top of each other. The cool air comes through vents closer to the bottom of the rack and rises upwards. It is pulled across the server as it rises up and that direction is horizontal (as shown in Figure 3). However as it moves up through the vertical direction, there is an increase in air temperature due to heat dissipation. There are also other mechanical impacts such as the differences in air pressure (cfm) at different server locations within a rack. In this section we explore if the server location and inlet temperature observed at each location correlates with the number of disk failures observed at that server location.

From Figure 6 (b), we see that for the cooler servers (Location 9, 10, 11, 12) that are on the bottom of the rack, the number of failures is lesser (closer to 5%) as compared to hotter servers (Location 2) at 6% failure rate. This shows a strong correlation between server locations inside a rack and the number of failures. This again reiterates our observation that temperature and air movement across a rack are significant determinants for server failures. The correlation coefficient computed for (*inlet temperature for server location within rack, number of failures*) pair is **R = 0.91.**

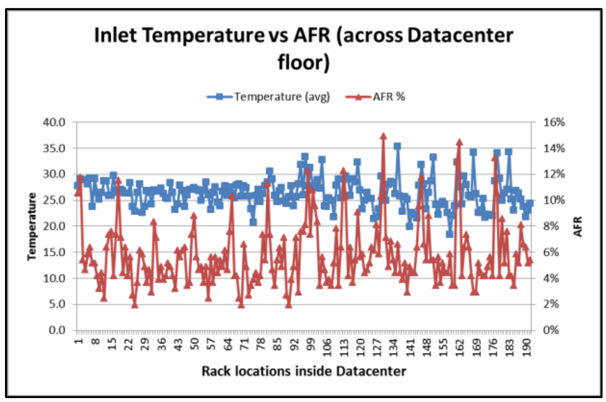
Observation: ***There is a significant correlation (R = 0.91) between the inlet temperatures observed with respect to the position of the server in the rack and number of failures for that server. The higher the average inlet temperature at a server location within a rack, the higher the number of failures.***



a) Within Server Chassis



b) Within a rack



c) Across rack locations within datacenter

Figure 6: Correlation across different location granularities

4.2.3 Correlation across multiple rack location

Having seen that drive bay location and server location temperatures are indeed major determinants for number of failures observed in that location, we also determine whether the temperatures observed across rack locations inside the datacenter are correlated to the number of failures observed. Figure 6 (c) presents the temperature observed at the particular rack location (averaged across the servers in the rack). Every cluster in the datacenter has columns of multiple racks. Each column has an inlet cold aisle and a corresponding hot aisle. Every rack has 12 servers.

One important observation from Figure 6 (c) is that we would expect the Temperature line to be fairly horizontal at a fixed datacenter set-point temperature. However this is not the case and there is significant variation in temperatures across the datacenter floor. This is possible due to a variety of reasons including inefficient hot aisle/cold aisle containment, other networking or server gear venting hot air into the cold aisle and hot air recirculation around the edges. There are other significant patterns observable from the Figure, especially that the rises in temperature are accompanied by rises in failures, however we note that there are several places in the figure where this is not the case. However, the correlation coefficient for the entire set of data *(temperature at datacenter location, failures at that location)* is **R = 0.30.** There is indeed a positive correlation and is statistically significant (critical value of R at df=120 is 0.232 for a two-tailed test at level of significance = 0.01). Also, it is clear that the lower temperature racks have lower failures and hence the motivation to be temperature-aware in datacenter and server design is still valid.

Observation: ***There can be varying degrees of deviation from the Datacenter set-point temperature in different parts of the datacenter floor. Hence hot and cold aisle containment solutions are needed for higher efficiency in traditional datacenters.***

4.3 Impact of variations in temperature on failures

Having observed the correlation of failures with average temperature measured at different granularities, we explore whether variations in the temperature experienced by the disk drive has any correlation with failures. Instead of just comparing variance or standard deviation which has no reference to the mean around which the variation occurs, we use the coefficient of variation as a representative metric. This metric is a normalized measure of dispersion of a probability distribution and it computes the variation of temperatures relative to the mean ($CV = \sigma/\mu$). We saw that average temperature experienced by disk drives already has a strong correlation to failures. We also want to answer if large variations in temperature impact failure rates.

Figure 7 shows the correlation between CoV (Coefficient of variation) clustered into discrete buckets (each with 0.001 CoV) and the corresponding AFR for all disks falling into this bucket. We also plot the PDF of the distribution to show places where there are high frequencies in the distribution. As can be seen from the figure, the actual variation of temperature measurements is around 0.8% - 3.4% of the mean for most of the hard disk drives. This number in itself is relatively small, since typical average temperature ranges between 35C-40C and this variation amounts to a small deviation from this mean. This is due to the fact that in a traditional datacenter, inlet temperature to the servers is tightly controlled by a chilled water loop [Patterson 2008], and is expected to show lesser variation. Moreover, the temperature difference that we observe is between different disk drive locations across the chassis and rack and is not localized to each disk drive. This also agrees with our observation that workload variations (seek requests) are expected to cause minimal variation to individual disk drive temperature.

