

DRAM Errors in the Wild: A Large-Scale Field Study

Bianca Schroeder
Dept. of Computer Science
University of Toronto
Toronto, Canada
bianca@cs.toronto.edu

Eduardo Pinheiro
Google Inc.
Mountain View, CA

Wolf-Dietrich Weber
Google Inc.
Mountain View, CA

ABSTRACT

Errors in dynamic random access memory (DRAM) are a common form of hardware failure in modern compute clusters. Failures are costly both in terms of hardware replacement costs and service disruption. While a large body of work exists on DRAM in laboratory conditions, little has been reported on real DRAM failures in large production clusters. In this paper, we analyze measurements of memory errors in a large fleet of commodity servers over a period of 2.5 years. The collected data covers multiple vendors, DRAM capacities and technologies, and comprises many millions of DIMM days.

The goal of this paper is to answer questions such as the following: How common are memory errors in practice? What are their statistical properties? How are they affected by external factors, such as temperature and utilization, and by chip-specific factors, such as chip density, memory technology and DIMM age?

We find that DRAM error behavior in the field differs in many key aspects from commonly held assumptions. For example, we observe DRAM error rates that are orders of magnitude higher than previously reported, with 25,000 to 70,000 errors per billion device hours per Mbit and more than 8% of DIMMs affected by errors per year. We provide strong evidence that memory errors are dominated by hard errors, rather than soft errors, which previous work suspects to be the dominant error mode. We find that temperature, known to strongly impact DIMM error rates in lab conditions, has a surprisingly small effect on error behavior in the field, when taking all other factors into account. Finally, unlike commonly feared, we don't observe any indication that newer generations of DIMMs have worse error behavior.

Categories and Subject Descriptors: B.8 [Hardware]: Performance and Reliability; C.4 [Computer Systems Organization]: Performance of Systems;

General Terms: Reliability.

Keywords: DRAM, DIMM, memory, reliability, data corruption, soft error, hard error, large-scale systems.

1. INTRODUCTION

Errors in dynamic random access memory (DRAM) devices have been a concern for a long time [3, 11, 15–17, 23]. A memory error is an event that leads to the logical state of one or multiple bits being read differently from how they

were last written. Memory errors can be caused by electrical or magnetic interference (e.g. due to cosmic rays), can be due to problems with the hardware (e.g. a bit being permanently damaged), or can be the result of corruption along the data path between the memories and the processing elements. Memory errors can be classified into soft errors, which randomly corrupt bits but do not leave physical damage; and hard errors, which corrupt bits in a repeatable manner because of a physical defect.

The consequence of a memory error is system dependent. In systems using memory without support for error correction and detection, a memory error can lead to a machine crash or applications using corrupted data. Most memory systems in server machines employ error correcting codes (ECC) [5], which allow the detection and correction of one or multiple bit errors. If an error is uncorrectable, i.e. the number of affected bits exceed the limit of what the ECC can correct, typically a machine shutdown is forced. In many production environments, including ours, a single uncorrectable error is considered serious enough to replace the dual in-line memory module (DIMM) that caused it.

Memory errors are costly in terms of the system failures they cause and the repair costs associated with them. In production sites running large-scale systems, memory component replacements rank near the top of component replacements [20] and memory errors are one of the most common hardware problems to lead to machine crashes [19]. Moreover, recent work shows that memory errors can cause security vulnerabilities [7, 22]. There is also a fear that advancing densities in DRAM technology might lead to increased memory errors, exacerbating this problem in the future [3, 12, 13].

Despite the practical relevance of DRAM errors, very little is known about their prevalence in real production systems. Existing studies are mostly based on lab experiments using accelerated testing, where DRAM is exposed to extreme conditions (such as high temperature) to artificially induce errors. It is not clear how such results carry over to real production systems. The few existing studies that are based on measurements in real systems are small in scale, such as recent work by Li *et al.* [10], who report on DRAM errors in 300 machines over a period of 3 to 7 months.

One main reason for the limited understanding of DRAM errors in real systems is the large experimental scale required to obtain interesting measurements. A detailed study of errors requires data collection over a long time period (several years) and thousands of machines, a scale that researchers cannot easily replicate in their labs. Production sites, which run large-scale systems, often do not collect and record error

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SIGMETRICS/Performance'09, June 15–19, 2009, Seattle, WA, USA.

Copyright 2009 ACM 978-1-60558-511-6/09/06 ...\$5.00.

data rigorously, or are reluctant to share it because of the sensitive nature of data related to failures.

This paper provides the first large-scale study of DRAM memory errors in the field. It is based on data collected from Google’s server fleet over a period of more than two years making up many millions of DIMM days. The DRAM in our study covers multiple vendors, DRAM densities and technologies (DDR1, DDR2, and FBDIMM).

The paper addresses the following questions: How common are memory errors in practice? What are their statistical properties? How are they affected by external factors, such as temperature, and system utilization? And how do they vary with chip-specific factors, such as chip density, memory technology and DIMM age?

We find that in many aspects DRAM errors in the field behave very differently than commonly assumed. For example, we observe DRAM error rates that are orders of magnitude higher than previously reported, with FIT rates (failures in time per billion device hours) of 25,000 to 70,000 per Mbit and more than 8% of DIMMs affected per year. We provide strong evidence that memory errors are dominated by hard errors, rather than soft errors, which most previous work focuses on. We find that, out of all the factors that impact a DIMM’s error behavior in the field, temperature has a surprisingly small effect. Finally, unlike commonly feared, we don’t observe any indication that per-DIMM error rates increase with newer generations of DIMMs.

2. BACKGROUND AND METHODOLOGY

2.1 Memory errors and their handling

Most memory systems in use in servers today are protected by error detection and correction codes. The typical arrangement is for a memory access word to be augmented with additional bits to contain the error code. Typical error codes in commodity server systems today fall in the single error correct double error detect (SECCDED) category. That means they can reliably detect and correct any single-bit error, but they can only detect and not correct multiple bit errors. More powerful codes can correct and detect more error bits in a single memory word. For example, a code family known as chip-kill [6], can correct up to 4 adjacent bits at once, thus being able to work around a completely broken 4-bit wide DRAM chip. We use the terms correctable error (CE) and uncorrectable error (UE) in this paper to generalize away the details of the actual error codes used.

If done well, the handling of correctable memory errors is largely invisible to application software. Correction of the error and logging of the event can be performed in hardware for a minimal performance impact. However, depending on how much of the error handling is pushed into software, the impact can be more severe, with high error rates causing a significant degradation of overall system performance.

Uncorrectable errors typically lead to a catastrophic failure of some sort. Either there is an explicit failure action in response to the memory error (such as a machine reboot), or there is risk of a data-corruption-induced failure such as a kernel panic. In the systems we study, all uncorrectable errors are considered serious enough to shut down the machine and replace the DIMM at fault.

Memory errors can be classified into soft errors, which randomly corrupt bits, but do not leave any physical damage; and hard errors, which corrupt bits in a repeatable manner

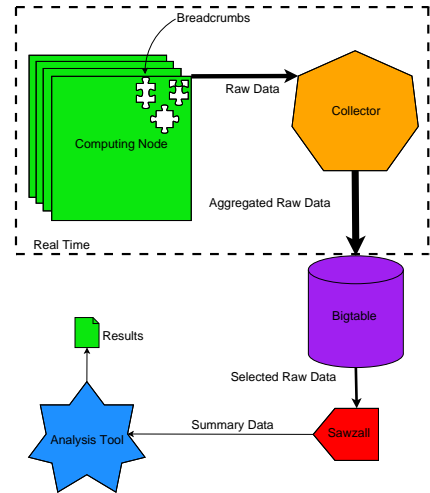


Figure 1: Collection, storage, and analysis architecture.

because of a physical defect (e.g. “stuck bits”). Our measurement infrastructure captures both hard and soft errors, but does not allow us to reliably distinguish these types of errors. All our numbers include both hard and soft errors.

