Technical Report CSRG-615

Temperature Management in Data Centers: Why Some (Might) Like It Hot

Nosayba El-Sayed Ioan Stefanovici George Amvrosiadis Andy A. Hwang
Bianca Schroeder
{nosayba, ioan, gamvrosi, hwang, bianca}@cs.toronto.edu
Department of Computer Science
University of Toronto

ABSTRACT

The energy consumed by data centers is starting to make up a significant fraction of the world's energy consumption and carbon emissions. A large fraction of the consumed energy is spent on data center cooling, which has motivated a large body of work on temperature management in data centers. Interestingly, a key aspect of temperature management has not been well understood: controlling the setpoint temperature at which to run a data center's cooling system. Most data centers set their thermostat based on (conservative) suggestions by manufacturers, as there is limited understanding of how higher temperatures will affect the system. At the same time, studies suggest that increasing the temperature setpoint by just one degree could save 2–5% of the energy consumption.

This paper provides a multi-faceted study of temperature management in data centers. We use a large collection of field data from different production environments to study the impact of temperature on hardware reliability, including the reliability of the storage subsystem, the memory subsystem and server reliability as a whole. We also use an experimental testbed based on a thermal chamber and a large array of benchmarks to study two other potential issues with higher data center temperatures: the effect on server performance and power. Based on our findings, we make recommendations for temperature management in data centers, that create the potential for saving energy, while limiting negative effects on system reliability and performance.

Categories and Subject Descriptors

B.8 [**Hardware**]: Performance and Reliability—*Temperature*; C.4 [**Computer Systems Organization**]: Performance of Systems—*Temperature*

Keywords

Data Center, Temperature, Reliability, Performance, Energy, LSE, Hard Drive, Memory, DRAM, CPU, Fans

1. INTRODUCTION

Data centers have developed into major energy hogs. The world's data centers are estimated to consume power equivalent to about seventeen 1,000 MW power plants, equaling more than 1% of total world electricity consumption, and to emit as much carbon dioxide as all of Argentina [16]. More than a third, sometimes up to one half of a data center's electricity bill is made up by electricity for cooling [6, 18]. For instance, for a data center consisting of 30,000 square feet and consuming 10MW, the yearly cost of running the cooling infrastructure can reach up to $4-8 million [22].

Not surprisingly, a large body of research has been devoted to reducing cooling cost. Approaches that have been investigated include, for example, methods to minimize air flow inefficiencies [22, 34], load balancing and the incorporation of temperature awareness into workload placement in data centers [7, 24, 27, 32], and power reduction features in individual servers [13, 14].

Interestingly, one key aspect in the thermal management of a data center is still not very well understood: controlling the setpoint temperature at which to run a data center's cooling system. Data centers typically operate in a temperature range between 20C and 22C, some are as cold as 13C degrees [8, 28]. Due to lack of scientific data, these values are often chosen based on equipment manufacturers' (conservative) suggestions. Some estimate that increasing the setpoint temperature by just one degree can reduce energy consumption by 2 to 5 percent [8, 9]. Microsoft reports that raising the temperature by two to four degrees in one of its Silicon Valley data centers saved $250,000 in annual energy costs [28]. Google and Facebook have also been considering increasing the temperature in their data centers [28].

While increasing data center temperatures might seem like an easy way to save energy and reduce carbon emissions, it comes with some concerns, the most obvious being its impact on system reliability. Unfortunately, the details of how increased data center temperatures will affect hardware reliability are not well understood and existing evidence is contradicting. A recent study [34] indicated that in order to avoid thermal redlining, a typical server needs to have the air temperature at its front inlets be in the range of 20C – 30C. Every 10C increase over 21C decreases the reliability of long-term electronics by 50% [23]. Other studies show that a 15C rise increases hard disk drive failure rates by a factor of two [4, 10]. On the other hand, a recent Google study [25] suggests that *lower* temperatures are actually more detrimental to disk reliability than higher temperatures.

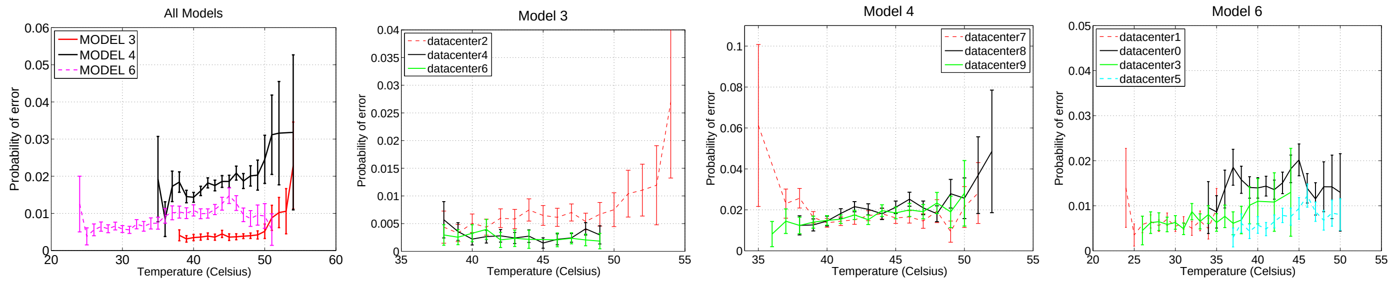

Figure 1: The monthly probability of LSEs as a function of temperature. In the three plots for individual models each line corresponds to the measurement from a different data center.

Other possible concerns of increasing data center temperatures include the effect on server performance, as many servers employ techniques such as CPU or memory throttling when temperatures reach a critical threshold, and the effect on server energy consumption, as increased temperatures will lead to increases in power leakage and higher (server internal) fan speeds.

The goal of this paper is to provide a better understanding of the issues involved in raising data center temperatures. As a first contribution, in Section 2 we perform a detailed study of the effect of temperature on hardware reliability by analyzing a large amount of field data. The data comes from three different organizations spanning several dozen data centers and covers a diverse set of common reliability issues, including hard disk failures, latent sector errors in hard disks, uncorrectable errors in DRAM, DRAM replacements, and general node outages. In Section 3 we perform an experimental study using a testbed based on a thermal chamber and a large set of different workloads to better understand the effects that temperature has on the performance and power usage of systems. Finally, in Section 4 we use the results of our study to derive some insights and guidelines for running data centers at higher temperatures, while limiting the impact on system performance and reliability.

2. TEMPERATURE AND RELIABILITY

We begin our study by analyzing a diverse set of field data collected at different organizations and data centers to better understand the effect of temperature on various aspects of hardware reliability. We first focus on two specific hardware components, hard disks and DRAM, since these are among the most frequently replaced components in modern data centers [29, 30]. In Sections 2.1 and 2.2, we study two common failure modes of hard disks, latent sector errors and complete disk failures, respectively, before moving to DRAM reliability in Section 2.3. Then, in Section 2.4 we use data on node outages in data centers to study the effect of temperature on overall server reliability.

2.1 Temperature and latent sector errors

2.1.1 *Background and data*

Latent sector errors (LSEs) are a common failure mode, where individual sectors on a disk become inaccessible, and the data stored on them is lost (unless the system can use redundancy mechanisms to recover it). LSEs happen at a significant rate in the field [5, 25], with 3-4% of all drives experiencing them at some point in their life, and are expected to grow more common as disk capacities increase. While recent work [5] has studied the prevalence and some statistical properties of LSEs, there is no prior work on how temperature affects this important error condition.

To study the effect of temperature on the prevalence of LSEs, we obtained data collected from January 2007 to May 2009 at 7 different data centers (DCs) at Google covering three different disk models. For each of the disks, we have monthly reports of the average (internal) disk temperature and temperature variance in that month, the count of latent sector errors, the number of read and write operations during that month, and the age of the disk. All data were collected by polling the disks' internal self-monitoring facility (SMART). The measurement infrastructure and methodology Google uses to collect such data are described in Pinheiro et al. [25]. The table below summarizes our data:

Model ID	#DCs	#Disks	#Disk Months	Avg. monthly LSE probability
3	3	18,692	300,000	0.0063
4	3	17,515	300,000	0.0177
6	4	36,671	300,000	0.0067

2.1.2 *Analysis*

Figure 1 (far left) shows for each of the three models the monthly probability of a disk experiencing an LSE as a function of the average temperature. The error bars in this figure (and in all other figures in this work) are computed using a 95% confidence level; larger bars for higher temperatures are due to lack of data. Since there are many data center-specific factors beyond temperature that might affect reliability (workload, humidity, power spikes, handling procedures, etc), we also break down the results for each model by data center. The three rightmost graphs in Figure 1 show the monthly LSE probabilities for the three models, where each line corresponds to a different data center.

As one might expect, we observe a trend of increasing LSE rates as temperature rises. However, the magnitude of increase is much smaller than expected based on common models and estimates, in particular when isolating the instances of LSEs per model per data center. Models for the effect of temperature on hardware components usually assume an exponential increase in failures as a function of temperature (based on the Arrhenius equation [15]), and predict roughly doubling failure rates for every 10-15C increase in temperature [4, 10, 34]. Visual inspection of our graphs shows for only 5 out of the 10 model/data center combinations a clear increase in errors with temperature: model 3, data center 2; model 4, data centers 8 and 9; model 6, data centers 3 and 5. We also observe that the increase in error rates tends to be linear, rather than exponential, except for very high temperatures (above 50C).

