# Comparison of off-chip interconnect validation to field failures

David Blankenbeckler, Adam Norman
Intel Corporation
Santa Clara, CA, USA
David.Blankenbeckler@Intel.com
Adam.j.Norman@Intel.com

Michael Shepherd
Dell Inc.
Round Rock, TX, USA
Michael_Shepherd@Dell.com

***Abstract*— Memory subsystem errors continue to be a common problem in modern computer systems. Through a large scale field study, this paper will introduce the interconnect transient margin validation metrics and compare to the observed field failures. The results will demonstrate that transient bus errors are not a dominant cause of system memory problems.**

***Keywords – DDR, DRAM, memory, bus margin***

## I. INTRODUCTION

Memory subsystem errors have remained a common form of failure since the advent of the computer. While much work has been done to reduce failures and gracefully handle them, they continue to be a significant problem in modern day computer architecture.

Over the past several years, at least two large scale studies have been conducted to help quantify the extent of memory bus related failures. The recent white paper "Dram Errors in the Wild: A Large-Scale Field Study" by Bianca Schroeder, et al, indicated the rate was as high as 1/3 of systems experiencing at least one memory error per year [1]. Another study found at least 11 systems out of 212 that show symptoms of memory errors [7]. But what is the cause of these high failure rates? Most large scale studies have focused on Soft Error Rates (SERs) due to alpha particles [8], junction/cell leakage, manufacturing defects and the rate of errors across die shrinks.

During the late 70's, alpha particles from decaying package contaminants were a dominant source of memory errors [2]. Around the same time, researchers at IBM found that cosmic rays were also a source of transient memory errors, even at sea level [3][4]. In one study it was reported that memory errors were about 100x more likely at the altitude typically used by commercial aircraft [5]. These radiation induced errors are generally referred to as soft errors and have been the subject of much research. Today, this phenomenon is generally understood and thus effective mitigation techniques have been and continue to be developed [6].

Besides soft errors, there are various types of hard errors which could occur in either the memory controller or DRAM device. The most common hard failures are due to defects produced during the wafer manufacturing process. These failures may also be caused by design marginalities or aging effects.

This study is uniquely different from prior work in that the goal was to better understand the relationship between bus related margins and their resulting error rates. As bus speeds have increased signal integrity has become a factor suspected of being a significant contributor to transient errors. A properly designed system should have a reliable interface between the memory controller and the DRAMs. However, in the real world, factors such as excessive manufacturing variation or unexpected environmental conditions may impact reliable data transfers. Over large volume, these variations may increase noise on the bus and thereby increase the chance of transient errors. As shown in Figure 1, the year over year incident rate of end user memory errors has remained relatively unchanged. This data indicates that even across DRAM technologies and speeds between 2006 and 2009; memory system failure rates have remained relatively stable. Why do we see this consistent failure rate and what is causing it?

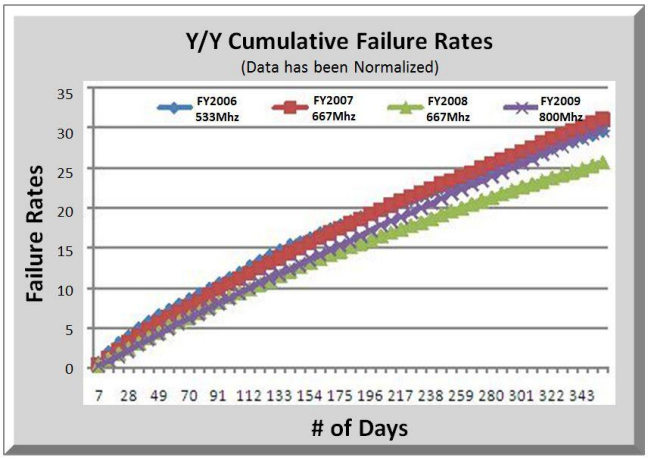


Figure 1. Year over year failure rates have remained relatively stable.

Could signal integrity now be a key factor in transient memory errors? This paper will explore this question through a large scale study comprised of nearly a quarter



 ISBN: 978-1-61208-168-7 

million systems and over a million DIMMs (dual in-line memory module). Section II will describe the concept of bus margin which provides a measure of bus noise susceptibility and the different sources of memory subsystem errors. Section III will describe the data collection methodology. This includes the measurement and collection of memory bus margin at the system assembly factory as well as the data regarding end user memory issues. Section IV provides analysis of the data leading to the conclusion that bus margin is not a dominant source of end user memory issues. Finally, Section V will summarize and conclude the paper.