Single-bit soft errors in the memory array can accumulate over time and turn into multi-bit errors. In order to avoid this accumulation of single-bit errors, memory systems can employ a hardware scrubber [14] that scans through the memory, while the memory is otherwise idle. Any memory words with single-bit errors are written back after correction, thus eliminating the single-bit error if it was soft. Three of the six hardware platforms (Platforms C, D and F) we consider make use of memory scrubbers. The typical scrubbing rate in those systems is 1GB every 45 minutes. In the other three hardware platforms (Platforms A, B, and E) errors are only detected on access.

2.2 The systems

Our data covers the majority of machines in Google’s fleet and spans nearly 2.5 years, from January 2006 to June 2008. Each machine comprises a motherboard with some processors and memory DIMMs. We study 6 different hardware *platforms*, where a platform is defined by the motherboard and memory generation.

The memory in these systems covers a wide variety of the most commonly used types of DRAM. The DIMMs come from multiple manufacturers and models, with three different capacities (1GB, 2GB, 4GB), and cover the three most common DRAM technologies: Double Data Rate (DDR1), Double Data Rate 2 (DDR2) and Fully-Buffered (FBDIMM). DDR1 and DDR2 have a similar interface, except that DDR2 provides twice the per-data-pin throughput (400 Mbit/s and 800 Mbit/s respectively). FBDIMM is a buffering interface around what is essentially a DDR2 technology inside.

2.3 The measurement methodology

Our collection infrastructure (see Figure 1) consists of locally recording events every time they happen. The logged events of interest to us are correctable errors, uncorrectable errors, CPU utilization, temperature, and memory allocated. These events (“breadcrumbs”) remain in the host machine

and are collected periodically (every 10 minutes) and archived in a Bigtable [4] for later processing. This collection happens continuously in the background.

The scale of the system and the data being collected make the analysis non-trivial. Each one of many ten-thousands of machines in the fleet logs every ten minutes hundreds of parameters, adding up to many TBytes. It is therefore impractical to download the data to a single machine and analyze it with standard tools. We solve this problem by using a parallel pre-processing step (implemented in Sawzall [18]), which runs on several hundred nodes simultaneously and performs basic data clean-up and filtering. We then perform the remainder of our analysis using standard analysis tools.

2.4 Analytical methodology

The metrics we consider are the rate and probability of errors over a given time period. For uncorrectable errors, we focus solely on probabilities, since a DIMM is expected to be removed after experiencing an uncorrectable error.

As part of this study, we investigate the impact of temperature and utilization (as measured by CPU utilization and amount of memory allocated) on memory errors. The exact temperature and utilization levels at which our systems operate are sensitive information. Instead of giving absolute numbers for temperature, we therefore report temperature values “normalized” by the smallest observed temperature. That is a reported temperature value of x , means the temperature was x degrees higher than the smallest observed temperature. The same approach does not work for CPU utilization, since the range of utilization levels is obvious (ranging from 0-100%). Instead, we report CPU utilization as multiples of the average utilization, i.e. a utilization of x , corresponds to a utilization level that is x times higher than the average utilization. We follow the same approach for allocated memory.

When studying the effect of various factors on memory errors, we often want to see how much higher or lower the monthly rate of errors is compared to an average month (independent of the factor under consideration). We therefore often report “normalized” rates and probabilities, i.e. we give rates and probabilities as multiples of the average. For example, when we say the normalized probability of an uncorrectable error is 1.5 for a given month, that means the uncorrectable error probability is 1.5 times higher than in an average month. This has the additional advantage that we can plot results for platforms with very different error probabilities in the same graph.

Finally, when studying the effect of factors, such as temperature, we report error rates as a function of percentiles of the observed factor. For example, we might report that the monthly correctable error rate is x if the temperature lies in the first temperature decile (i.e. the temperature is in the range of the lowest 10% of reported temperature measurements). This has the advantage that the error rates for each temperature range that we report on are based on the same number of data samples. Since error rates tend to be highly variable, it is important to compare data points that are based on a similar number of samples.

3. BASELINE STATISTICS

We start our study with the basic question of how common memory errors are in the field. Since a single uncorrectable error in a machine leads to the shut down of the entire ma-

Table 1: Memory errors per year:

Platf.	Tech.	Per machine				
		CE Incid. (%)	CE Rate Mean	CE Rate C.V.	CE Median Affct.	UE Incid. (%)
A	DDR1	45.4	19,509	3.5	611	0.17
B	DDR1	46.2	23,243	3.4	366	–
C	DDR1	22.3	27,500	17.7	100	2.15
D	DDR2	12.3	20,501	19.0	63	1.21
E	FBD	–	–	–	–	0.27
F	DDR2	26.9	48,621	16.1	25	4.15
Overall	–	32.2	22,696	14.0	277	1.29

Platf.	Tech.	Per DIMM				
		CE Incid. (%)	CE Rate Mean	CE Rate C.V.	CE Median Affct.	UE Incid. (%)
A	DDR1	21.2	4530	6.7	167	0.05
B	DDR1	19.6	4086	7.4	76	–
C	DDR1	3.7	3351	46.5	59	0.28
D	DDR2	2.8	3918	42.4	45	0.25
E	FBD	–	–	–	–	0.08
F	DDR2	2.9	3408	51.9	15	0.39
Overall	–	8.2	3751	36.3	64	0.22

chine, we begin by looking at the frequency of memory errors per machine. We then focus on the frequency of memory errors for individual DIMMs.

3.1 Errors per machine

Table 1 (top) presents high-level statistics on the frequency of correctable errors and uncorrectable errors per machine per year of operation, broken down by the type of hardware platform. Blank lines indicate lack of sufficient data.

Our first observation is that memory errors are not rare events. About a third of all machines in the fleet experience at least one memory error per year (see column CE Incid. %) and the average number of correctable errors per year is over 22,000. These numbers vary across platforms, with some platforms (e.g. Platform A and B) seeing nearly 50% of their machines affected by correctable errors, while in others only 12–27% are affected. The median number of errors per year for those machines that experience at least one error ranges from 25 to 611.

Interestingly, for those platforms with a lower percentage of machines affected by correctable errors, the average number of correctable errors per machine per year is the same or even higher than for the other platforms. We will take a closer look at the differences between platforms and technologies in Section 3.2.

We observe that for all platforms the number of errors per machine is highly variable with coefficients of variation between 3.4 and 20¹. Some machines develop a very large number of correctable errors compared to others. We find that for all platforms, 20% of the machines with errors make up more than 90% of all observed errors for that platform. One explanation for the high variability might be correlations between errors. A closer look at the data confirms this hypothesis: in more than 93% of the cases a machine that sees a correctable error experiences at least one more correctable error in the same year.

¹These are high C.V. values compared, for example, to an exponential distribution, which has a C.V. of 1, or a Poisson distribution, which has a C.V. of 1/mean.

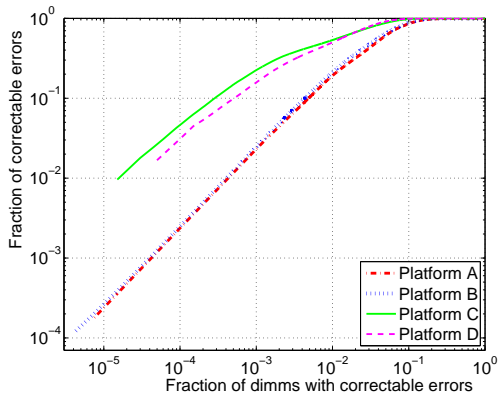


Figure 2: The distribution of correctable errors over DIMMs: The graph plots the fraction Y of all errors in a platform that is made up by the fraction X of DIMMs with the largest number of errors.

While correctable errors typically do not have an immediate impact on a machine, uncorrectable errors usually result in a machine shutdown. Table 1 shows, that while uncorrectable errors are less common than correctable errors, they do happen at a significant rate. Across the entire fleet, 1.3% of machines are affected by uncorrectable errors per year, with some platforms seeing as many as 2-4% affected.

3.2 Errors per DIMM

Since machines vary in the numbers of DRAM DIMMs and total DRAM capacity, we next consider per-DIMM statistics (Table 1 (bottom)).