To formalize our observation above, we fitted two different

Model	DC	Monthly Probability	Linear fit			Exponential fit		
			a1	a2	SSE	b1	b2	SSE
3	2	$7.99 \cdot 10^{-3}$	$-2.726 \cdot 10^{-2}$	$7.664 \cdot 10^{-4}$	$2.331 \cdot 10^{-4}$	$2.561 \cdot 10^{-1}$	$-1.637 \cdot 10^{+2}$	$2.402 \cdot 10^{-4}$
	4	$2.93 \cdot 10^{-3}$	$7.519 \cdot 10^{-3}$	$-1.055 \cdot 10^{-4}$	$1.157 \cdot 10^{-5}$	$6.613 \cdot 10^{-4}$	$6.192 \cdot 10^{+1}$	$1.112 \cdot 10^{-5}$
	6	$2.51 \cdot 10^{-3}$	$7.322 \cdot 10^{-3}$	$-1.111 \cdot 10^{-4}$	$2.328 \cdot 10^{-6}$	$3.730 \cdot 10^{-4}$	$8.092 \cdot 10^{+1}$	$2.402 \cdot 10^{-6}$
4	7	$2.06 \cdot 10^{-2}$	$6.517 \cdot 10^{-2}$	$-1.025 \cdot 10^{-3}$	$1.720 \cdot 10^{-3}$	$3.624 \cdot 10^{-3}$	$7.054 \cdot 10^{+1}$	$1.595 \cdot 10^{-3}$
	8	$2.28 \cdot 10^{-2}$	$-5.614 \cdot 10^{-2}$	$1.755 \cdot 10^{-3}$	$3.994 \cdot 10^{-4}$	$5.256 \cdot 10^{-1}$	$-1.429 \cdot 10^{+2}$	$3.920 \cdot 10^{-4}$
	9	$1.73 \cdot 10^{-2}$	$-2.346 \cdot 10^{-2}$	$9.482 \cdot 10^{-4}$	$6.192 \cdot 10^{-5}$	$1.955 \cdot 10^{-1}$	$-1.047 \cdot 10^{+2}$	$6.218 \cdot 10^{-5}$
6	0	$1.43 \cdot 10^{-2}$	$8.730 \cdot 10^{-3}$	$1.317 \cdot 10^{-4}$	$1.282 \cdot 10^{-4}$	$2.543 \cdot 10^{-2}$	$-2.481 \cdot 10^{+1}$	$1.275 \cdot 10^{-4}$
	1	$6.67 \cdot 10^{-3}$	$1.067 \cdot 10^{-2}$	$-1.356 \cdot 10^{-4}$	$7.784 \cdot 10^{-5}$	$4.695 \cdot 10^{-3}$	$8.477 \cdot 10^{+0}$	$7.944 \cdot 10^{-5}$
	3	$7.61 \cdot 10^{-3}$	$-4.752 \cdot 10^{-3}$	$3.616 \cdot 10^{-4}$	$2.918 \cdot 10^{-5}$	$3.131 \cdot 10^{-2}$	$-4.863 \cdot 10^{+1}$	$3.235 \cdot 10^{-5}$
	5	$7.10 \cdot 10^{-3}$	$-1.001 \cdot 10^{-2}$	$3.934 \cdot 10^{-4}$	$3.586 \cdot 10^{-5}$	$1.180 \cdot 10^{-1}$	$-1.236 \cdot 10^{+2}$	$3.820 \cdot 10^{-5}$

Table 1: Parameters from fitting linear and exponential models to monthly LSE probabilities as a function of avg. temperature.

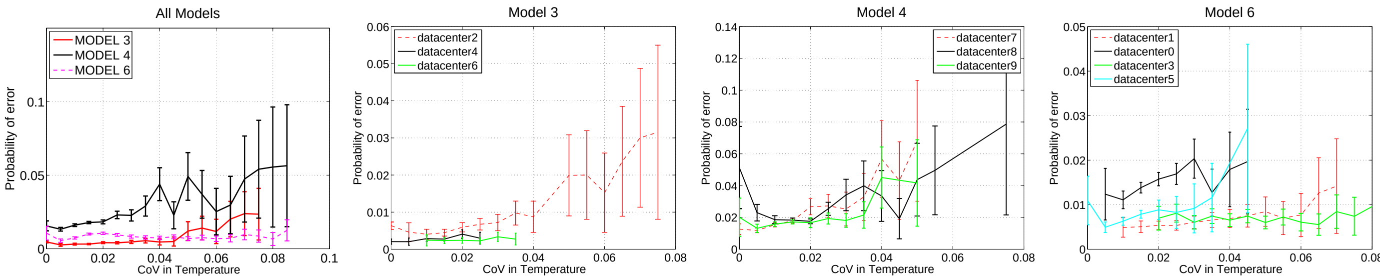

Figure 2: The monthly probability of LSEs as a function of variability in temperature, measured by the coefficient of variation. In the three plots for individual models each line corresponds to the measurement from a different data center.

models to the data. The first is a simple linear model, i.e. we try to model the error rate y as a function of temperature t as $y = a_1 + a_2 \cdot t$. Since one of the most common models for effects of temperature on hardware reliability, the Arrhenius model, is an exponential one, we also fit an exponential model to our data, i.e. we model the failure rate y as a function of temperature t as follows: $y = a_1 \cdot e^{-a_2/t}$. The detailed results (including values for the parameters a_1, a_2, b_1, b_2, and the corresponding sum of squared errors (SSE)) are presented in Table 1. We find that in all cases the linear model provides a fit of comparable or even better accuracy, as measured by the SSE. The only exception is model 3, data center 2, where the exponential model provides a better fit. We attribute this to the sudden increase in LSEs for temperatures above 50C. When repeating our analysis for only data points below 50C, also for model 3, data center 2, the linear model provides a better fit.

Observation 1: For the temperature range that our data covers with statistical significance ($< 50C$), the prevalence of latent sector errors increases much more slowly with temperature, than reliability models suggest. Half of our model/data center pairs show no evidence of an increase, while for the others the increase is linear rather than exponential.

In addition to comparing the quality of the linear versus the exponential fit, it is interesting to look at the slope of the linear increase in errors (parameter a_2), i.e. the rate at which errors increase. One interpretation of a_2 is that it gives the additional fraction of drives that will develop LSEs for each 1 degree increase in temperature, e.g. $a_2 = 0.01$ means that for a 1 degree increase in temperature an additional 1% of the drive population in a data center would develop LSEs in a given month (that would not have had LSEs otherwise). We find that for 4 of the 10 model/data center combinations a_2 actually has a small negative value, indicating a small decrease in error rates with temperature. For the remaining positive values, it is important to put the value of a_2 in relation to the average probability of a drive developing an LSE (provided in the third column in Table 1). Studying the values of a_2 for those cases where it is positive, we see that a_2 is always at least an order of magnitude smaller than the average LSE probability for that model/data center combination. That means the fraction of drives in the population that will develop LSEs due to a one degree increase in temperature, will be an order of magnitude smaller than the average observed in the dataset. However, an increase in the range of ten degrees or more in data center temperature would probably warrant some extra measures to protect against data loss due to LSEs.

In addition to the average temperature that a drive is exposed to, another important factor is the variability in temperature, since large variations in temperature can negatively affect IT equipment. To study the impact of temperature variability on LSEs we plot the monthly LSE probabilities as a function of the coefficient of variation (CoV) [1] (see Figure 2). We chose the CoV, rather than variance or standard deviation, since it is normalized by the mean. A positive correlation between LSEs and temperature variance could just be due to the positive correlation between LSEs and mean temperature. Figure 2 shows a clear increase in LSE probabilities with increasing CoV for all models. We verify those visual trends by fitting a linear model to capture the relationship between LSEs and the CoV, and find a positive slope (a_2) for all model/data center pairs.

Observation 2: The variability in temperature tends to have a more pronounced and consistent effect on LSE rates than mere average temperature.

Our analysis so far has exclusively focused on the probability of a drive developing LSEs. Another interesting question is whether higher temperature leads to a higher number of LSEs, once a drive starts developing LSEs. To answer

[1] Recall that the coefficient of variation is defined as the standard deviation divided by the mean.

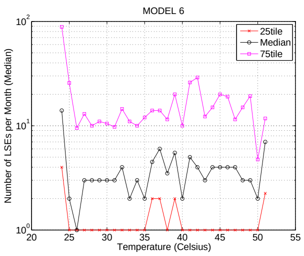

Figure 3: The quartiles of number of LSEs for drives with LSEs as a function of temperature.

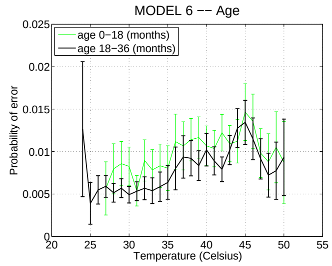

Figure 4: The monthly probability of LSEs as a function of temperature by drive age.

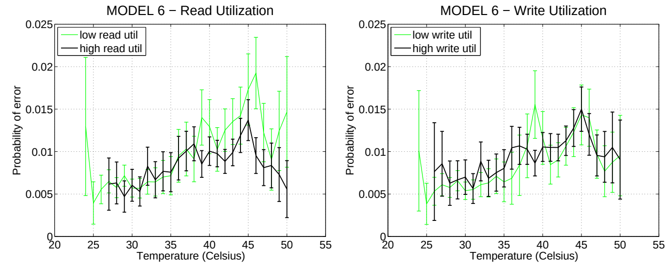

Figure 5: The monthly probability of LSEs as a function of temperature for drives with high and low write loads (right) and read loads (left).

this question Figure 3 plots for those disk months that have errors the 25th and 75th percentile, and the mean. (We only include results for model 6, all others have comparable trends). We focus on the quartiles, rather than the mean, since we find the mean number of LSEs to be highly variable and hence easily biased by outliers. We observe that the line for all quartiles is flat, indicating that hotter drives with errors do not experience a higher frequency of errors than colder drives with errors.