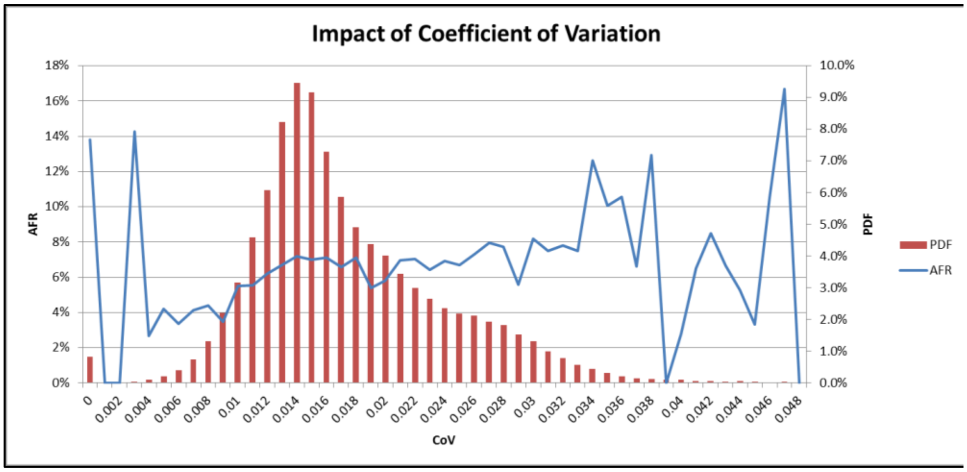


Figure 7: Correlation between AFR (failure rate) and CoV (Coefficient of Variation)

From the figure, we observe that there is no significant correlation between the CoV and the resulting AFR (R value of 0.21 is lower than critical value of R required for statistical correlation), though there is a slight uptrend and a positive correlation at certain CoV. For comparison purposes, note that correlation coefficient of average temperature at different chassis and rack locations with failures was in the 0.8-0.91 range. Note that this population is from an identical server design, housing a homogeneous load-balanced datacenter application, and hence has little variation in terms of age, disk drive model or workload intensities.

Observation: ***This analysis shows that 1) temperature variation relative to average temperature in large datacenters is minimal (less than 5%) and 2) temperature variation does not show a strong correlation to hard disk drive failures in the population under study.***

4.4 Impact of Workload on temperature and failures

In the above section we identified that variations in temperature do not correlate with failures. However, we also want to independently see whether workload variations were the cause of either temperature or failure. In this section, we compare workload measurements to temperature and failure measurements separately.

4.4.1 Workload Intensity and Temperature

The collected data set also contains the total number of read and writes operations done on the disk drive at every collection interval. This is a useful metric to have, since we can figure out the disks that were stressed more when compared to other disks. We can then correlate the observed temperature at the disk drive to see whether the disk that had a lot of workload requests was at a higher temperature than other disks.

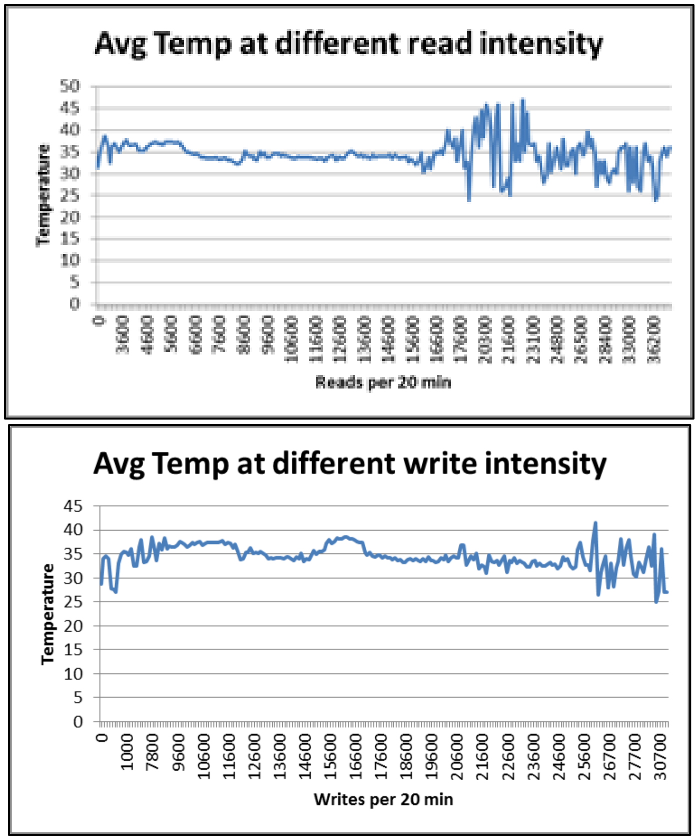


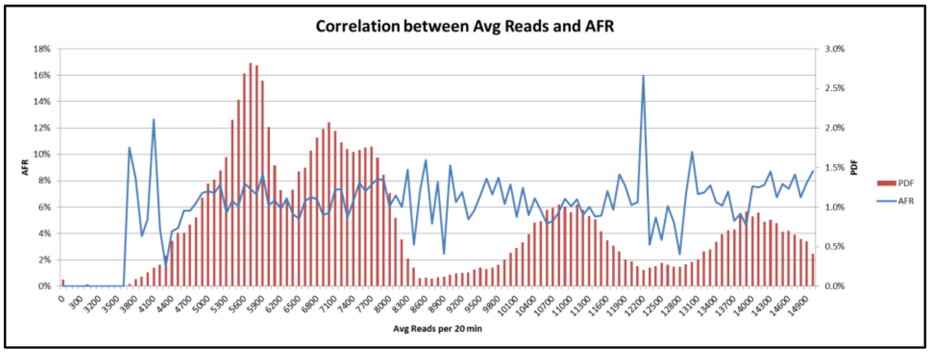
Figure 8: Temperature at increasing read and write intensities at all disk drives

In Figure 8, we plot the workload intensity at each drive and the corresponding average temperature experienced by the drive at that particular workload intensity. We plot both read and write intensities. Note that these intensities are also measured for a 20 minute interval due to the datacenter data collection limitations. However, we expect heavily accessed disks to have consistent high access rates during the entire period of operation, since essentially intensity is a sum of all requests over a 20 minute window and to have a sum that is large, the individual intensity measured every second (IO operations per second or IOPS) should have been large. From the figure, we are able to note that for both read and write operations, increasing intensities do not show a relative increase in temperature of the drive. As we move to the higher write and read intensities, we see temperature swings that are very high – this is due to the fact that the sample size at those high intensities is low and averaging them yields skewed temperature numbers. We show this data in the graph for completeness, but at most intensities where there is sufficient number of samples; we see no direct correlation between temperature and intensities. This confirms our earlier hypothesis that enterprise workloads have very little idle time [Gurumurthi et al. 2003], and the resulting continuous operation typically shows little or no change in temperature behavior of the disk drive [Gurumurthi et al. 2005], such that it deviates by a significant amount from the average temperature experienced throughout.