Not surprisingly, the per-DIMM numbers are lower than the per-machine numbers. Across the entire fleet, 8.2% of all DIMMs are affected by correctable errors and an average DIMM experiences nearly 4000 correctable errors per year. These numbers vary greatly by platform. Around 20% of DIMMs in Platform A and B are affected by correctable errors per year, compared to less than 4% of DIMMs in Platform C and D. Only 0.05–0.08% of the DIMMs in Platform A and Platform E see an uncorrectable error per year, compared to nearly 0.3% of the DIMMs in Platform C and Platform D. The mean number of correctable errors per DIMM are more comparable, ranging from 3351–4530 correctable errors per year.

The differences between different platforms bring up the question of how chip-hardware specific factors impact the frequency of memory errors. We observe that there are two groups of platforms with members of each group sharing similar error behavior: there are Platform A, B, and E on one side, and Platform C, D and F on the other. While both groups have mean correctable error rates that are on the same order of magnitude, the first group has a much higher fraction of DIMMs affected by correctable errors, and the second group has a much higher fraction of DIMMs affected by uncorrectable errors.

We investigated a number of external factors that might explain the difference in memory rates across platforms, including temperature, utilization, DIMM age and capacity. While we will see (in Section 5) that all these affect the frequency of errors, they are not sufficient to explain the differences we observe between platforms.

Table 2: Errors per DIMM by DIMM type/manufacturer

Pf	Mfg	GB	Incid. CE (%)	Incid. UE (%)	Mean CE rate	C.V. CE	CEs/GB
A	1	1	20.6	0.03	4242	6.9	4242
	1	2	19.7	0.07	4487	5.9	2244
	2	1	6.6		1496	11.9	1469
	3	1	27.1	0.04	5821	6.2	5821
	4	1	5.3	0.03	1128	13.8	1128
B	1	1	20.3	–	3980	7.5	3980
		2	18.4	–	5098	6.8	2549
	2	1	7.9	–	1841	11.0	1841
		2	18.1	–	2835	8.9	1418
C	1	1	3.6	0.21	2516	69.7	2516
	4	1	2.6	0.43	2461	57.2	2461
	5	2	4.7	0.22	10226	12.0	5113
D	6	2	2.7	0.24	3666	39.4	1833
		4	5.7	0.24	12999	23.0	3250
E	1	2	–	0	–	–	–
		4	–	0.13	–	–	–
	2	2	–	0.05	–	–	–
		4	–	0.27	–	–	–
	5	2	–	0.06	–	–	–
		4	–	0.14	–	–	–
F	1	2	2.8	0.20	2213	53.0	1107
		4	4.0	1.09	4714	42.8	1179

A closer look at the data also lets us rule out memory technology (DDR1, DDR2, or FBDIMM) as the main factor responsible for the difference. Some platforms within the same group use different memory technology (e.g. DDR1 versus DDR2 in Platform C and D, respectively), while there are platforms in different groups using the same memory technology (e.g. Platform A, B and C all use DDR1). There is not one memory technology that is clearly superior to the others when it comes to error behavior.

We also considered the possibility that DIMMs from different manufacturers might exhibit different error behavior. Table 2 shows the error rates broken down by the most common DIMM types, where DIMM type is defined by the combinations of platform and manufacturer. We note that, DIMMs within the same platform exhibit similar error behavior, even if they are from different manufacturers. Moreover, we observe that DIMMs from some manufacturers (Mfg1, Mfg4) are used in a number of different platforms with very different error behavior. These observations show two things: the differences between platforms are not mainly due to differences between manufacturers and we do not see manufacturers that are consistently good or bad.

While we cannot be certain about the cause of the differences between platforms, we hypothesize that the observed differences in correctable errors are largely due to board and DIMM design differences. We suspect that the differences in uncorrectable errors are due to differences in the error correction codes in use. In particular, Platforms C and D are the only platforms that do not use a form of chip-kill [6]. Chip-kill is a more powerful code, that can correct certain types of multiple bit errors, while the codes in Platforms C and D can only correct single-bit errors.

We observe that for all platforms the number of correctable errors per DIMM per year is highly variable, with coefficients of variation ranging from 6 to 46. One might suspect that

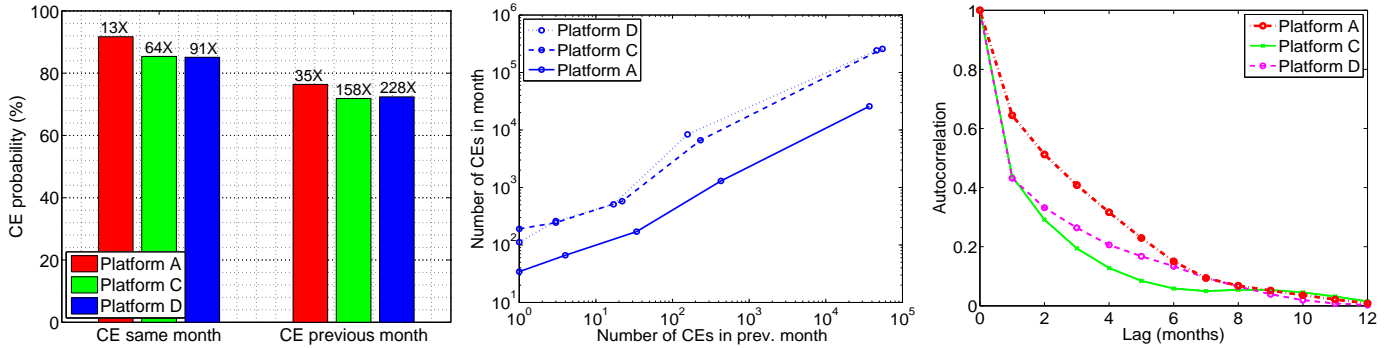


Figure 3: Correlations between correctable errors in the same DIMM: The left graph shows the probability of seeing a CE in a given month, depending on whether there were other CEs observed in the same month and the previous month. The numbers on top of each bar show the factor increase in probability compared to the CE probability in a random month (three left-most bars) and compared to the CE probability when there was no CE in the previous month (three right-most bars). The middle graph shows the expected number of CEs in a month as a function of the number of CEs in the previous month. The right graph shows the autocorrelation function for the number of CEs observed per month in a DIMM.

this is because the majority of the DIMMs see zero errors, while those affected see a large number of them. It turns out that even when focusing on only those DIMMs that have experienced errors, the variability is still high (not shown in table). The C.V. values range from 3–7 and there are large differences between the mean and the median number of correctable errors: the mean ranges from 20,000 – 140,000, while the median numbers are between 42 – 167.

Figure 2 presents a view of the distribution of correctable errors over DIMMs. It plots the fraction of errors made up by the top x percent of DIMMs with errors. For all platforms, the top 20% of DIMMs with errors make up over 94% of all observed errors. For Platform C and D, the distribution is even more skewed, with the top 20% of DIMMs comprising more than 99.6% of all errors. Note that the graph in Figure 2 is plotted on a log-log scale and that the lines for all platforms appear almost straight indicating a power-law distribution.

To a first order, the above results illustrate that errors in DRAM are a valid concern in practice. This motivates us to further study the statistical properties of errors (Section 4) and how errors are affected by various factors, such as environmental conditions (Section 5).

4. A CLOSER LOOK AT CORRELATIONS

In this section, we study correlations between correctable errors within a DIMM, correlations between correctable and uncorrectable errors in a DIMM, and correlations between errors in different DIMMs in the same machine.

Understanding correlations between errors might help identify when a DIMM is likely to produce a large number of errors in the future and replace it before it starts to cause serious problems.

4.1 Correlations between correctable errors

Figure 3 (left) shows the probability of seeing a correctable error in a given month, depending on whether there were correctable errors in the same month or the previous month. As the graph shows, for each platform the monthly correctable error probability increases dramatically in the presence of prior errors. In more than 85% of the cases a correctable

error is followed by at least one more correctable error in the same month. Depending on the platform, this corresponds to an increase in probability between 13X to more than 90X, compared to an average month. Also seeing correctable errors in the previous month significantly increases the probability of seeing a correctable error: The probability increases by factors of 35X to more than 200X, compared to the case when the previous month had no correctable errors.