Observation 3: Higher temperatures do not increase the expected number of LSEs once a drive develops LSEs, possibly indicating that the mechanisms that cause LSEs are the same under high or low temperatures.

Figure 1 provides another interesting observation: The rate of LSEs for the same model can vary greatly across data centers. For example, model 3's error rate is significantly higher (more than 2x difference) for data center 2 than for the other data centers, and model 6's error rates are significantly higher for data center 0 than for other data centers (again, more than 2x difference). This brings up the question whether factors, such as environmental conditions or the age or usage of a drive affect how it reacts to temperature. While we have no data on environmental factors, such as humidity or the quality of the power, we have information on the age of each drive and its utilization and study the effect of those factors in Figures 4 and 5.

Our study of age and temperature in Figure 4 focuses on model 6, since the disks for this model span the widest range in age. We divide the drives into two groups, those that are less than 18 months old and those that are 18-36 months old, and plot LSE probabilities as a function of temperature separately for each group. We find that both lines show similar trends with no evidence that older drives are more sensitive to higher temperatures.

Observation 4: Within a range of 0-36 months, older drives are not more likely to develop LSEs under temperature than younger drives.

Figure 5 studies the effect of workload intensity. Figure 5 (left) divides disks into two groups, one with high read utilization and one with low read utilization, and plots the LSE probabilities separately for the two groups. We measure read utilization by the number of read operations per month and assign a disk to the low read utilization group if the number of read operations is below the median for the dataset, and to the high read utilization group otherwise. Figure 5 (right) performs the corresponding analysis for write utilization. Results are shown for model 6 only, but trends were similar for other models as well.

We find that drives with higher utilization are not more sensitive to higher temperatures. That is an interesting result beyond the study of temperature effects, as it has been an open question as to how workload intensity affects LSEs. Methods that are intended to protect against data loss due to LSEs, such as running a periodic "scrubber" that reads the entire disk to proactively detect LSEs, place additional load on a system, and a concern is that this added load might increase the rate of LSEs. Our results indicate that such worries are, likely, unfounded.

Observation 5: High utilization does not increase LSE rates under temperatures.

To add statistical rigour to Observations 4 and 5, we performed an ANOVA test. The results indicate no correlation between LSEs and write utilization. There is evidence for a correlation with read utilization and age, however this is due to drives with *lower* read utilization and *lower* age experiencing slightly increased rates of LSEs.

2.2 Temperature and disk failures

2.2.1 Background and data

Hard disk failures include any kind of disk problems that are considered serious enough to replace the disk in question. Hard disk failures are a serious condition since they create the potential for data loss and happen at a significant rate: typically 1-5% of drives in a data center need to be replaced in a given year [25, 30]. The only existing work that includes trends for the effect of temperature on hard disk failures based on field data is the work by Pinheiro et al. [25]. Surprisingly, this work found a strong *drop* in disk failure rates with *increasing* temperature, except for very high temperatures (above 45C). This is in contrast with common reliability models, which estimate disk failure rates to increase exponentially with temperature.

The goal of this section is to revisit the question of how temperature affects disk failure rates. In addition to obtaining a more conclusive answer to this question, we also look at the question from a broader angle, studying the effect of utilization, differences between models and data centers, and the age of a disk. For our study, we have obtained data on disk replacements collected from January 2007 to May 2009 at 19 different data centers (DCs) at Google covering 5 different disk models. For each disk we know the age of the disk, the average temperature and average utilization over

		All temperatures							< 50'C Temperatures						
Model	DC	Monthly Prob. ($\cdot 10^{-3}$)	Linear fit a1 ($\cdot 10^{-3}$)	Linear fit a2 ($\cdot 10^{-4}$)	Linear fit SSE ($\cdot 10^{-6}$)	Exponential fit b1 ($\cdot 10^{-2}$)	Exponential fit b2	Exponential fit SSE ($\cdot 10^{-6}$)	Monthly Prob. ($\cdot 10^{-3}$)	Linear fit a1 ($\cdot 10^{-3}$)	Linear fit a2 ($\cdot 10^{-5}$)	Linear fit SSE ($\cdot 10^{-6}$)	Exponential fit b1 ($\cdot 10^{-4}$)	Exponential fit b2	Exponential fit SSE ($\cdot 10^{-7}$)
1	9	2.82	−6.387	1.958	6.357	3.242	−116.7	6.702	2.47	0.02274	5.315	1.844	45.88	−29.34	18.96
	13	3.79	−7.253	2.273	5.110	4.057	−116.4	5.241	3.34	0.2640	6.499	1.348	79.08	−41.24	13.52
2	3	3.32	−6.602	2.157	12.30	4.500	−123.3	12.69	2.87	−0.02586	6.384	5.376	94.40	−56.07	55.32
	9	3.09	−8.462	2.485	15.56	6.065	−142.5	16.37	2.75	−1.297	8.807	9.901	96.89	−60.98	102.1
4	8	1.07	−0.523	0.3841	1.996	0.7987	−84.98	2.168	1.07	−1.032	5.201	1.421	129.0	−102.4	15.93
	15	1.64	−4.042	1.481	5.488	4.093	−128.8	5.607	1.41	−3.813	14.15	5.399	353.2	−123.9	56.82
6	0	0.625	0.5464	0.025	0.3250	0.076	−7.242	0.3340	0.625	0.5464	0.2496	0.325	7.600	−7.242	3.340
	1	0.869	0.9486	−0.0183	0.9065	0.06928	7.947	0.9194	0.869	0.9486	−0.1833	0.9065	6.928	7.947	9.194
	2	0.919	2.559	−0.455	0.7095	0.0179	54.33	0.8768	0.919	2.559	−4.555	0.7095	1.798	54.33	8.768
	3	1.45	−1.172	0.5886	6.440	0.3750	−45.18	7.123	1.20	2.117	−2.123	0.9326	5.812	30.03	9.105

Table 2: Parameters from fitting a linear and an exponential model to monthly disk failures as a function of avg. temperature.

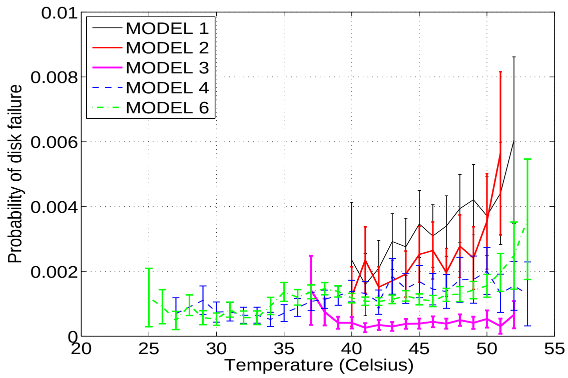

Figure 6: The monthly probability of a disk failure as a function of temperature separated by disk model.

the observation period as reported by the drive's SMART system, and whether the disk was replaced during the observation period. While the time period is different from the study in [25] (there is actually no overlap in time), the measurement methodology and infrastructure used to collect the data is the same as the one Google used in their study.

The following table provides some summary information.

Model	#DCs	#Disks	#Disk Months	Monthly disk fail prob.
1	5	7972	173,945	0.0028
2	4	5906	143,456	0.0023
3	5	93498	752,579	0.0004
4	3	69421	829,859	0.0011
6	5	95226	2,953,123	0.0012

2.2.2 Analysis

Figure 6 plots the monthly failure rate for each of the five models averaged across all data centers. Except for one model (model 3) we observe increasing failure rates with rising temperature. However, we observe that the increase in failures with temperature tends to be linear rather than exponential, except for very high temperatures (above 50C). We validate this observation by fitting a linear and an exponential model to the data, following the same methodology as described in Section 2.1. Results are shown in Table 2. Since the slope of the curves tends to change for very high temperatures, we also repeated the analysis by including only data points below 50C (see right half of Table 2). We find that in all cases the linear model provides a significantly better fit than the exponential model.

As explained in Section 2.1, when studying the rate at which failures increase with temperature (as given by the a_2 parameter) it is important to put the amount of increase in failures, in relation to the average failure rate in a system. When looking at the values for a_2 when fitting the linear model to data points below 50C (see Table 2), we notice that for all model/data center combinations a_2 is by two orders of magnitude smaller than the average failure rate (with the exception of one data point, model 4, data center 15). While average monthly failure rates are typically on the order of 0.1-0.2%, the additional fraction of drives one would expect to fail for each degree increase in temperature is on the order of one thousandth of a percent.

Observation 6: For temperatures below 50C, disk failure rates grow more slowly with temperature than common models predict. The increase tends to be linear rather than exponential, and the expected increase in failure rates for each degree increase in temperature is small compared to the magnitude of existing failure rates.

We also note that, unlike the Google study [25], we do not see a general trend for higher failure rates at lower temperatures. For example, the Google study reports more than a 50% drop in failure rate when moving from 25 to 35C. We believe that the reason is the aggregation of data for different models and data centers in the same curve in [25]. Since different drive models run at different temperatures (due to differences in their design) and different drive models can also vary greatly in their failure rate, it is possible that the data points at the lower end of the temperature spectrum contain more drives of a model that happened to run colder and have higher failure rates, hence biasing the results. Figure 7 shows the distribution of failure counts as a function of temperature for each model, in our data.