Observation: ***Workload variations do not impact temperature variations significantly for our load-balanced datacenter application.***

4.4.2 Workload Intensity and Failures

In this section, we correlate workload intensity experienced at each disk drive, with the failures experienced by disk drives. Figure 9 shows the correlation between average reads/ average writes and the corresponding failures for disk drives that experience that read/write intensity. In both the charts, X-axis plots increasing read and write intensities measured per 20 min. Y-Axis plots the failure rate for all disk drives that experienced that read intensity. We also plot the pdf to show the distribution of read and write intensities over the measured population. We see from both the charts that there is no correlation between the read or write intensity to the failures experienced by the disk drives. This conclusively shows that workload variation in itself does not impact hard disk drive failure at datacenters.



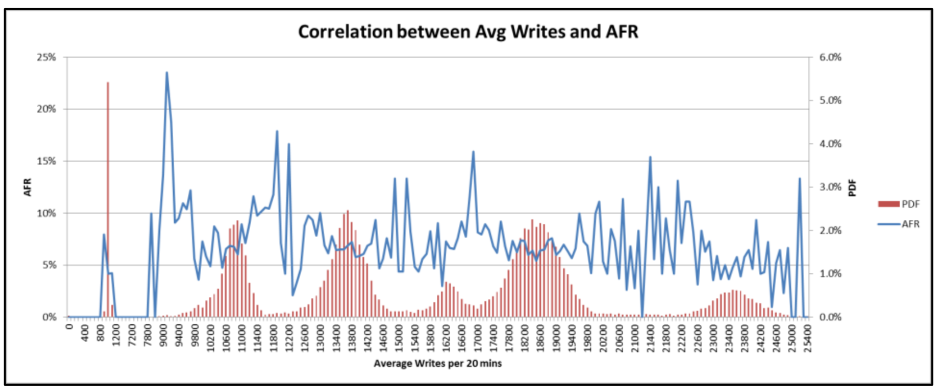


Figure 9: Correlation between workload intensities and failure rates

Observation: *There is no significant correlation between workload intensities and failure rates in the datacenter population.*

4.5 Summary of observations from Datacenter data

In summary, we see that average temperature has a strong correlation to disk failures at different locations inside chassis, rack and across datacenter floor. However, we do not observe a significant correlation between variations in temperature and disk failures. The variations in temperature are within 5% of the average and hence are not significant enough a concern for datacenter design. We also see that workload variations have minimal impact on temperature variations or hard disk drive failures in the datacenter.

5. EVALUATION OF TEMPERATURE CONTROL KNOBS

Given the results of the datacenter study, that showed significant correlation between temperature and failure rates, we evaluate the validity of our observations through controlled lab experiments. In this section, we control knobs that can influence temperature and experimentally quantify the benefit of each knob. This evaluation is done on a real system resembling the actual production system in a controlled lab environment. We evaluate the following temperature control knobs in this section:

1) Workload knobs (Intensity, Different workloads)
2) Chassis design knobs (Disk placement, Fan speeds)

5.1 Workload Knobs

In Section 4, we saw that workload variations have minimal impact on temperature. We want to validate this with the help of an experiment, where we control two workload knobs – we modulate the workload intensity by controlling inter-arrival rates; and we also run different workloads that have different access patterns on the same experimental system. We then compare the impact of these two knobs on disk drive temperature.

5.1.1 Impact of Workload Intensity on disk temperature

We modulate the inter-arrival rate of the workload by delaying the time between every IO request. We simulate various inter-arrival rate from 1 ms, 10 ms, 100 ms, 1000 ms and 10000 ms. We also simulate an artificial workload with 0 inter-arrival time – basically, the workload sends as much requests as it can limited by the queue size specified (1024 in this case). Figure 10 plots the temperature measured by our thermal sensors in each of the 34 drives in our experimental setup. As can be seen from the figure, different workload intensities do not impact the drive temperature at each drive. The reason for this can be attributed to the fact that the spindle motor contributes to a significant portion of the power consumption of the disk drive [Sankar et al. 2008] and as long as there is any activity on the VCM that moves the read/write heads, the intensity of the operation does not have an impact on temperature.

5.1.2 Impact of workload patterns on disk temperature

In order to identify whether there is a difference between workload access patters we run our suite of different workloads on the experimental system. We do not change the RAID 1+0 partitions to maintain uniform infrastructure for all our experiments. Figure 11 shows the drive temperature for the different workloads. As we can see from the chart, there is no significant change in temperature experienced by the disk drives running different workloads. The maximum difference is a delta of 3C between Email and Messenger (OLTP) workloads. OLTP workloads have a higher read:write ratio and has a higher inter-arrival time compared to Email. They are also largely random and hence has a slightly higher seek activity that can result in the minor difference between temperatures.

In this section, we saw that workload intensities or variations in actual profiles do not cause significant changes in temperature behavior at the disk drive. This agrees

with our observations from our datacenter study that shows low correlation between workload behavior and temperature of the hard disk drives. Given this low correlation at the disk drives, we believe that investing in workload modulation to control temperature at disk drives yield low benefit with respect to reducing temperature or increasing reliability. Compared to CPU temperature control using DVFS schemes or t-states in processors [Govindan et al. 2009], workload modulation achieves lower reduction in disk temperatures.