Seeing errors in the previous month not only affects the probability, but also the expected number of correctable errors in a month. Figure 3 (middle) shows the expected number of correctable errors in a month, as a function of the number of correctable errors observed in the previous month. As the graph indicates, the expected number of correctable errors in a month increases continuously with the number of correctable errors in the previous month.

Figure 3 (middle) also shows that the expected number of errors in a month is significantly larger than the observed number of errors in the previous month. For example, in the case of Platform D, if the number of correctable errors in the previous month exceeds 100, the expected number of correctable errors in this month is more than 1,000. This is a 100X increase compared to the correctable error rate for a random month.

We also consider correlations over time periods longer than from one month to the next. Figure 3 (right) shows the autocorrelation function for the number of errors observed per DIMM per month, at lags up to 12 months. We observe that even at lags of several months the level of correlation is still significant.

4.2 Correlations between correctable and uncorrectable errors

Since uncorrectable errors are simply multiple bit corruptions (too many for the ECC to correct), one might wonder whether the presence of correctable errors increases the probability of seeing an uncorrectable error as well. This is the question we focus on next.

The three left-most bars in Figure 4 (left) show how the probability of experiencing an uncorrectable error in a given month increases if there are correctable errors in the same month. The graph indicates that for all platforms, the prob-

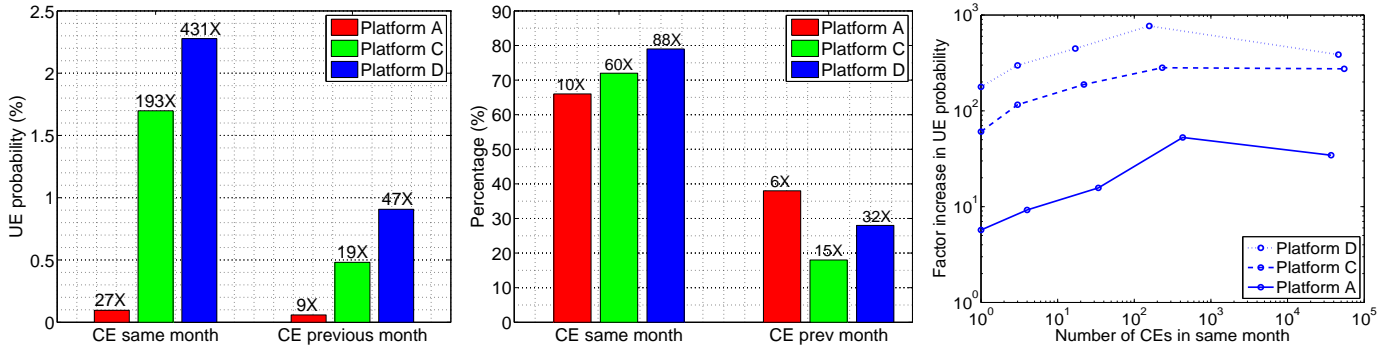


Figure 4: Correlations between correctable and uncorrectable errors in the same DIMM: *The left graph shows the UE probability in a month depending on whether there were CEs in the same month or in the previous month. The numbers on top of the bars give the increase in UE probability compared to a month without CEs (three left-most bars) and the case where there were no CEs in the previous month (three right-most bars). The middle graph shows how often a UE was preceded by a CE in the same/previous month. The right graph shows the factor increase in the probability of observing an UE as a function of the number of CEs in the same month.*

ability of an uncorrectable error is significantly larger in a month with correctable errors compared to a month without correctable errors. The increase in the probability of an uncorrectable error ranges from a factor of 27X (for Platform A) to more than 400X (for Platform D). While not quite as strong, the presence of correctable errors in the preceding month also affects the probability of uncorrectable errors. The three right-most bars in Figure 4 (left) show that the probability of seeing an uncorrectable error in a month following a month with at least one correctable errors is larger by a factor of 9X to 47X than if the previous month had no correctable errors.

Figure 4 (right) shows that not only the presence, but also the rate of observed correctable errors in the same month affects the probability of an uncorrectable error. Higher rates of correctable errors translate to a higher probability of uncorrectable errors. We see similar, albeit somewhat weaker trends when plotting the probability of uncorrectable errors as a function of the number of correctable errors in the previous month (not shown in figure). The uncorrectable error probabilities are about 8X lower than if the same number of correctable errors had happened in the same month, but still significantly higher than in a random month.

Given the above observations, one might want to use correctable errors as an early warning sign for impending uncorrectable errors. Another interesting view is therefore what fraction of uncorrectable errors are actually preceded by a correctable error, either in the same month or the previous month. Figure 4 (middle) shows that 65-80% of uncorrectable errors are preceded by a correctable error in the same month. Nearly 20-40% of uncorrectable errors are preceded by a correctable error in the previous month. Note that these probabilities are significantly higher than seeing a correctable error in an average month.

The above observations lead to the idea of early replacement policies, where a DIMM is replaced once it experiences a significant number of correctable errors, rather than waiting for the first uncorrectable error. However, while uncorrectable error probabilities are greatly increased after observing correctable errors, the absolute probabilities of an uncorrectable error are still relatively low (e.g. 1.7-2.3% in the case of Platform C and Platform D, see Figure 4 (left)).

We also experimented with more sophisticated methods for predicting uncorrectable errors, for example by building CART (Classification and regression trees) models based on parameters such as the number of CEs in the same and previous month, CEs and UEs in other DIMMs in the machine, DIMM capacity and model, but were not able to achieve significantly better prediction accuracy. Hence, replacing DIMMs solely based on correctable errors might be worth the price only in environments where the cost of downtime is high enough to outweigh the cost of the relatively high rate of false positives.

The observed correlations between correctable errors and uncorrectable errors will be very useful in the remainder of this study, when trying to understand the impact of various factors (such as temperature, age, utilization) on the frequency of memory errors. Since the frequency of correctable errors is orders of magnitudes higher than that of uncorrectable errors, it is easier to obtain conclusive results for correctable errors than uncorrectable errors. For the remainder of this study we focus mostly on correctable errors and how they are affected by various factors. We assume that those factors that increase correctable error rates, are likely to also increase the probability of experiencing an uncorrectable error.

4.3 Correlations between DIMMs in the same machine

So far we have focused on correlations between errors within the same DIMM. If those correlations are mostly due to external factors (such as temperature or workload intensity), we should also be able to observe correlations between errors in different DIMMs in the same machine, since these are largely subject to the same external factors.

Figure 5 shows the monthly probability of correctable and uncorrectable errors, as a function of whether there was an error in another DIMM in the same machine. We observe significantly increased error probabilities, compared to an average month, indicating a correlation between errors in different DIMMs in the same machine. However, the observed probabilities are lower as when an error was previously seen in the same DIMM (compare with Figure 3 (left) and Figure 4 (left)).

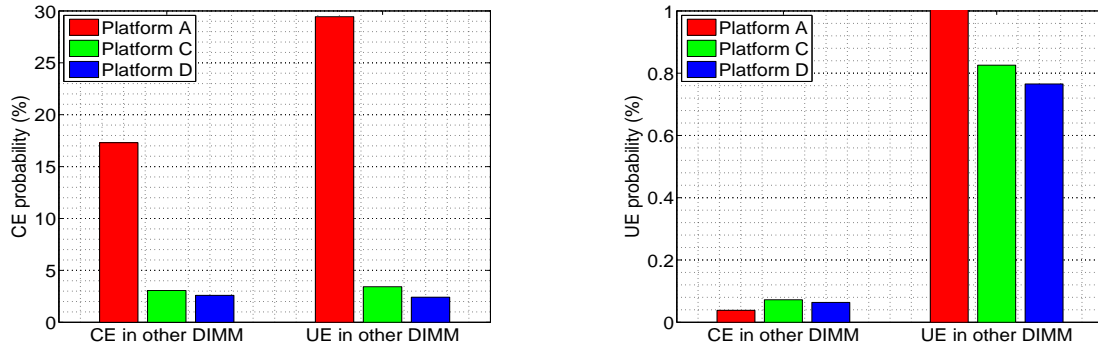


Figure 5: Correlations between errors in different DIMMs in the same machine: The graphs show the monthly CE probability (left) and UE probability (right) as a function of whether there was a CE or a UE in another DIMM in the same machine in the same month.