As was the case for LSEs, we find that for the same model, the monthly failure probabilities can vary greatly across data centers, even for the same temperature (see Figure 8). This points to other factors, beyond temperature, that have an equally strong or stronger effect on disk lifetimes and motivates us to study two possible factors that we have data on: age and utilization. We followed the same methodology as for LSEs, and divided the drives for each model into those with high and low read utilization, high and low write utilization, and based on age. We found that the behavior of a drive under temperature did not change depending on either utilization or age (with statistically significant data only up to 36 months).

Observation 7: Neither utilization nor the age of a drive significantly affect drive failure rates as a function of temperature.

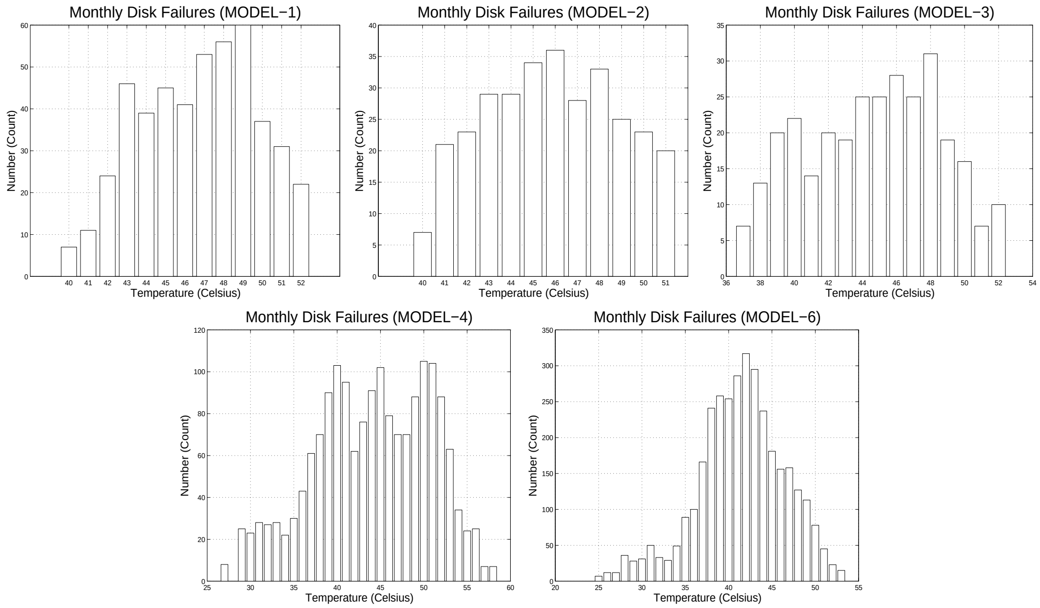

Figure 7: Actual counts of months with disk failures as a function of temperature separated by disk model.

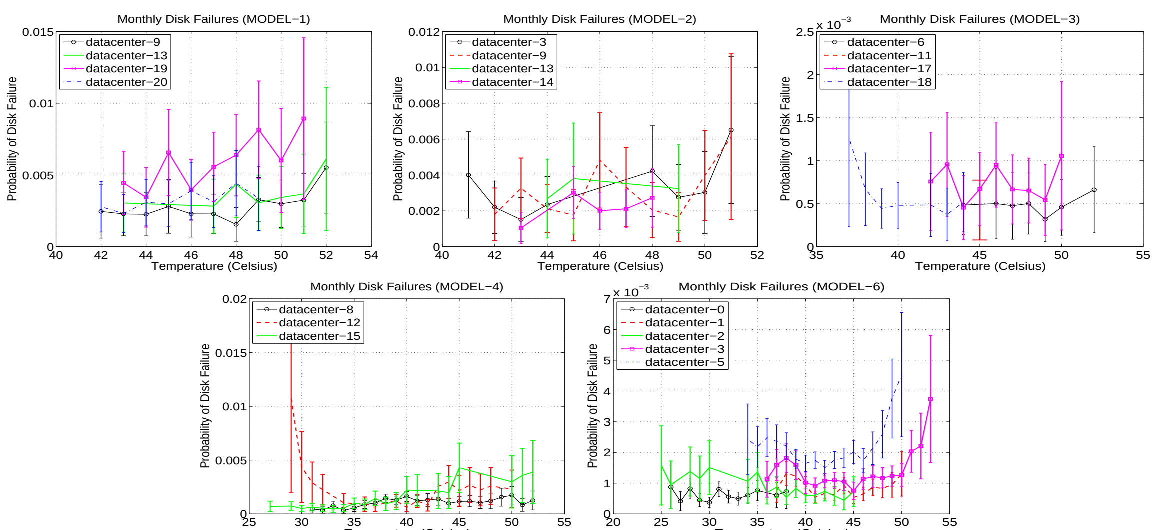

Figure 8: The monthly probability of a disk failure as a function of temperature separated by disk model. Each line in a graph corresponds to a different data center

2.3 Temperature and DRAM reliability

2.3.1 Background and data

In this section, we study how temperature affects the reliability of DRAM, which is one of the most commonly replaced hardware components in data centers and the most common hardware related cause of node outages [29, 30]. DRAM has two different error modes: correctable errors (CEs), where a bit on a DRAM chip is flipped, but can be corrected with internal error correcting codes (ECC); and uncorrectable errors (UEs), where multiple bits are flipped, and the number of erroneous bits is too large for the ECC to correct, causing a machine crash or shutdown. CEs can be caused, by external disturbances, such as cosmic rays, or by hardware defects, such as a stuck bit. UEs usually involve underlying hardware defects, since it is highly unlikely that cosmic rays would simultaneously flip a large enough number of bits to cause an uncorrectable error. Therefore in many data centers it is a common policy to immediately replace a DRAM DIMM after the first occurrence of a UE.

Work in [31] looked at correctable errors in DRAM and showed that their frequency goes up with temperature, but found that this correlation disappears once one controls for utilization. In this section, we ask how temperature affects the long-term reliability of DRAM, rather than the likelihood of transient problems, i.e. do higher temperatures increase the rate at which DRAM wears out and needs to be replaced. We study the long-term reliability of DRAM by analyzing data on DIMM replacements, data on node outages that were attributed to DRAM, and data on uncorrectable errors (since the latter two tend to be indicative of hardware problems and typically lead to replacement of a DIMM). We have collected data from three different sources:

Google: Google routinely collects data on the occurrence of correctable and uncorrectable errors in all of their data centers, as well as periodic temperature measurements based on sensors on the motherboard. An overview of Google's measurement infrastructure is provided in [31]. For our study we have obtained data for a sample set of Google's systems, comprising a dozen different data centers. The data centers are based on five different hardware platforms, where a hardware platform is defined by the motherboard and memory generation. Details on the hardware platforms are considered confidential and we hence just refer to them as Platforms A, B, C, D, E, F.

Los Alamos National Lab (LANL): LANL has made available data on node outages for more than 20 of their high-performance computing clusters, including information on the root cause of an outage and the duration of the outage. The data can be downloaded from LANL's web page [1] and a more detailed description of the data and systems can be found in [29]. Uncorrectable DRAM errors are one of the most common root causes for node outages, and in this section we use only the subset of the data that consists of node outages due to DRAM.

For one of LANL's clusters periodic temperature measurements from a motherboard sensor are available, allowing us to directly study the relationship between temperature and outages. We refer to this system as LANL-system-20, since the ID for this system on LANL's web page is 20. For another 12 clusters information on the data center layout is available, including each node's position in a rack. We use rack position as a proxy for temperature, since due to the cooling system design in those clusters the top of the rack tends to be hotter than the bottom. We have verified that this is the case by analyzing the data for LANL-system-20, where actual temperature measurements are available, and found a difference of 4C between the top and bottom of the rack. The 12 clusters are based on two different hardware platforms, which we refer to as *LANL-Type-1* and *LANL-Type-2*.

LANL-Type-1 comprises seven clusters at LANL totalling 2720 nodes and 20880 processors. The nodes in the system are SMPs with 4 processors per node and are all based on the same hardware platform. The data for these systems spans the years 2002-2008 and corresponds to systems with IDs 3,4,5,6,18,19, and 20 on the LANL web page.

LANL-Type-2 comprises six clusters at LANL totalling 1664 nodes and 3328 processors. The nodes are SMPs with 2 processors per node and the data for these systems spans the years 2003-2008. The data is also available at LANL's web page and corresponds to the systems with IDs 9,10,11,12,13, and 14 on the web page.

SciNet-GPC: The SciNet High Performance Computing Consortium provides computing facilities to researchers in Canada. Their General Purpose Cluster (GPC) is currently the largest supercomputer in the country [2]. We obtained parts replacement data from this system which is manually entered by an administrator when broken hardware is replaced. The replacement log we obtained spans 19 months. The GPC consists of 3870 IBM iDataPlex nodes grouped into 45 racks. Each node contains 2 Intel Xeon E5540 CPUs totaling 8 cores and 16GB of ECC memory.

2.3.2 Analysis

Figures 9 show the monthly probability for node outages at LANL that are attributed to memory as a function of the node's average temperature. In Figure 9 (left) we us the data for LANL-system-20, which has actual temperature measurements, and for Figure 9 (middle,right) we use the server's position in a rack as a proxy for temperature for LANL-Type-1 and LANL-Type-2 systems. We find that none of the graphs shows clear evidence for increasing rate of node outages with increasing temperatures.