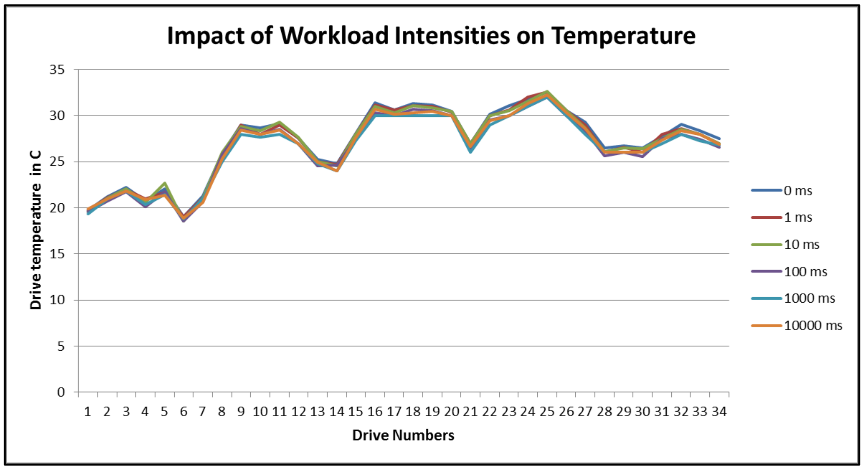


Figure 10: Temperature of 34 drives at different workload intensities

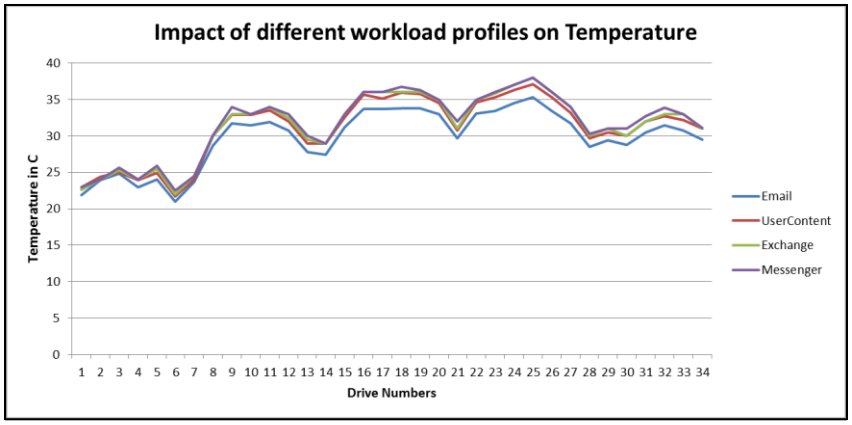


Figure 11: Impact of running different workload profiles on disk temperature

5.2 Chassis Knobs

Server chassis is composed of several components including the sheet metal casing that includes all the individual components like the CPU, motherboard, power supply, fans, memory and the hard disk drives, in addition to all the cables connecting the different components. The layout of each component on the chassis is deliberated and positioned in a way that optimizes the floor plan, signal integrity and cost of the overall solution. The thermal behavior of each component in the server system is impacted by the position of the system relative to inlet temperature at the cold aisle and also the cooling solution employed. CPUs have heat sinks that absorb heat produced from the processor. Hard disks however do not contain heat sinks in

the typical enterprise scenario; however they are cooled by chassis level fans that move air through the chassis. The pressure difference maintained across the chassis by the rotating fans results in air flow that removes heat from the system. Typically, the components closer to the cold aisle have a lower temperature, and due to the preheating effect and the direction of air flow, the components at the back of the chassis have higher temperature. In Section 4.2.1 we saw the impact of difference in temperature across the chassis impact disk failures differently. In this section, we measure the impact of control knobs that can impact temperature differences across the chassis, including disk placement and fan speeds, on the temperature experienced by disk drives.

5.2.1 Impact of disk placements inside the chassis

In our experimental system, there are a total of 35 disk bays where disk drives could be connected. However, there is a requirement for only 34 disk drives in our system. We use the one available open slot to experiment the impact of disk placement on temperature experienced by the disk drives. We use the column positions 1 till 5 to place the empty slot in the middle of the chassis (refer to Section 3.2). Figure 12 shows the impact of an empty slot in the system. We denote the temperature of the empty slot to be 0 in the chart. Note that whenever there is a sharp dip in the series, after 7 consecutive positions, there is another small dip in temperatures. Since there are 7 disk drives arranged in each column, the second dip in each series corresponds to the disk drive directly behind the empty slot. This experiment shows that the position of hard disk drives and empty slot influence air flow and can result in reducing temperature in storage enclosures. We see that an empty slot can reduce the temperature experienced by the disk drive behind the empty slot by close to 5C-7C. Hence based on the requirement of enterprise applications, it might be beneficial to allow empty slots with the purpose of cooling disk drives that experience a higher temperature.

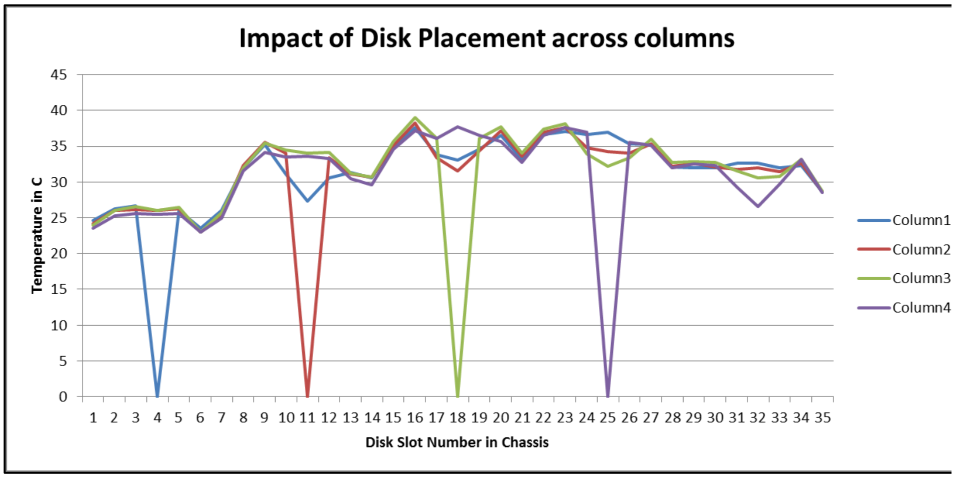


Figure 12: Empty slots with temperature of zero creates a dip in temperature of the drive directly behind the slot (after 7 places)