The fact that correlations between errors in different DIMMs are significantly lower than those between errors in the same DIMM might indicate that there are strong factors in addition to environmental factors that affect error behavior.

5. THE ROLE OF EXTERNAL FACTORS

In this section, we study the effect of various factors on correctable and uncorrectable error rates, including DIMM capacity, temperature, utilization, and age. We consider all platforms, except for Platform F, for which we do not have enough data to allow for a fine-grained analysis, and Platform E, for which we do not have data on CEs.

5.1 DIMM Capacity and chip size

Since the amount of memory used in typical server systems keeps growing from generation to generation, a commonly asked question when projecting for future systems, is how an increase in memory affects the frequency of memory errors. In this section, we focus on one aspect of this question. We ask how error rates change, when increasing the capacity of individual DIMMs.

To answer this question we consider all DIMM types (type being defined by the combination of platform and manufacturer) that exist in our systems in two different capacities. Typically, the capacities of these DIMM pairs are either 1GB and 2GB, or 2GB and 4GB (recall Table 2). Figure 6 shows for each of these pairs the factor by which the monthly probability of correctable errors, the correctable error rate and the probability of uncorrectable errors changes, when doubling capacity².

Figure 6 indicates a trend towards worse error behavior for increased capacities, although this trend is not consistent. While in some cases the doubling of capacity has a clear negative effect (factors larger than 1 in the graph), in others it has hardly any effect (factor close to 1 in the graph). For example, for Platform A -Mfg1 and Platform F -Mfg1 doubling the capacity increases uncorrectable errors, but not correctable errors. Conversely, for Platform D -Mfg6 doubling the capacity affects correctable errors, but not uncorrectable error.

The difference in how scaling capacity affects errors might be due to differences in how larger DIMM capacities are

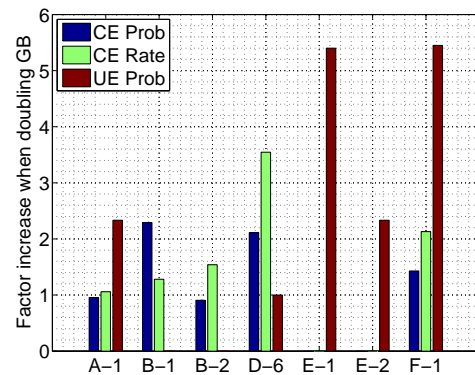


Figure 6: Memory errors and DIMM capacity: The graph shows for different Platform-Manufacturer pairs the factor increase in CE rates, CE probabilities and UE probabilities, when doubling the capacity of a DIMM.

built, since a given DIMM capacity can be achieved in multiple ways. For example, a one gigabyte DIMM with ECC can be manufactured with 36 256-megabit chips, or 18 512-megabit chips or with 9 one-gigabit chips.

We studied the effect of chip sizes on correctable and uncorrectable errors, controlling for capacity, platform (dimmm technology), and age. The results are mixed. When two chip configurations were available within the same platform, capacity and manufacturer, we sometimes observed an increase in average correctable error rates and sometimes a decrease. This either indicates that chip size does not play a dominant role in influencing CEs or there are other, stronger confounders in our data that we did not control for.

In addition to a correlation of chip size with error rates, we also looked for correlations of chip size with incidence of correctable and uncorrectable errors. Again we observe no clear trends. We also repeated the study of chip size effect without taking information on the manufacturer and/or age into account, again without any clear trends emerging.

The best we can conclude therefore is that any chip size effect is unlikely to dominate error rates given that the trends are not consistent across various other confounders such as age and manufacturer.

²Some bars are omitted, as we do not have data on UEs for Platform B and data on CEs for Platform E.

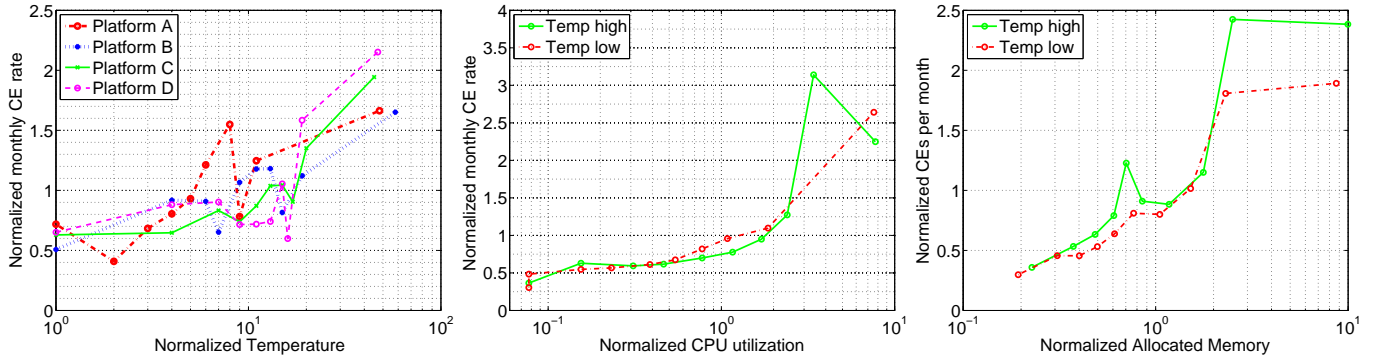


Figure 7: The effect of temperature: The left graph shows the normalized monthly rate of experiencing a correctable error as a function of the monthly average temperature, in deciles. The middle and right graph show the monthly rate of experiencing a correctable error as a function of memory usage and CPU utilization, respectively, depending on whether the temperature was high (above median temperature) or low (below median temperature). We observe that when isolating temperature by controlling for utilization, it has much less of an effect.

5.2 Temperature

Temperature is considered to (negatively) affect the reliability of many hardware components due to the strong physical changes on materials that it causes. In the case of memory chips, high temperature is expected to increase leakage current [2, 8] which in turn leads to a higher likelihood of flipped bits in the memory array.

In the context of large-scale production systems, understanding the exact impact of temperature on system reliability is important, since cooling is a major cost factor. There is a trade-off to be made between increased cooling costs and increased downtime and maintenance costs due to higher failure rates.

Our temperature measurements stem from a temperature sensor on the motherboard of each machine. For each platform, the physical location of this sensor varies relative to the position of the DIMMs, hence our temperature measurements are only an approximation of the actual temperature of the DIMMs.

To investigate the effect of temperature on memory errors we turn to Figure 7 (left), which shows the normalized monthly correctable error rate for each platform, as a function of temperature deciles (recall Section 2.4 for the reason of using deciles and the definition of normalized probabilities). That is the first data point (x_1, y_1) shows the monthly correctable error rate y_1 , if the temperature is less than the first temperature decile (temperature x_1). The second data point (x_2, y_2) shows the correctable error rate y_2 , if the temperature is between the first and second decile (between x_1 and x_2), and so on.

Figure 7 (left) shows that for all platforms higher temperatures are correlated with higher correctable error rates. In fact, for most platforms the correctable error rate increases by a factor of 3 or more when moving from the lowest to the highest temperature decile (corresponding to an increase in temperature by around 20C for Platforms B, C and D and an increase by slightly more than 10C for Platform A).

It is not clear whether this correlation indicates a causal relationship, i.e. higher temperatures inducing higher error rates. Higher temperatures might just be a proxy for higher system utilization, i.e. the utilization increases leading independently to higher error rates and higher temperatures. In

Figure 7 (middle) and (right) we therefore isolate the effects of temperature from the effects of utilization. We divide the utilization measurements (CPU utilization and allocated memory, respectively) into deciles and report for each decile the observed error rate when temperature was “high” (above median temperature) or “low” (below median temperature). We observe that when controlling for utilization, the effects of temperature are significantly smaller. We also repeated these experiments with higher differences in temperature, e.g. by comparing the effect of temperatures above the 9th decile to temperatures below the 1st decile. In all cases, for the same utilization levels the error rates for high versus low temperature are very similar.