Results are similar for hardware replacement rates at SciNet. Figure 10 shows a node's monthly probability of requiring a DIMM replacement as a function of its position in the rack. Again, we see no evidence of higher failure rates for higher (and hence hotter) rack positions.

Unfortunately, due to the size of the datasets the error bars in those graphs are relatively high. We therefore turn to the Google data on uncorrectable errors, which is a larger data set. Figure 11 (left) shows the monthly probability of an uncorrectable DRAM error for the five different hardware platforms at Google. We observe that for two of the models, model C and model F, error rates remain stable throughout the available range of temperature data (which is quite large ranging from 15C to 60C). Maybe surprisingly, model D and model A show contradicting trends, with the former exhibiting decreasing rates as temperature increases and the latter showing increasing rates as temperature rises. To investigate the possible cause we break down the data by data center. Figure 11 (right) shows the resulting breakdown by data center for model D. We find that the error rates for individual data centers are mostly flat with temperature, with

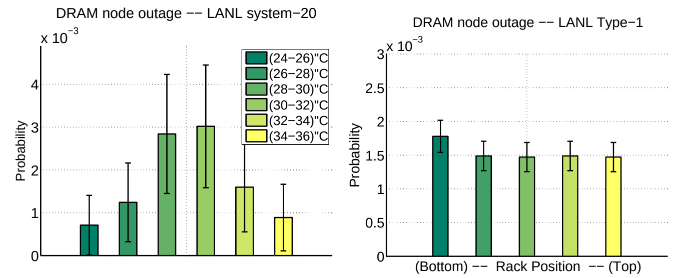

Figure 9: Probability of node outages at LANL due to DRAM problems as a function of temperature (left) and rack positions as a proxy for temperature (middle, right).

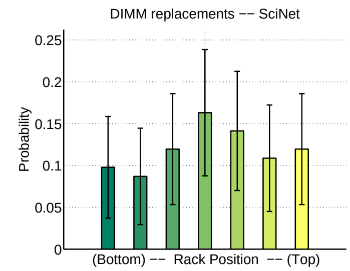

Figure 10: DIMM replacements at SciNet.

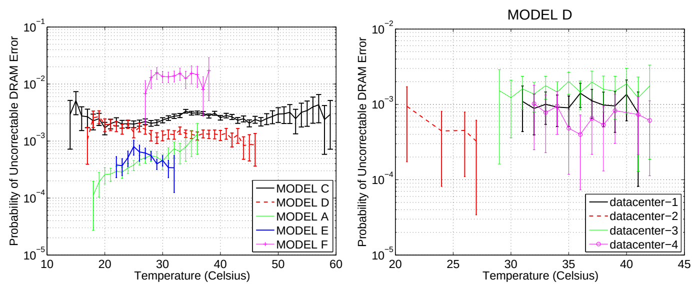

Figure 11: Probability of uncorrectable DRAM errors at Google as a function of temperature.

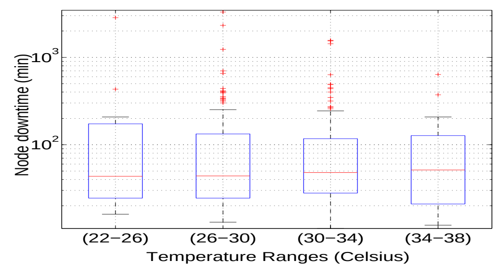

Figure 12: Box plots for per node downtime as a function of temperature.

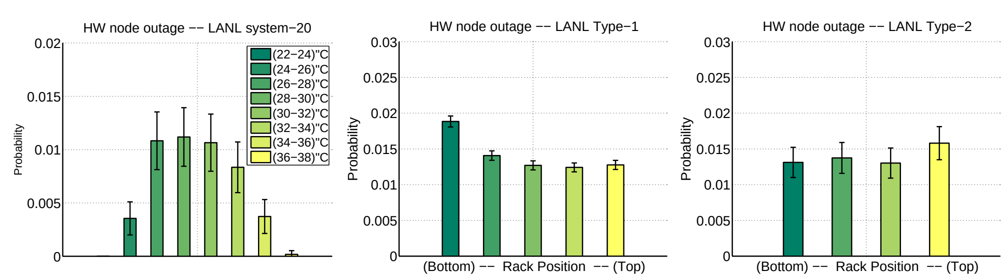

Figure 13: Probability of node outages at LANL as a function of temperature (left) and rack positions as a proxy for temperature (middle, right).

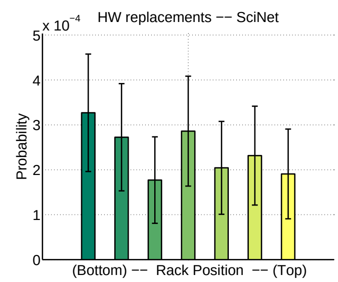

Figure 14: Probability of hardware replacements at SciNet.

the exception of one data center (datacenter-2). It is the aggregation of data from different data centers that creates those apparently contradicting trends. Similarly, we observe for model A that higher temperature points are biased by one data center that is running at a higher temperature and tends to have generally higher error rates (even for low temperatures).

Observation 8: We do not observe evidence for increasing rates of uncorrectable DRAM errors, DRAM DIMM replacements or node outages caused by DRAM problems as a function of temperature (within the range of temperature our data comprises).

2.4 Temperature and node outages

2.4.1 Background and data

Rather than focusing on a particular hardware component, this section looks at overall system reliability and availability as a function of temperature. For our study we use data from two different sources. The first source comprises the LANL datasets LANL-Type-1 and LANL-Type-2. Rather than focusing on records of node outages due to DRAM, we now include in our analysis any node outage that was attributed to a hardware problem. The second dataset is the SciNet-GPC replacement data, but rather than focusing on DRAM replacements we consider replacements of any hardware components.

2.4.2 Analysis

Figure 13 shows the effect of temperature on the rate of node outages at LANL. Figure 13 (left) shows the monthly probability of a node outages as a function of the node's average temperature for system 20 in the LANL data set, as for this system temperature measurements are available. Figure 13 (middle, right) show the monthly probability of a node outages as a function of a node's position in the rack (bottom to top position, i.e. colder to hotter) for LANL-Type-1 and LANL-Type-2 systems. We observe within the temperature range that our data spans no indication that hotter nodes have a higher probability of failing than colder nodes. Results are similar for hardware replacements observed at SciNet (Figure 14): no indication that nodes at

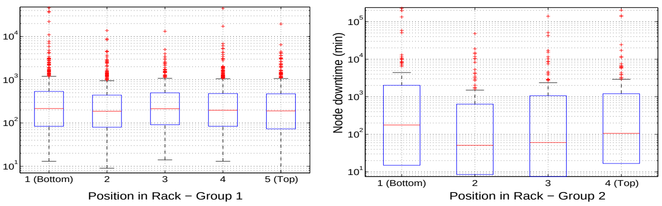

Figure 15: Per node downtime as a function of position of the node within a rack.

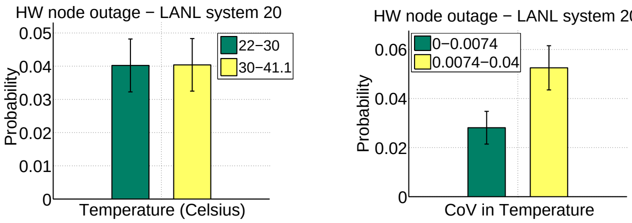

Figure 16: Probability of node outages by temperature (left) and by coefficient of variation (right).

the top of the rack experience more hardware replacements than those at the bottom of the rack.

For the LANL data, we also have information on the length of a node outage, i.e. how long did it take to bring the node back up. Figure 12 shows box plots [2] for the total amount of downtime experienced by a node per month for system 20. (Similar plots for the aggregate of LANL-Type-1 systems and LANL-Type-2 systems can be seen in Figure 15). We find that the downtime experienced by hot nodes does not differ significantly from the downtime experienced by cold nodes, as both medians and lower and upper quartiles of downtime tend to be similar.

Observation 9: We observe no evidence that hotter nodes have a higher rate of node outages, node downtime or hardware replacements than colder nodes.

One might ask whether node outages might be more strongly affected by variability in temperature, rather than average temperature. The only dataset that allows us to study this question is the LANL data for system 20. Figure 16 (right) shows the monthly probability of a node outage for LANL-system-20 as a function of the coefficient of variation in temperature. The figure compares the node outage probability for the top 50% of nodes with the highest CoV and the bottom 50% of nodes with lowest CoV. We observe that nodes with a higher CoV in temperature have significantly increased rates of node outages. For comparison, we also plotted the probability of node outages as a function of average temperature in the same way (Figure 16 (left)) and observe no difference between hot and cold nodes.

Observation 10: We find that high variability in temperature seems to have a stronger effect on node reliability than average temperature.

3. OTHER CONCERNS WITH HIGH TEMPERATURES

Beyond potentially affecting server reliability, there are other concerns with raising data center temperatures. Higher temperatures might affect server performance, increase a server's energy consumption, and lead to smaller safety margins in case of AC or fan failures. We are studying each of these concerns in the remainder of this section.

3.1 Temperature and performance

While it is widely known that higher temperatures might negatively affect the reliability and lifetime of hardware devices, less attention is paid to the fact that high temperatures can also negatively affect the *performance* of systems. For example, in order to protect themselves against a possibly increasing rate of LSEs, some hard disk models enable Read-after-Write (RaW) when a certain temperature threshold is reached. Under RaW, every write to the disk is converted to a Write Verify command, or a Write followed by a Verify operation, reading the sector that has just been written and verifying its contents [35, 36] [3]. Also, when CPU and memory temperatures reach a certain threshold, most advanced servers employ CPU throttling (dynamic voltage frequency scaling) and memory throttling (of the memory bus).