5.2.2 Impact of fan speeds on disk temperature

Fans are the most common solution used in servers for moving cold air across the server chassis to cool hard disk drives. In this section, we measure the impact of different fan RPMs on the temperature of disk drives in our experimental setup. An increase in fan RPM results in increase in power consumed by the fans since power is proportional to the cube of the RPM. Hence we need to evaluate the benefit of reducing temperature on reliability compared to the cost of increased power for increasing fan RPM. Figure 13 shows the relationship between fan speeds and temperature. In our setup, we can control the fan speed RPM from 7000 RPM (denoted by 7000-wkld) to 12000 RPM (denoted by 12000-wkld) and we increase the fan RPM in steps of 1000. We see a drop in temperature of 5C when we increase fan speed from 7000 RPM to 12000 RPM.

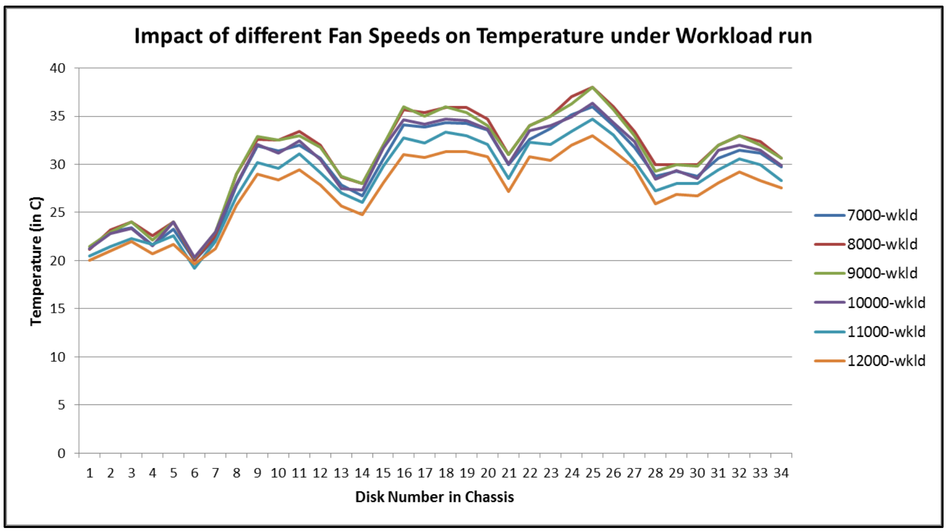


Figure 13: Increasing fan speeds reduces temperature of disk drives

6. MODEL FOR HARD DISK RELIABILITY

From our real datacenter study and experimental evaluation, we identify that average temperature has stronger correlation to disk failures. In order to quantify the impact of different datacenter inlet temperatures experienced by the servers, we needed to come up with a model for measuring the reliability of the hard disk drives that are the primary failure components in the system. We used a physical Arrhenius model and estimated the activation energy based on the failures from the field. Earlier studies have estimated duty cycle has a negative effect on AFR (higher duty cycles have higher accelerated failures) [Cole et al. 2000]. We factor in the effect of duty cycle in the proportional multiplier for Arrhenius model in the next section. Using this model, we estimate the AFR (Annualized Failure Rate) and consider that to be a baseline for comparison between different datacenter inlet temperature decisions.

6.1 Arrhenius model for Acceleration Factor

The failure rate due to elevated temperature is governed by the Arrhenius equation [Cole et al. 2000]. The Arrhenius acceleration factor (AF) can be expressed as:

$$AF = A * e^{\left(-\frac{E_a}{kT}\right)}$$

Where,

A, is a proportional multiplier
Ea, is the activation energy determined empirically
K, is the Boltzmann's constant that relates energy at the particle level with temperature observed at macro level
T, is the absolute temperature observed at elevated temperature points respectively

Acceleration Factor (AF) can also be expressed as the ratio between the time it took to fail under normal temperature versus the elevated temperature. Rewriting above equation,

$$Ln({}^{t2}/_{t1}) = \left(\frac{E_a}{k}\right) * \left({}^{1}/_{T_2} - {}^{1}/_{T_1}\right)$$

Where, t2 is the time for failure with elevated temperature and t1 is the time to failure with normal temperature.

Activation energy E_a, can be calculated from the above equation. We know empirically from Section 4 that we had almost twice the number of failures with 12 C increase in temperature. Substituting the values in the equation, we get **E_a = 0.464 eV.** We estimate the proportional multiplier (A) for the Arrhenius Acceleration Factor equation to be 1.25 based on workload duty cycle expectations. This multiplier is calculated as a function of the duty cycle expected and the duty cycle rated by the manufacturer (similar to [Cole et al. 2000]). We base our calculations on the worst case duty cycle for the workload (100%). We use the above empirically calculated value to compose the Arrhenius model for estimating Acceleration Factor at different temperatures.