5.3 Utilization

The observations in the previous subsection point to system utilization as a major contributing factor in memory error rates. Ideally, we would like to study specifically the impact of memory utilization (i.e. number of memory accesses). Unfortunately, obtaining data on memory utilization requires the use of hardware counters, which our measurement infrastructure does not collect. Instead, we study two signals that we believe provide indirect indication of memory activity: CPU utilization and memory allocated.

CPU utilization is the load activity on the CPU(s) measured instantaneously as a percentage of total CPU cycles used out of the total CPU cycles available and are averaged per machine for each month.

Memory allocated is the total amount of memory marked as used by the operating system on behalf of processes. It is a value in bytes and it changes as the tasks request and release memory. The allocated values are averaged per machine over each month.

Figure 8 (left) and (right) show the normalized monthly rate of correctable errors as a function of CPU utilization and memory allocated, respectively. We observe clear trends of increasing correctable error rates with increasing CPU utilization and allocated memory. Averaging across all platforms, it seems that correctable error rates grow roughly logarithmically as a function of utilization levels (based on the roughly linear increase of error rates in the graphs, which have log scales on the X-axis).

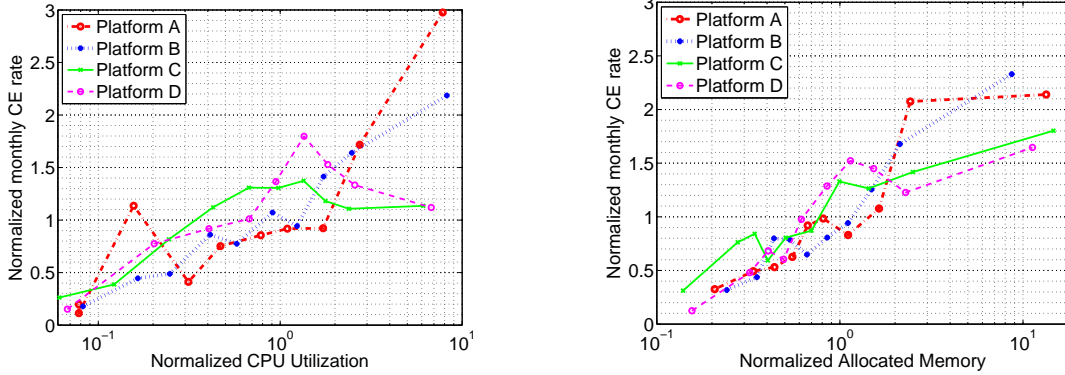


Figure 8: The effect of utilization: *The normalized monthly CE rate as a function of CPU utilization (left) and memory allocated (right).*

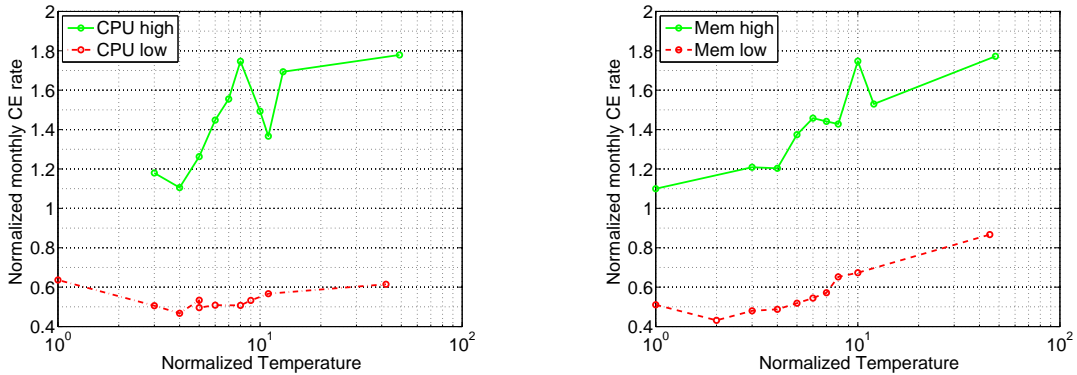


Figure 9: Isolating the effect of utilization: *The normalized monthly CE rate as a function of CPU utilization (left) and memory allocated (right), while controlling for temperature.*

One might ask whether utilization is just a proxy for temperature, where higher utilization leads to higher system temperatures, which then cause higher error rates. In Figure 9, we therefore isolate the effects of utilization from those of temperature. We divide the observed temperature values into deciles and report for each range the observed error rates when utilization was "high" or "low". High utilization means the utilization (CPU utilization and allocated memory, respectively) is above median and low means the utilization was below median. We observe that even when keeping temperature fixed and focusing on one particular temperature decile, there is still a huge difference in the error rates, depending on the utilization. For all temperature levels, the correctable error rates are by a factor of 2–3 higher for high utilization compared to low utilization.

The higher error rate for higher utilization levels might simply be due to a higher detection rate of errors, not an increased incidence of errors. For Platforms A and B, which do not employ a memory scrubber, this might be the case. However, we note that for Platforms C and D, which do use memory scrubbing, the number of reported *soft* errors should be the same, independent of utilization levels, since errors that are not found by a memory access, will be detected by the scrubber. The higher incidence of memory errors at higher utilizations must therefore be due to a different error mechanism, such as hard errors or errors induced on the datapath, either in the DIMMs or on the motherboard.

5.4 Aging

Age is one of the most important factors in analyzing the reliability of hardware components, since increased error rates due to early aging/wear-out limit the lifetime of a device. As such, we look at changes in error behavior over time for our DRAM population, breaking it down by age, platform, technology, correctable and uncorrectable errors.

5.4.1 Age and Correctable Errors

Figure 10 shows normalized correctable error rates as a function of age for all platforms (left) and for four of the most common DIMM configurations (platform, manufacturer and capacity). We observe that age clearly affects the correctable error rates for all platforms.

For a more fine-grained view of the effects of aging, we consider the mean cumulative function (MCF) of errors. Intuitively, the MCF value for a given age x represents the expected number of errors a DIMM will have seen by age x . That is, for each age point, we compute the number of DIMMs with errors divided by the total number of DIMMs at risk at that age and add this number to the previous running sum, hence the term cumulative. The use of a cumulative mean function helps visualizing trends, as it allows us to plot points at discrete rates. A regular age versus rate plot would be very noisy if plotted at such a fine-granularity.

The left-most graph in Figure 11 shows the MCF for all DIMMs in our population that were in production in Jan-

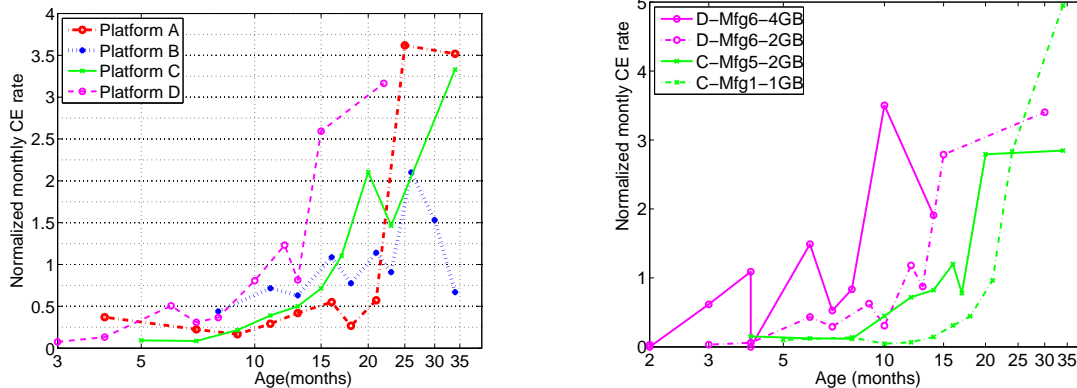


Figure 10: The effect of age: The normalized monthly rate of experiencing a CE as a function of age by platform (left) and for four common DIMM configurations (right). We consider only DIMMs manufactured after July 2005, to exclude very old platforms (due to a rapidly decreasing population).

uary 2007 and had a correctable error. We see that the correctable error rate starts to increase quickly as the population ages beyond 10 months up until around 20 months. After around 20 months, the correctable error incidence remains constant (flat slope).