Unfortunately, features such as RaW are often considered trade secrets and are not well documented. In fact, even within a company manufacturing hardware those features and associated parameters are regarded confidential and not shared outside product groups. The goal in this part of our work is to investigate experimentally how performance of different components changes with increasing temperatures.

3.1.1 Experimental setup

To study the performance of a server under increasing ambient temperatures, we set up a testbed using a thermal chamber. The thermal chamber is large enough to fit an entire server inside it, and allows us to exactly control temperature within a range of -10C to 60C with a precision of 0.1C. How ambient temperature affects the temperature of server-internal components depends on many factors, including the design of the cooling system and the server and rack architecture. Therefore, instead of directly predicting the impact of data center ambient temperature on a system, we present our results as a function of the temperature of server internal components.

The server we use in our study is a Dell PowerEdge R710, a model that is commonly used in data center server racks. The server has a quad-core 2.26 GHz Intel Xeon 5520 with 8MB L3, with 16GB DDR3 ECC memory, running Ubuntu 10.04 Server with the 2.6.32-28-server Linux kernel. We also equipped the server with a large variety of different hard disk drives, including both SAS and SATA drives and covering all major manufacturers:

Manufacturer	Model	Interface	Capacity	RPM
Hitachi	Deskstar	SATA	750GB	7200
Western Digital	Caviar	SATA	160GB	7200
Seagate	Barracuda	SATA	1TB	7200
Seagate	Constellation	SAS	500GB	7200
Seagate	Cheetah	SAS	73GB	15000
Fujitsu	MAX3073RC	SAS	73GB	15000
Hitachi	Ultrastar	SAS	300GB	15000

[2] Recall that in a box plot the bottom and top of the box are always the 25th and 75th percentile, respectively, and the band near the middle of the box is always the 50th percentile (the median).

[3] Note that Write Verify is not specified in the ATA standard, which might explain the absence of a performance hit for most SATA drives, in the following subsections.

We use a wide range of workloads in our experiments, including a set of synthetic microbenchmarks designed to stress different parts of the system, and a set of macrobenchmarks aiming to model a number of real world applications:

STREAM: A microbenchmark measuring bandwidth of sequential memory accesses [19]. We used an implementation from the lmbench suite [20, 33] and benchmarked the performance of accessing 4GB of memory.

GUPS: Microbenchmark that measures memory random accesses, in giga-updates-per-second, as defined by the High Performance Computing Challenge [26]. We tested the performance of 8KB-chunk updates randomly to 4GB of memory.

Dhrystone: A well-known microbenchmark that evaluates the CPU performance for integer operations [39].

Whetstone: A well-known CPU benchmark for floating-point performance [11]. Our implementations of *Dhrystone* and *Whetstone* were obtained from the Unixbench suite [21].

Random-Read/Write: A synthetic workload comprised of independent 64KB read (or write) requests issued back-to-back at random disk sectors.

Sequential-Read/Write: Since a pure sequential workload would stress the on-disk cache, we opt for a synthetic workload with a high degree of sequentiality, instead. We pick a random disk sector, and issue back-to-back 64KB read (or write) requests on consecutive sectors for 8MB following the initial request.

OLTP-Mem: We configured TPC-C [37], a commonly used database benchmark modeling on-line transaction processing (OLTP), with 30 warehouses resulting in a 3GB *memory-resident* database.

OLTP-Disk: Models the same workload as *OLTP-Mem*. To make the workload I/O-bound, we configured the database with 70 warehouses (7GB), using 4GB RAM.

DSS-Mem: We configured TPC-H [38], a commonly used database benchmark modeling decision support workloads (DSS), with a 1GB memory-resident MySQL InnoDB database.

DSS-Disk: Another TPC-H based workload, this time configured with a database of 10GB and a 3.4GB buffer pool, resulting in a disk-bound workload.

PostMark [17]: A common file system benchmark, which we configured to generate $50 - 5000$KB files, and modified it to avoid using the OS cache entirely, so that all transactions are directed to disk.

BLAST [3]: An application used by computational biology researchers, acting as a high-performance computing benchmark that stresses both the CPU and memory. We used the parallel *mpiBLAST* implementation [12] and ran 10 representative queries on a 5GB library.

3.1.2 Temperature and disk performance

To study the effect of temperature on disk performance, we ran our disk-bound workloads against each of the drives in our testbed, while placing the drive in the heat chamber and gradually increasing the temperature inside the chamber. The two graphs in Figure 17 show the results for the random-read and random-write microbenchmarks, as a function of the drive internal temperatures, as reported by the drives' SMART statistics. (Results for sequential-read and sequential-write were similar and are omitted for lack of space). We observe that all SAS drives and one SATA drive (the Hitachi Deskstar) experience some drop in throughput

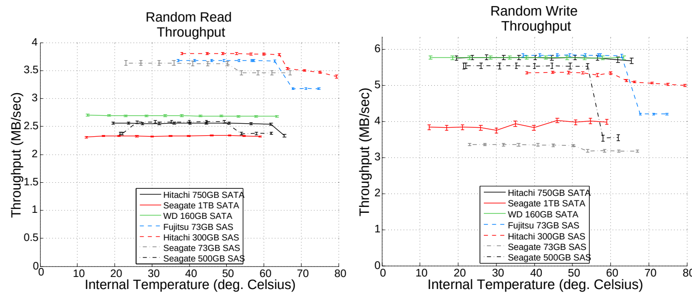

Figure 17: Disk throughput under a synthetic random read and random write workload, respectively, as a function of disk internal temperature. Results for sequential read and sequential write workloads were comparable.

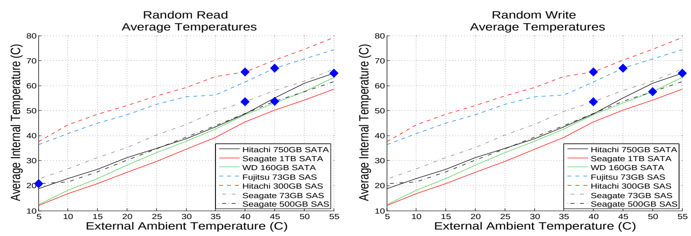

Figure 18: Disk internal temperature as a function of ambient temperature for different drive models and random reads (left) and random writes (right).

for high temperatures. The drop in throughput is usually in the 5-10% range, but can be as high as 30%. Because of the fact that the throughput drop for a drive happens consistently at the same temperature, rather than randomly or gradually, and that none of the drives reported any errors, we speculate that it is due to protective mechanisms enabled by the drive. For example, in the case of the write workloads (which show a more significant drop in throughput) this drop in throughput might be due to the enabling of features such as RaW.

An interesting question is: at what temperature does the throughput start to drop? We observe in Figure 17 drops at either around 50C (for the Seagate SAS drives) or 60C (for the Fujitsu and Hitachi SAS drives). However, these are disk internal temperatures. The two graphs in Figure 18 translate ambient temperatures (inside the heat chamber) to the observed drives' internal temperatures. The markers along the lines mark the points where we observed a drop in throughput. We observe a drop in throughput for temperatures as low as 40C (for the Seagate 73GB and Hitachi SAS drives), 45C for the Fujitsu and Seagate 500GB SAS drives, and 55C for the Hitachi Deskstar, ranges that are significantly lower than the maximum of 50-60C that manufacturers typically rate hard disks for.

While data centers will rarely run at an average inlet temperature of 40C or above, most data centers have hot spots (see Section 3.3), which are significantly hotter than the rest of the data center, and which might routinely reach such temperatures.

Figure 19 shows how temperature affects the throughput of two of our disk-intensive applications, Postmark and OLTP-disk. We observe similar trends as for the microbenchmarks, with throughput drops at the same temperature point. However, the magnitude of lost throughput tends to be big-

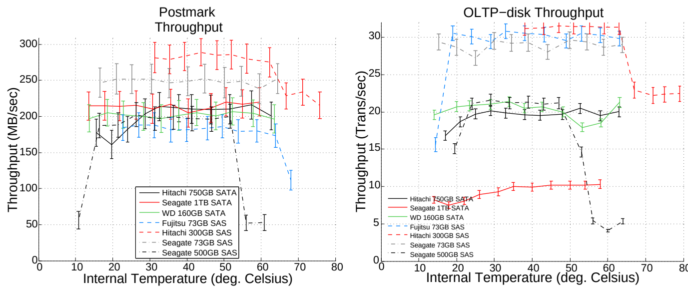

Figure 19: Throughput under two different I/O intensive workloads (Postmark, OLTP-disk) as a function of disk internal temperature.

ger, typically in the 10-20% range, sometimes as high as 40-80%. The drops observed for DSS-disk looked more similar in magnitude to those for the synthetic benchmarks.

3.1.3 *Temperature and CPU/memory performance*

Most enterprise class servers support features to protect the CPU and memory subsystem from damage or excessive errors due to high temperatures. These include scaling of the CPU frequency, reducing the speed of the memory bus, and employing advanced error correcting codes (ECC) for DRAM. For example, our server supports a continuous range of CPU frequencies, bus speeds of either 800MHz or 1066MHz, and three memory protection schemes: single-error-correction and double-error-detection (SEC-DED), advanced ECC (AdvEcc), which allows the detection and correction of multi-bit errors, and mirroring, which provides complete redundancy. Server manuals tend to be purposely vague as to when such features are enabled (CPU and memory bus throttling can be automatically activated by the server), or possible performance impact. In particular, for the memory options it is difficult to predict how they affect performance and power consumption. Since running data centers at higher temperatures might necessitate the use of such features more frequently, we use our testbed to study their impact on performance and power consumption.