Table 2 shows the increase in Acceleration Factor and the corresponding impact on reliability (AFR). We use the 40C row as the baseline temperature and AFR value since it is derived from typical HDD manufacturer data sheets (eg: [Seagate ES 2011]). We see that operating the hard disk drive at 55C increases the AFR by almost twice when compared to the AFR quoted by manufacturers at 40C. The table provides a handy reference sheet for expected failures when the hard disk drive experiences a particular temperature. Given a chassis design, it is straightforward to compute the delta T observed by the hard disk drives at different location inside the chassis. We computed the delta T from SMART logs and when running a constant workload at specific temperatures. We can then use Table 2 to estimate the failure rate for the particular chassis design, given the corresponding datacenter inlet temperature. Thus, this provides a methodology for selecting datacenter setpoint based on expected reliability.

HDD Temp	Acc Factor (AF)	AFR	AFR relative to 40 C
40 C	1.257	2.75	100%
41 C	1.328	2.91	106%
42 C	1.402	3.07	112%
43 C	1.480	3.24	118%
44 C	1.562	3.42	124%
45 C	1.648	3.61	131%
46 C	1.737	3.80	138%
47 C	1.831	4.01	146%
48 C	1.930	4.23	153%
49 C	2.033	4.45	162%
50 C	2.141	4.69	170%
51 C	2.254	4.94	179%
52 C	2.372	5.19	189%
53 C	2.495	5.46	198%
54 C	2.625	5.75	209%
55 C	2.759	6.04	219%

Table 2. HDD temperature and corresponding AFR (40C is baseline)

6.2 Application to Datacenter Setpoint Selection

This section discusses the application of Table 2 in selecting the datacenter setpoint temperature. The setpoint temperature determines the chilled water temperature. The lower the setpoint temperature required, higher the energy required by the chiller units to bring down the temperature. Hence, fixing an optimal setpoint temperature by a data-driven reliability-aware approach would lead to energy conservation and better efficiency at the datacenter.

Table 3 presents two server chassis design. One design contains the HDDs in the front, and therefore is exposed to the cold aisle. The delta T between the temperatures experienced by the front HDDs and the datacenter setpoint temperature is minimal (1 C). The other server design consists of the inner HDDs, which has HDDs arranged one behind the other. We present only the case of the worst HDD in the design. Because of preheating, the delta T in cold temperatures is 20C. However as the air gets hotter, the chassis fans will be sped up to prevent the HDDs from overheating with a delta T of 10C. For temperatures in-between, the delta T will be assumed to be linear. Hence at inlet of 50C, the hottest drive experiences a temperature of 60C. We assume that for temperatures below 40C there is no AFR increase and we keep that as baseline AFR and compute the relative AFR from that data point.

Inlet Temp		HDD's in Front, ΔT 1°C		Buried HDDs Design, ΔT 20°C cold de-rated to ΔT 10°C hot	
		HDD Case Temp	Relative AFR	HDD Case Temp	Relative AFR
10 C	50 F	11 C	100%	30 C	100%
15 C	59 F	16 C	100%	34 C	100%
20 C	68 F	21 C	100%	38 C	100%
25 C	77 F	26 C	100%	41 C	106%
30 C	86 F	31 C	100%	45 C	131%
35 C	95 F	36 C	100%	49 C	153%
40 C	104 F	41 C	106%	53 C	189%
45 C	113 F	46 C	138%	56 C	231%
50 C	122 F	51 C	179%	60 C	281%

Table 3. Choosing Datacenter Setpoint for a) HDDs in Front, b) Buried HDDs

As we can observe from the table, a front facing hard disk drive design experiences fewer failure events at 50C inlet temperature. However, the buried HDD design has significant increase in the relative AFR of the disk drives. Hence we need to make the decision about housing the second design in a datacenter more carefully. If the threshold for disk failures can be fixed, (say at 1.05X the advertised AFR rates, a 5% increase over baseline), then we need to adjust the datacenter setpoint inlet temperature for a datacenter having the second design at 25C. However, if all our servers had the first design, then the setpoint temperature could be 40C. The 15C delta between these two setpoints is a significant temperature delta to operate a datacenter. A 15C difference in setpoint temperature is close to 150KW difference on the datacenter floor. Hence it is useful to have such a methodology in place for setting datacenter setpoint temperature.

Observation: ***Datacenter setpoint temperature should be selected in a reliability-aware manner to avoid potential increases in server failures due to temperature impacts.***

7. COST ANALYSIS OF TEMPERATURE OPTIMIZATIONS

In the preceding sections we saw different temperature optimizations that control disk temperature. In order to quantify the cost of different optimizations, we use available power costs from [Hamilton 2008] and publicly available sources to compare the cost of optimizations. On the other hand, we also evaluate the cost of increased failures by using the Arrhenius model to predict failure increase with temperature. The cost calculations are a measurement of the relative differences between different options. We use publicly available sources for estimating the cost, and this should not be viewed as the actual cost in a typical datacenter. A different deployment would have a different cost calculation, specific to that deployment.

In this section, we consider two optimizations: 1. Cost of fan speed increase and 2. Cost of datacenter chiller costs. In order to estimate the cost of fan speed increase, we identify the total power increase experimentally. The increase from 7000 RPM to 12000 RPM increases the power of the system by 137 watts and reduces temperature by 5C. We need to calculate both the power cost and the cost of increased failures to see which of the cost we should incur. For the cost analysis, we assume a typical power cost of 0.10$ per Kilowatt-hour, and a constant number of servers in the datacenter (we assume 10000 servers each rated at 1000 Watts to contribute to an overall power capacity of 10 Megawatts. 10 Megawatts is a typical datacenter size for large enterprises). To compute the increase in power cost alone,

Power cost = number of servers * increase in power * power cost (adjusted to 1 year) = **1.5 Million/ year.**