The flat slope means that the error incidence rates reach a constant level, implying that older DIMMs continue to have correctable errors (even at an increased pace as shown by Figure 10), but there is not a significant increase in the incidence of correctable error for other DIMMs. Interestingly, this may indicate that older DIMMs that did not have correctable errors in the past, possibly will not develop them later on.

Since looking at the MCF for the entire population might confound many other factors, such as platform and DRAM technology, we isolate the aging effect by focusing on one individual platform. The second graph from the left in Figure 11 shows the MCF for correctable errors for Platform C, which uses only DDR1 RAM. We see a pattern very similar to that for the entire population. While not shown, due to lack of space, the shape of the MCF is similar for all other platforms. The only difference between platforms is the age when the MCF begins to steepen.

We also note the lack of infant mortality for almost all populations: none of the MCF figures shows a steep incline near very low ages. We attribute this behavior to the weeding out of bad DIMMs that happens during the burn-in of DIMMs prior to putting them into production.

In summary, our results indicate that age severely affects correctable error rates: one should expect an increasing incidence of errors as DIMMs get older, but only up to a certain point, when the incidence becomes almost constant (few DIMMs start to have correctable errors at very old ages). The age when errors first start to increase and the steepness of the increase vary per platform, manufacturer and DRAM technology, but is generally in the 10–18 month range.

5.4.2 Age and Uncorrectable Errors

We now turn to uncorrectable errors and aging effects. The two right-most graphs in Figure 11 show the mean cumulative function for uncorrectable errors for the entire population of DIMMs that were in production in January 2007, and for all DIMMs in Platform C, respectively. In these

figures, we see a sharp increase in correctable errors at early ages (3–5 months) and then a subsequent flattening of error incidence. This flattening is due to our policy of replacing DIMMs that experience uncorrectable errors, and hence the incidence of uncorrectable errors at very old ages is very low (flat slope in the figures).

In summary, uncorrectable errors are strongly influenced by age with slightly different behaviors depending on the exact demographics of the DIMMs (platform, manufacturer, DIMM technology). Our replacement policy enforces the survival of the fittest.

6. RELATED WORK

Much work has been done in understanding the behavior of DRAM in the laboratory. One of the earliest published work comes from May and Woods [11] and explains the physical mechanisms in which alpha-particles (presumably from cosmic rays) cause soft errors in DRAM. Since then, other studies have shown that radiation and errors happens at ground level [16], how soft error rates vary with altitude and shielding [23], and how device technology and scaling [3, 9] impact reliability of DRAM components. Baumann [3] shows that per-bit soft-error rates are going down with new generations, but that the reliability of the system-level memory ensemble has remained fairly constant.

All the above work differs from ours in that it is limited to laboratory studies and focused on only soft errors. Very few studies have examined DRAM errors in the field, in large populations. One such study is the work by Li *et al.* which reports soft-error rates for clusters of up to 300 machines. Our work differs from Li’s in the scale of the DIMM-days observed by several orders of magnitude. Moreover, our work reports on uncorrectable as well as correctable errors, and includes analysis of covariates commonly thought to be correlated with memory errors, such as age, temperature, and workload intensity.

We observe much higher error rates than previous work. Li *et al* cite error rates in the 200–5000 FIT per Mbit range from previous lab studies, and themselves found error rates of < 1 FIT per Mbit. In comparison, we observe mean correctable error rates of 2000–6000 per GB per year, which translate to 25,000–75,000 FIT per Mbit. Furthermore, for

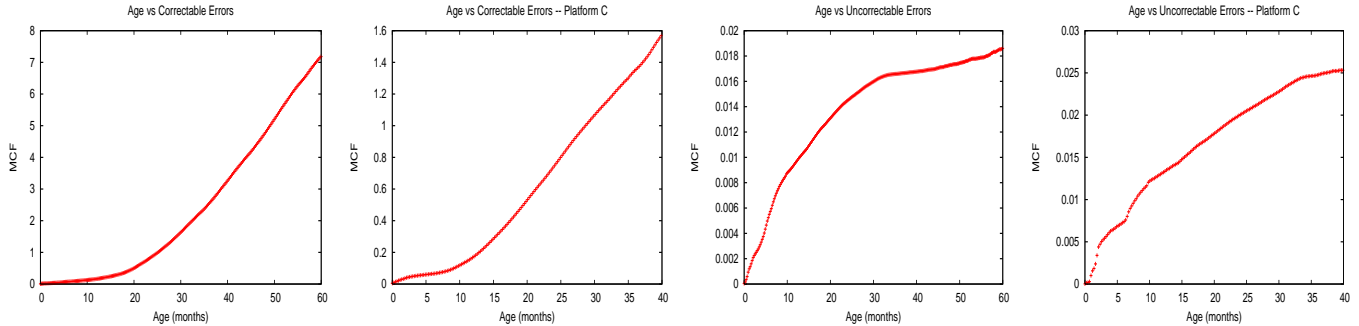


Figure 11: The effect of age: The two graphs on the left show the mean cumulative function for CEs for all DIMMs in production in January 2007 until November 2008, and for Platform C, respectively. The two graphs on the right show for the same two populations the mean cumulative function for UEs.

DIMMs with errors we observe median CE rates from 15 – 167 per month, translating to a FIT range of 778 – 25,000 per Mbit. A possible reason for our wider range of errors might be that our work includes both hard and soft errors.

7. SUMMARY AND DISCUSSION

This paper studied the incidence and characteristics of DRAM errors in a large fleet of commodity servers. Our study is based on data collected over more than 2 years and covers DIMMs of multiple vendors, generations, technologies, and capacities. All DIMMs were equipped with error correcting logic (ECC) to correct at least single bit errors.

Our study includes both correctable errors (CE) and uncorrectable errors (UE). Correctable errors can be handled by the ECC and are largely transparent to the application. Uncorrectable errors have more severe consequences, and in our systems lead to a machine shut-down and replacement of the affected DIMM. The error rates we report include both soft errors, which are randomly corrupted bits that can be corrected without leaving permanent damage, and hard errors, which are due to a physical defect and are permanent. Below we briefly summarize our results and discuss their implications.

Conclusion 1: We found the incidence of memory errors and the range of error rates across different DIMMs to be much higher than previously reported.

About a third of machines and over 8% of DIMMs in our fleet saw at least one correctable error per year. Our per-DIMM rates of correctable errors translate to an average of 25,000–75,000 FIT (failures in time per billion hours of operation) per Mbit and a median FIT range of 778 – 25,000 per Mbit (median for DIMMs with errors), while previous studies report 200-5,000 FIT per Mbit. The number of correctable errors per DIMM is highly variable, with some DIMMs experiencing a huge number of errors, compared to others. The annual incidence of uncorrectable errors was 1.3% per machine and 0.22% per DIMM.

The conclusion we draw is that error correcting codes are crucial for reducing the large number of memory errors to a manageable number of uncorrectable errors. In fact, we found that platforms with more powerful error codes (chip-kill versus SECDED) were able to reduce uncorrectable error rates by a factor of 4–10 over the less powerful codes. Nonetheless, the remaining incidence of 0.22% per DIMM

per year makes a crash-tolerant application layer indispensable for large-scale server farms.

Conclusion 2: Memory errors are strongly correlated.

We observe strong correlations among correctable errors within the same DIMM. A DIMM that sees a correctable error is 13–228 times more likely to see another correctable error in the same month, compared to a DIMM that has not seen errors. There are also correlations between errors at time scales longer than a month. The autocorrelation function of the number of correctable errors per month shows significant levels of correlation up to 7 months.