For the temperature range we experimented with (heat chamber temperatures up to 55C, significantly higher than the 35C inlet temperature at which most servers are rated) we did not observe any throttling triggered by the server. To study the effect of different memory features, we manually configure the server to run with different combinations of memory bus speed (800MHz vs. 1066MHz) and ECC schemes (SEC-DED, AdvEcc, Mirror). The effect on throughput for the different benchmarks is shown in Figure 20 (left). Throughput is normalized by the maximum attainable throughput, i.e. the throughput achieved when combining a 1066MHz bus speed with the SEC-DED ECC scheme. The results for the two microbenchmarks designed to stress the memory (GUPS and Stream) show that drops in throughput can potentially be huge. Switching to the lower bus speed can lead to a 20% reduction in throughput. The effect of the ECC scheme is even bigger: enabling AdvECC can cost 40% in throughput. The combination of features can cause a drop of more than 50%. For the macrobenchmarks modeling real-world applications the difference in throughput is (not surprisingly) not quite as large, but can reach significant levels at 3–4%. We also measured the server's power consumption (Figure 20 (right)), and

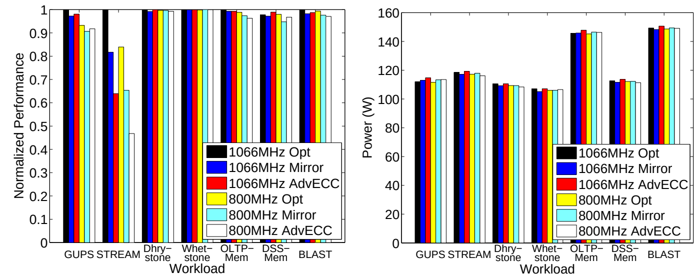

Figure 20: The effect of memory error protection and bus speed on performance (left) and power consumption (right).

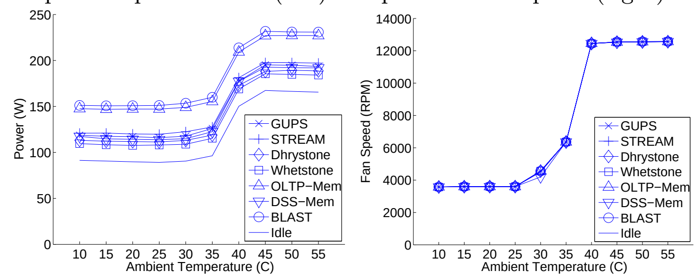

Figure 21: The effect of ambient temperature on power consumption (left) and server fan speeds (right).

found that the impact of memory configurations on server power is small (1-3%) compared to the increases we will observe in the next section during increasing temperatures.

3.2 Increased server energy consumption

Increasing the air intake temperature of IT equipment can have an impact on the equipment's power dissipation. Many IT manufacturers start to increase the speed of internal cooling fans once inlet air temperatures reach a certain threshold to offset the increased ambient air temperature. Also, leakage power of a processor increases with higher temperatures, and can make up a significant fraction of a processor's total power consumption. To study the effect of increasing ambient temperatures on a server's power consumption, we repeated all our earlier experiments with a power meter attached to our server and, in addition, monitored fan speeds.

Figure 21 (left) shows the server's power usage as a function of the ambient (thermal chamber) temperature for the CPU and memory intensive workloads. While the absolute energy used by different workloads varies widely, we observe the same basic trend for all workloads: power consumption stays constant up to 30C and then begins to continually increase, until it levels off at 40C. The increase in power consumption is quite dramatic: up to 50%.

An interesting question is whether this increase in power comes from an increase in fan speed (something that can be controlled by the server) or from increased leakage power (which is governed by physical laws). Unfortunately, it is not possible to measure leakage power directly. Nevertheless, there is strong evidence that the increase in power is dominated by fan power: Figure 21 (right) plots the fan speed as a function of the ambient temperature for all workload experiments. We observe that, the temperature thresholds we notice for which fan speeds increase, line up exactly with the temperatures at which when power consumption increases. We also observe that power consumption levels off once fan speeds level off, while leakage power would continue to grow

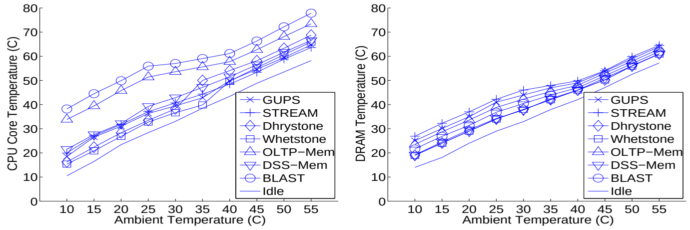

Figure 22: The effect of ambient temperature on CPU temperature (left) and memory temperature (right).

with rising temperatures.

Observation 11: As ambient temperature increases, the resulting increase in power is significant and can be mostly attributed to fan power. In comparison, leakage power is negligible.

Another interesting observation is that power usage starts to increase at the same ambient temperature point for all workloads, although server internal temperatures vary widely across workloads, which means fan speeds increase based on ambient rather than internal temperature. Figure 22 shows the CPU and memory temperature as a function of the ambient temperature for the different workloads and an idle server. We see, for example, that CPU core temperature is more than 20C higher for BLAST and OLTP-Mem than for most other workloads. That means for many workloads the server internal temperatures are still quite low (less than 40C) when the fan speeds start to increase. In particular, we observe that for an idle server, the temperature measured at CPU and memory is still at a very modest 25-30C [4] when the fan speeds start to increase. This is an important observation, since most servers in data centers spend a large fraction of their lives idle.

Observation 12: Smart control of server fan speeds is imperative to run data centers hotter. A significant fraction of the observed increase in power dissipation in our experiments could likely be avoided by more sophisticated algorithms controlling the fan speeds.

3.3 Reduced safety margins

One concern with increasing data center temperatures is that most data centers tend to have hot spots that are significantly hotter than the average temperature in the facility. When raising the temperature setpoint in a data center's cooling system, it is important to also keep in mind how this will affect the hottest part of the system, rather than just the system average. In addition to hot spots, another concern are reduced safety margins: most servers are configured with a critical temperature threshold and will shut down when that threshold is reached, in order to avoid serious equipment damage. As the ambient temperature in a data center increases, equipment will be operating closer to the maximum temperature, reducing the time available to shut down a server cleanly or take protective measures in the case of data center events, such as AC or fan failures.

To better understand temperature imbalances we analyzed the differences in temperature within the data centers in our datasets. We study the distribution of per-disk temperatures in different data centers at Google (using the dataset from Section 2.1) and the per-node temperatures for nodes in LANL's system 20 (using the dataset from Section 2.3). We consider how much hotter the disk/node in the 95th and 99th percentile of the distribution in the data center is, compared to the median disk/node.

[4]For reference, DRAM, for example, is typically rated for up to 95C.

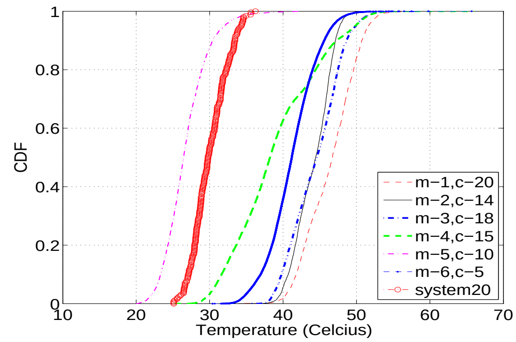

Figure 23: The cumulative distribution function of the per node/disk average temperatures for the Google data centers in our study and LANL's system 20.

Interestingly, the trends for temperature imbalances are very similar across data centers, despite the fact that they have been designed and managed by independent entities. We find that for all of Google's data centers in our study, and LANL's system 20, the node/disk in the 95th percentile is typically around 5 degrees C hotter than the median node/disk, and that the 99th percentile is around 8–10 degrees hotter than the median node/disk. Figure 23 shows the full CDFs of the per node/disk distribution for both the Google data centers in our study and LANL's system 20.

Observation 13: The degree of temperature variation across the nodes in a data center is surprisingly similar for all data centers in our study. The hottest 5% nodes tend to be more than 5C hotter than the typical node, while the hottest 1% nodes tend to be more than 8–10C hotter.

4. SUMMARY AND IMPLICATIONS

Increasing data center temperatures creates the potential for large energy savings and reductions in carbon emissions. Unfortunately, the pitfalls possibly associated with increased data center temperatures are not very well understood, and as a result most data centers operate at very conservative, low temperature levels. This work sheds some light on the issues involved in raising data center temperatures, and comes to some surprising conclusions.

Based on our study of data spanning more than a dozen data centers at three different organizations, and covering a broad range of reliability issues, we find that the effect of high data center temperatures on system reliability are smaller than often assumed. For some of the reliability issues we study, namely DRAM failures and node outages, we do not find any evidence for a correlation with higher temperatures (within the range of temperatures in our datasets). For those error conditions that show a correlation (latent sector errors in disks and disk failures), the correlation is much weaker than expected. For (device internal) temperatures below 50C, errors tend to grow linearly with temperature, rather than exponentially, as existing models suggest.