We use the Arrhenius equation to determine the difference between 5C decrease in temperature. We calculate the difference in acceleration factor and estimate the number of failing disks in the population. We assume that the average operating temperature was 45C and a 5C decrease in temperature resulted in a 40C operation. We plug these temperature values in the Arrhenius equation, and show that Acceleration Factor (AF) for 45C = 1.648. Compared to the AF at 40C (1.257), we see a 31% increase in Acceleration factor and hence the AFR% also increases by 31%. The average AFR quoted by disk drive vendors for enterprise class disk drives is

close to 3% of the population. Assuming this value, we estimate the AFR at 45C to be 3.93%. Applying to the population of 10000 servers with 34 disk drives each, we expect an extra 3162 drives to fail every year (0.93% of overall population). The cost of replacing 3162 drives that are under warranty is minimal; however datacenter environments do not return disk drives to manufacturers to protect sensitive information within the disk drives. Instead, they shred the disk drives to protect data. Hence, the cost of total replacement is $632,400 (at 200$ per drive). To this number, we add additional service cost of replacement (we assume 15% of the disk cost to be the cost of service for each replacement – note that we do not have typical numbers for this service, since it is negotiated differently by each vendor for specific use case [Vishwanath et al. 2010]). The total cost of service then becomes $94860. The total cost to the datacenter operator for the increased failures is $727,260.

The power cost that a datacenter operator would pay for a year to increase the fan speeds ($1.5 Million) is almost twice that of the cost of increased failures ($0.73 Million). However note that we do not include the cost of service downtime. For datacenter operators, service downtime is a critical metric, and to obtain a 1% increase in that metric, they might be willing to incur this extra cost.

Another optimization that can be used at the datacenter level is to change the datacenter setpoint temperature. This also lowers the temperature of the entire server chassis in addition to lowering the temperature for the entire datacenter. The power consumption of the datacenter increases, and hence there is power cost associated with this knob. We compare the cost for the additional power required to increase the datacenter setpoint temperature. A decrease of 5C incurs additional power usage of 60 Kilowatt for a typical datacenter facility. This power is spent in reducing the datacenter chilled water temperature, and to maintain the temperature at 5C below the earlier operating temperature. Our observations correlate with a study conducted by [Namek et al. 2011] where they consider a 4000 ton chilled water plant serving a 100,000 square feet data center at 150 W/sf. According to their results, they estimate a total power consumption of 695 Megawatt-hour annually for 2F decrease in temperature. For a 5 C decrease, we can compute the resulting power usage from their calculations (estimated at 3129 Megawatt-hour resulting in a power cost of $312,900). This cost is lesser when compared with the cost of increase in failures computed earlier ($0.73 Million).

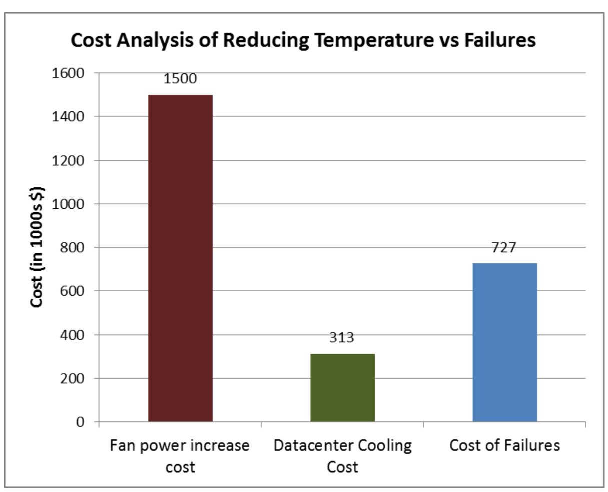


Figure 14: Cost comparison between reducing temperature and increase in failures

From Figure 14, we can see that the cost of datacenter cooling increase is significantly lower than the cost of server fan power increase, and is also lower than the cost of increased failures. In this case, datacenter setpoint temperature reduction is recommended to decrease failures. However note that running chiller plants at lower temperatures reduces their efficiency since the temperature delta increases between the chilled loop temperature and external temperature [Microsoft 2009] and hence this methodology should be reevaluated for lower temperatures.

Summary:
There is a cost associated with increasing cooling to facilitate lower temperatures at disk drives. Similarly there is a cost associated with reducing cooling to increase datacenter power efficiency, attributed to the cost of the resulting increase in failures. We propose that these costs should be factored in before datacenter design decisions are taken. In our cost analysis we show that certain chassis design knobs and datacenter knobs are better at overall temperature control, and workload knobs do not provide significant benefit to disk drive temperature control.

8. FUTURE WORK

There has been a significant move to efficient cooling mechanisms in datacenters like airside economizers, such as by Microsoft [Microsoft 2009] and more recently by Facebook [Facebook 2011]. The underlying principle behind the cooling mechanism is that outside air is cold enough for a majority of hours during the year to cool servers inside datacenter, and water based chiller units can be removed. During the hotter summer months, these datacenters use adiabatic cooling in addition to free-air cooling [Intel 2008]. This methodology of cooling datacenters causes significant variations in inlet temperature and the datacenter setpoint temperature is not maintained at a constant level like we saw earlier in traditional datacenters. Every server component experiences variations in temperature according to outside temperature, in addition to relative humidity differences based on outside humidity and adiabatic cooling. This presents a completely different set of challenges in terms of quantifying reliability and is subject of future work.

9. CONCLUSION

Server and datacenter reliability are first order constraints that determine profit margins for large enterprises. Previous works on hard disk drive failures and temperature impact are highly variant in their claims and do not evaluate variations in temperature or the inter-relationship between workload, temperature and failures. In this work, we evaluate the impact of temperature on hard disk drive reliability, model real world data on temperature and failures and also focus on the resulting impact on server design and datacenter cost. This work highlights the need for temperature aware server design for increased datacenter efficiency.