We also observe strong correlations between correctable errors and uncorrectable errors. In 70-80% of the cases an uncorrectable error is preceded by a correctable error in the same month or the previous month, and the presence of a correctable error increases the probability of an uncorrectable error by factors between 9–400. Still, the absolute probabilities of observing an uncorrectable error following a correctable error are relatively small, between 0.1–2.3% per month, so replacing a DIMM solely based on the presence of correctable errors would be attractive only in environments where the cost of downtime is high enough to outweigh the cost of the expected high rate of false positives.

Conclusion 3: The incidence of CEs increases with age, while the incidence of UEs decreases with age (due to replacements).

Given that DRAM DIMMs are devices without any mechanical components, unlike for example hard drives, we see a surprisingly strong and early effect of age on error rates. For all DIMM types we studied, aging in the form of increased CE rates sets in after only 10–18 months in the field. On the other hand, the rate of incidence of uncorrectable errors continuously declines starting at an early age, most likely because DIMMs with UEs are replaced (survival of the fittest).

Conclusion 4: There is no evidence that newer generation DIMMs have worse error behavior.

There has been much concern that advancing densities in DRAM technology will lead to higher rates of memory errors in future generations of DIMMs. We study DIMMs in six different platforms, which were introduced over a period of several years, and observe no evidence that CE rates increase with newer generations. In fact, the DIMMs used in

the three most recent platforms exhibit lower CE rates, than the two older platforms, despite generally higher DIMM capacities. This indicates that improvements in technology are able to keep up with adversarial trends in DIMM scaling.

Conclusion 5: Within the range of temperatures our production systems experience in the field, temperature has a surprisingly low effect on memory errors.

Temperature is well known to increase error rates. In fact, artificially increasing the temperature is a commonly used tool for accelerating error rates in lab studies. Interestingly, we find that differences in temperature in the range they arise naturally in our fleet’s operation (a difference of around 20C between the 1st and 9th temperature decile) seem to have a marginal impact on the incidence of memory errors, when controlling for other factors, such as utilization.

Conclusion 6: Error rates are strongly correlated with utilization.

Conclusion 7: Error rates are unlikely to be dominated by soft errors.

We observe that CE rates are highly correlated with system utilization, even when isolating utilization effects from the effects of temperature. In systems that do not use memory scrubbers this observation might simply reflect a higher detection rate of errors. In systems with memory scrubbers, this observations leads us to the conclusion that a significant fraction of errors is likely due to mechanism other than soft errors, such as hard errors or errors induced on the datapath. The reason is that in systems with memory scrubbers the reported rate of soft errors should not depend on utilization levels in the system. Each soft error will eventually be detected (either when the bit is accessed by an application or by the scrubber), corrected and reported. Another observation that supports Conclusion 7 is the strong correlation between errors in the same DIMM. Events that cause soft errors, such as cosmic radiation, are expected to happen randomly over time and not in correlation.

Conclusion 7 is an interesting observation, since much previous work has assumed that soft errors are the dominating error mode in DRAM. Some earlier work estimates hard errors to be orders of magnitude less common than soft errors [21] and to make up about 2% of all errors [1]. Conclusion 7 might also explain the significantly higher rates of memory errors we observe compared to previous studies.

Acknowledgments

We would like to thank Luiz Barroso, Urs Hoelzle, Chris Johnson, Nick Sanders and Kai Shen for their feedback on drafts of this paper. We would also like to thank those who contributed directly or indirectly to this work: Kevin Bartz, Bill Heavlin, Nick Sanders, Rob Sprinkle, and John Zapisek. Special thanks to the System Health Infrastructure team for providing the data collection and aggregation mechanisms. Finally, the first author would like to thank the System Health Group at Google for hosting her during the summer of 2008.

8. REFERENCES

- [1] Mosys adds soft-error protection, correction. *Semiconductor Business News*, 28 Jan. 2002.
- [2] Z. Al-Ars, A. J. van de Goor, J. Braun, and D. Richter. Simulation based analysis of temperature effect on the faulty behavior of embedded drams. In *ITC’01: Proc. of the 2001 IEEE International Test Conference*, 2001.
- [3] R. Baumann. Soft errors in advanced computer systems. *IEEE Design and Test of Computers*, pages 258–266, 2005.
- [4] F. Chang, J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach, M. Burrows, T. Chandra, A. Fikes, and R. E. Gruber. Bigtable: A distributed storage system for structured data. In *Proc. of OSDI’06*, 2006.
- [5] C. Chen and M. Hsiao. Error-correcting codes for semiconductor memory applications: A state-of-the-art review. *IBM J. Res. Dev.*, 28(2):124–134, 1984.
- [6] T. J. Dell. A white paper on the benefits of chipkill-correct ECC for PC server main memory. *IBM Microelectronics*, 1997.
- [7] S. Govindavajhala and A. W. Appel. Using memory errors to attack a virtual machine. In *SP ’03: Proc. of the 2003 IEEE Symposium on Security and Privacy*, 2003.
- [8] T. Hamamoto, S. Sugiura, and S. Sawada. On the retention time distribution of dynamic random access memory (dram). *IEEE Transactions on Electron Devices*, 45(6):1300–1309, 1998.
- [9] A. H. Johnston. Scaling and technology issues for soft error rates. In *Proc. of the 4th Annual Conf. on Reliability*, 2000.
- [10] X. Li, K. Shen, M. Huang, and L. Chu. A memory soft error measurement on production systems. In *Proc. of USENIX Annual Technical Conference*, 2007.
- [11] T. C. May and M. H. Woods. Alpha-particle-induced soft errors in dynamic memories. *IEEE Transactions on Electron Devices*, 26(1), 1979.
- [12] Messer, Bernadat, Fu, Chen, Dimitrijevic, Lie, Mannaru, Riska, and Milojevic. Susceptibility of commodity systems and software to memory soft errors. *IEEE Transactions on Computers*, 53(12), 2004.
- [13] D. Milojevic, A. Messer, J. Shau, G. Fu, and A. Munoz. Increasing relevance of memory hardware errors: a case for recoverable programming models. In *Proc. of the 9th ACM SIGOPS European workshop*, 2000.
- [14] S. S. Mukherjee, J. Emer, T. Fossum, and S. K. Reinhardt. Cache scrubbing in microprocessors: Myth or necessity? In *PRDC ’04: Proceedings of the 10th IEEE Pacific Rim International Symposium on Dependable Computing*, 2004.
- [15] S. S. Mukherjee, J. Emer, and S. K. Reinhardt. The soft error problem: An architectural perspective. In *HPCA ’05: Proc. of the 11th International Symposium on High-Performance Computer Architecture*, 2005.
- [16] E. Normand. Single event upset at ground level. *IEEE Transaction on Nuclear Sciences*, 6(43):2742–2750, 1996.
- [17] T. J. O’Gorman, J. M. Ross, A. H. Taber, J. F. Ziegler, H. P. Muhlfield, C. J. Montrose, H. W. Curtis, and J. L. Walsh. Field testing for cosmic ray soft errors in semiconductor memories. *IBM J. Res. Dev.*, 40(1), 1996.
- [18] R. Pike, S. Dorward, R. Griesemer, and S. Quinlan. Interpreting the data: Parallel analysis with sawzall. *Scientific Programming Journal, Special Issue on Grids and Worldwide Computing Programming Models and Infrastructure*, 13(4), 2005.
- [19] B. Schroeder and G. A. Gibson. A large scale study of failures in high-performance-computing systems. In *DSN 2006: Proc. of the International Conference on Dependable Systems and Networks*, 2006.
- [20] 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 *5th USENIX FAST Conference*, 2007.
- [21] K. Takeuchi, K. Shimohigashi, H. Kozuka, T. Toyabe, K. Itoh, and H. Kurosawa. Origin and characteristics of alpha-particle-induced permanent junction leakage. *IEEE Transactions on Electron Devices*, March 1999.
- [22] J. Xu, S. Chen, Z. Kalbarczyk, and R. K. Iyer. An experimental study of security vulnerabilities caused by errors. In *DSN 2001: Proc. of the 2001 International Conference on Dependable Systems and Networks*, 2001.
- [23] J. F. Ziegler and W. A. Lanford. Effect of Cosmic Rays on Computer Memories. *Science*, 206:776–788, 1979.