It is important to note that this does not mean that high temperatures have no effect on hardware reliability or that the Arrhenius model is flawed. But it might mean that the effects of other factors dominate failure rates. The Arrhenius model tries to solely capture the effect of heat on hardware components without taking into account other possible factors that impact hardware reliability in the field. Anecdotal evidence from discussions with data center operators suggests for example that poor handling procedures for equipment are a major factor in the field (which is hard to capture in measurement data). Our results indicate that, all things considered, the effect of temperature on hardware reliability is actually weaker than commonly thought.

We also find that, rather than average temperature, the variability in temperature might be the more important factor. Even failure conditions, such as node outages, that did not show a correlation with temperature, did show a clear correlation with the variability in temperature. Efforts in controlling such factors might be more important in keeping hardware failure rates low, than keeping temperatures low.

We also make some observations that might be helpful in protecting against temperature-induced hardware issues. The error mode that was most strongly correlated with high temperatures are LSEs. Common method for protecting against data loss due to LSEs include Read-after-Write and periodic "scrubbing" of the hard disk to proactively detect such errors. In experiments with our testbed based on a thermal chamber, we observe evidence that (enterprise class) hard disks do employ mechanisms, such as RaW, but we find that they tend to kick in only at very high temperatures and are associated with significant performance penalties. On the other hand, we find that one of the concerns often associated with scrubbing does not seem to be a valid concern in practice, which might make scrubbing the better approach to defend against LSEs: some fear that the extra workload placed on a drive by the scrub process might lead to early wear-out of the drive, but we see no correlation between a drive's workload intensity and it's failure probability.

Our encouraging results on the impact of temperature on hardware reliability move the focus to other potential issues with increasing data center temperatures. One such issue is an increase in the power consumption of individual servers as inlet air temperatures go up. The two most commonly cited reasons for such an increase are increased power leakage in the processor and increased (server internal) fan speeds. Our experimental results show that power leakage seems to be negligible compared to the effect of server fans. In fact, we find that even for relatively low ambient temperatures (on the orders that are commonly found in the hotter areas of an otherwise cool data center) fan power consumption makes up a significant fraction of total energy consumption. Much of this energy might be spent unnecessarily, due to poorly designed algorithms for controlling fan speed.

We would oversimplify the problem if we tried to make generalized recommendations or predictions on what exactly data center temperatures should be and how much energy precisely could be saved. The answer to these questions will depend on too many factors, that are data center or application specific. However, we see our results as strong evidence that most organizations could run their data centers hotter than they currently are without making significant sacrifices in system reliability. We hope that this paper will motivate future work in this area and encourage more organizations to share field data or results from analyzing their data.

5. ACKNOWLEDGMENTS

We would like to thank Google for hosting the last author of this paper as a visiting faculty in the System Health Group during the summer of 2009. In particular, we thank Wolf-Dietrich Weber for his efforts in making the data available for us and Chris Mayfield who provided scripts to extract some of the data. We would like to thank the members of SciNet, in particular Chris Loken and Ching-Hsing Yu, for providing us with the data from their GPC system. And we would like to thank LANL for making their data publicly available, and in particular Gary Grider and James Nunez who helped answer our questions about the data. We would also like to thank Jiri Schindler for the ideas for some of the thermal chamber experiments and the anonymous reviewers for their detailed feedback, especially our shepherd Abhishek Sharma. This work has been funded by an NSERC discovery grant and a Netapp Faculty Fellowship.

6. REFERENCES

[1] Operational Data to Support and Enable Computer Science Research, Los Alamos National Laboratory. http://institute.lanl.gov/data/fdata/.
[2] SciNet. http://www.scinet.utoronto.ca/.
[3] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410, 1990.
[4] D. Anderson, J. Dykes, and E. Riedel. More than an interface—SCSI vs. ATA. In *Proc. of FAST 2003*.
[5] L. N. Bairavasundaram, G. R. Goodson, S. Pasupathy, and J. Schindler. An Analysis of Latent Sector Errors in Disk Drives. In *Proc. of SIGMETRICS'07*, 2007.
[6] C. Belady, A. Rawson, J. Pfleuger, and T. Cader. The Green Grid Data Center Power Efficiency Metrics: PUE & DCiE. Technical report, Green Grid, 2008.
[7] D. J. Bradley, R. E. Harper, and S. W. Hunter. Workload-based power management for parallel computer systems. *IBM J.Res.Dev.*, 47:703–718, 2003.
[8] J. Brandon. Going Green In The Data Center: Practical Steps For Your SME To Become More Environmentally Friendly. *Processor*, 29, Sept. 2007.
[9] California Energy Commission. Summertime energy-saving tips for businesses. *consumerenergycenter.org/tips/business_summer.html*.
[10] G. Cole. Estimating drive reliability in desktop computers and consumer electronics systems. TP-338.1. Seagate. 2000.
[11] H. J. Curnow, B. A. Wichmann, and T. Si. A synthetic benchmark. *The Computer Journal*, 19:43–49, 1976.
[12] A. E. Darling, L. Carey, and W.-c. Feng. The design, implementation, and evaluation of mpiblast. In *In Proc. of ClusterWorld 2003*, 2003.
[13] K. Flautner and T. Mudge. Vertigo: automatic performance-setting for linux. In *Proc. of OSDI*, 2002.
[14] A. Gandhi, M. Harchol-Balter, R. Das, and C. Lefurgy. Optimal power allocation in server farms. In *Proc. of Sigmetrics '09*, 2009.
[15] JEDEC Global Standards for the Microelectronics Industry. Arrhenius equation for reliability. *http://www.jedec.org/standards-documents/dictionary/terms/arrhenius-equation-reliability*.
[16] J. M. Kaplan, W. Forrest, and N. Kindler. Revolutionizing data center energy efficiency. Technical report, McKinsey & Company, July 2008.
[17] J. Katcher. Postmark: a new file system benchmark. Network Appliance Tech Report TR3022, Oct. 1997.
[18] Lawrence Berkeley National Labs. Benchmarking Data Centers. *http://hightech.lbl.gov/benchmarking-dc.html*, December 2007.
[19] J. D. McCalpin. Memory bandwidth and machine balance in current high performance computers. *IEEE*

Comp. Soc. TCCA Newsletter, pages 19–25, Dec. 1995.

[20] L. McVoy and C. Staelin. lmbench: portable tools for performance analysis. In *Proc. of USENIX ATC*, 1996.

[21] D. C. Niemi. Unixbench. *http://www.tux.org/pub/tux/niemi/unixbench/*.

[22] C. D. Patel, C. E. Bash, R. Sharma, and M. Beitelmal. Smart cooling of data centers. In *Proc. of IPACK*, 2003.

[23] M. Patterson. The effect of data center temperature on energy efficiency. In *Proc. of ITHERM*, May 2008.

[24] E. Pinheiro, R. Bianchini, E. V. Carrera, and T. Heath. Load balancing and unbalancing for power and performance in cluster-based systems. In *Proc. of Workshop on Compilers and Operating Systems for Low Power (COLP)*, 2001.

[25] E. Pinheiro, W. D. Weber, and L. A. Barroso. Failure trends in a large disk drive population. In *Proc. of Usenix FAST 2007*.

[26] S. J. Plimpton, R. Brightwell, C. Vaughan, K. Underwood, and M. Davis. A Simple Synchronous Distributed-Memory Algorithm for the HPCC RandomAccess Benchmark. In *IEEE Cluster Computing*, Sept. 2006.

[27] K. Rajamani and C. Lefurgy. On evaluating request-distribution schemes for saving energy in server clusters. In *Proc. of the IEEE ISPASS*, 2003.

[28] Rich Miller. Google: Raise your data center temperature. *http://www.datacenterknowledge.com/archives/2008/10/14/google-raise-your-data-center-temperature/*, 2008.

[29] B. Schroeder and G. Gibson. A large-scale study of failures in high-performance computing systems. In *Proc. of DSN'06*, 2006.

[30] B. Schroeder and G. A. Gibson. Disk failures in the real world: What does an MTTF of 1,000,000 hours mean to you? In *Proc. of USENIX FAST*, 2007.

[31] B. Schroeder, E. Pinheiro, and W.-D. Weber. DRAM errors in the wild: a large-scale field study. In *Proc. of SIGMETRICS '09*, 2009.

[32] R. Sharma, C. Bash, C. Patel, R. Friedrich, and J. Chase. Balance of power: dynamic thermal management for internet data centers. *IEEE Internet Computing*, 9(1):42–49, 2005.

[33] C. Staelin. Lmbench3: Measuring scalability. Technical report, HP Laboratories Israel, 2002.

[34] R. F. Sullivan. Alternating Cold and Hot Aisles Provides More Reliable Cooling for Server Farms. In *Uptime Institute*, 2000.

[35] T10 Technical Committee. SCSI Block Commands – 3, Rev.25. Work. Draft T10/1799-D, ANSI INCITS.

[36] T13 Technical Committee. ATA 8 - ATA/ATAPI Command Set, Rev.4a. Work. Draft T13/1699-D, ANSI INCITS.

[37] Transaction Processing Performance Council. TPC Benchmark C - Rev. 5.11. Standard, Feb. 2010.

[38] Transaction Processing Performance Council. TPC Benchmark H - Rev. 2.14.2. Standard, June 2011.

[39] R. P. Weicker. Dhrystone: a synthetic systems programming benchmark. *Communications of the ACM*, 27(10):1013–1030, Oct. 1984.