REFERENCES

COLE, G. 2000. Estimating drive reliability in desktop computers and consumer electronics systems. *SeagateTechnology Paper TP-338.1*, November 2000
ELERATH, J. G. AND SHAH, S. 2004. "Server class disk drives: How reliable are they?" *In Proceedingsof the Annual Symposium on Reliability and Maintainability*, pages 151 – 156, January 2004.

EL-SAYED, N., STEFANOVICI, I. A., AMVROSIADIS, G., HWANG, A. A., AND SCHROEDER. B., 2012. "Temperature management in data centers: why some (might) like it hot". *In Proceedings of the 12th ACM SIGMETRICS/ PERFORMANCE joint international conference on Measurement and Modeling of Computer Systems* (SIGMETRICS '12). ACM, New York, NY, USA

GOVINDAN, M. S. S., LEFURGY, C., AND DHOLAKIA, A. 2009 Using On-line Power Modeling for Server Power Capping. *In Proceedings of the 2009 Workshop on Energy-Efficient Design (WEED),* June 2009.

GRAY, J. AND VAN INGEN, C. 2005. Empirical measurements of disk failure rates and error rates. *Technical Report MSR-TR-2005-166*, December 2005.

GREENBERG, S., MILLS, E., TSCHUDI, W., RUMSEY, P., AND MYATT, B. 2006. Best practices for data centers: Lessons learned from benchmarking 22 data centers. *ACEEE Summer Study on Energy Efficiency in Buildings*, 2006.

GUO, G. AND ZHANG, J. 2003. Feedforward Control for Reducing Disk-Flutter-Induced Track Misregistration. *IEEE Transactions on Magnetics*, 39(4):2103–2108, July 2003

GURUMURTHI, S., ZHANG, J., SIVASUBRAMANIAM, A., KANDEMIR, M., FRANKE, H., VIJAYKRISHNAN, N., AND IRWIN, M. 2003. 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 (ISPASS*), pages 123–132, March 2003.

GURUMURTHI, S., SIVASUBRAMANIAM, A., AND NATARAJAN, V. 2005 Disk Drive Roadmap from the Thermal Perspective: A Case for Dynamic Thermal Management. *In Proceedings of the International Symposium on Computer Architecture (ISCA*), pages 38–49, June 2005.

HAMILTON, J. 2007. "An Architecture for Modular Data Centers". In *CIDR 2007*

HAMILTON, J. 2008. "Datacenter TCO Model*"http://perspectives.mvdirona.com* , 2008

HOELZLE, U., BARROSO, L. A. 2009. The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines, *Morgan and Claypool Publishers*, 2009

HP White Paper. 2003. "Assessing and Comparing Serial Attached SCSI and Serial ATA Hard Disk Drives and SAS interface", October 2003

HP SSA70 Storage Disk Enclosure, 2011. h18006.www1.hp.com/storage/disk_storage/index.html

Intel Whitepaper, 2008. "Reducing Data Center Cost with an Air Economizer". August 2008.

IOMeter 2011. IOMeter project – www.iometer.org

KIM, Y., GURUMURTHI, S., AND SIVASUBRAMANIAM, A. 2006. Understanding the Performance-Temperature Interactions in Disk I/O of Server Workloads, *In the Proceedings of the International Symposium on High Performance Computer Architecture*, pages 179-189, February 2006.

Microsoft 2009. "Microsoft's Chiller-less Data Center", *Datacenter Knowledge*, Sep 2009.

NAMEK, R.Y., AND FOURNIER, E. 2011 "Two strategies to reduce Chiller power and Plant energy consumption in Data centers", *DatacenterDynamics*, March 2011.

PARK, I., AND BUCH, R. 2007. "Improve Debugging and Performance Tuning with ETW", Microsoft Corporation, April 2007

PATTERSON, M. K. 2008. "The Effect of Data Center Temperature on Energy Efficiency" in *Proceedings of the 11th Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, May 2008, pp. 1167–1174.

PINHEIRO, E., WEBER, W. D., AND BARROSO, L. A. 2007. Failure trends in a large disk drive population. *In Proc. of the FAST '07 Conference on File and Storage Technologies*, 2007

Seagate ES 2011. Seagate Constellation ES drive datasheet.

SANKAR, S., GURUMURTHI, S., AND STAN, M.R. 2008. Intra-Disk Parallelism: An Idea Whose Time Has Come, *Proceedings of the International Symposium on Computer Architecture*, June 2008.

SCHROEDER, B., AND GIBSON, G. 2006. "A large scale study of failures in high-performance-computing systems." *International Symposium on Dependable Systems and Networks (DSN 2006)*.

SCHROEDER, B., AND GIBSON, G. 2007. "Disk failures in the real world: what does an MTTF of 1,000,000 hours mean to you?", *Proceedings of the 5th USENIX conference on File and Storage Technologies*, February 13-16, 2007, San Jose, CA

SCHROEDER, B., PINHEIRO, E., AND WEBER, W. 2009. DRAM errors in the wild: a large-scale field study, *Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems*, June 15-19, 2009, Seattle, WA, USA

SCHWARTZ, T., BAKER, M., BASSI, S., BAUMGART, B., FLAGG, W., VAN INGEN, C., JOSTE, K., MANASSE, M., AND SHAH, M. 2006. Disk failure investigations at the internet archive. 14th NASA Goddard, *23rd IEEE Conference on Mass Storage Systems and Technologies*, May 2006.

Facebook 2011. Open compute project at Facebook http://opencompute.org/

VISHWANATH, K. V. AND NAGAPPAN, N. 2010. Characterizing cloud computing hardware reliability. *In Proceedings of the 1st ACM Symposium on Cloud Computing* (Indianapolis, Indiana, USA, June 10 - 11, 2010). SoCC '10.

YANG, J., AND SUN, F. 1999. A comprehensive review of hard-disk drive reliability. *In Proceedings of the Annual Symposium on Reliability and Maintainability*, pages 403 - 409, January 1999