## II. MEMORY SUBSYSTEM FAILURE CHARACTERIZATION

There are four main sources of memory errors as shown in Figure 2. At a high level, memory subsystem errors can be categorized as:

1. Internal Memory Controller Errors include logic or timing faults inside the memory controller.
2. Internal DRAM Errors include logic faults, timing faults, and cell faults.
3. Bus errors, which occur when one device (memory controller or DRAM) transmits one state but, due to noise or other factors, the other device receives the data in the opposite state. The susceptibility to transient bus errors is commonly measured by bus voltage and timing margin.
4. Soft Errors, which refer to radiation-induced transient events whereby a bit is flipped due to interference from energy sources such as cosmic rays or alpha particles. These are random events, which are very unlikely to repeat and cause an end user DIMM replacement. The primary focus of this paper is to distinguish the relative contributions of the other three sources of memory subsystem errors. The other three sources often appear random but are usually repeatable.

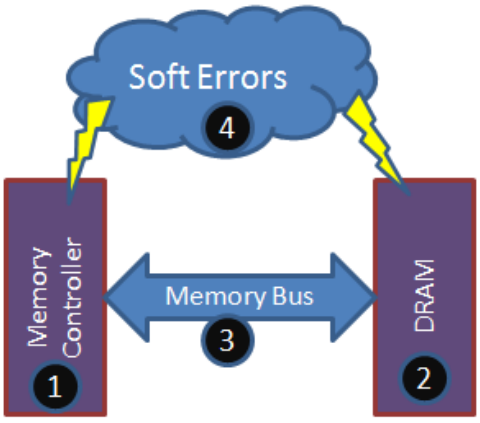


Figure 2. Sources of memory subsystem errors.

### A. DDR3 Bus Margin

There are many sources of noise which can create transient errors on a high speed bus such as DDR3. These include crosstalk, inter symbol interference (ISI), and power delivery issues. In addition, there is manufacturing process variation that impacts the performance of the bus. Examples include impedance variation of the printed circuit board, variation in the nominal supply voltage, and variation in the transmitter and receiver characteristics. To account for noise and high volume manufacturing variation, system designers commonly use the concept of bus margin.

Bus margin, in concept, is a measure of the amount of noise a bus can sustain before an error is induced. For many buses, such as DDR3, this concept is implemented in practice by measuring the voltage and timing margin. The measured margin is then compared against a minimum expectation of margin, or guardband, to account for factors such as different data patterns, high volume manufacturing variation, and other factors not included in the measurement.

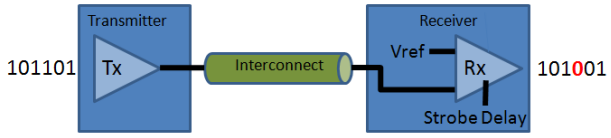


Figure 3. Data (DQ) bus topology for DDR. Vref and Strobe Delay are the bus margining offsets that can be adjusted to produce a bit error.

A typical implementation of measuring bus margin is to transfer a set of patterns over the bus while adjusting either the voltage reference (Vref) or internal timing controls to alter the relationship between the clock and the transmitted or received data, as shown in Figure 3. The voltage or timing is shifted to the point that a data error occurs. The voltage or timing offset required to induce an error establishes the voltage or timing margin for that specific configuration and conditions. Voltage and timing margin are measured for both the positive and negative direction, as shown in Figure 4. For example, the Vref is adjusted up until failure and is also adjusted down until failure. This establishes a high side and low side voltage margin and provides a source of parametric data which can be analyzed. Likewise, the sampling position (strobe delay) is moved both left and right to establish timing margin in both directions. Voltage and timing margins are measured at both ends of the bus, the memory controller in the CPU as well as the DRAM's receiver.



 ISBN: 978-1-61208-168-7 

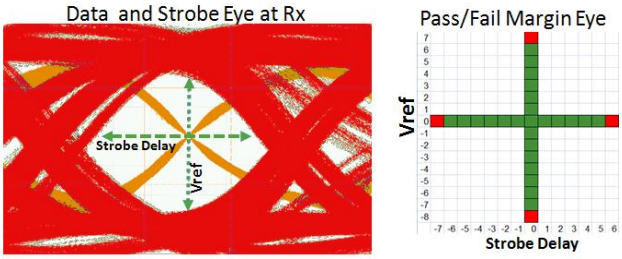


Figure 4. Bus margin example. The left plot shows the data and strobe signals at the Rx pad. The right plot shows the timing and voltage bus margin after adjusting the Vref and strobe delay until a bit error.

Bus margin is a system level metric which can be impacted by multiple factors. Specific areas that can have an impact include: varying characteristics of the transmitters drive strength for the Memory Controller and DIMM, the receivers jitter tolerance, the interconnect, the board, and the connector.

The focus of this paper is on understanding the relative impact of bus margin on the overall population of memory subsystem errors. In this study, the specific factor contributing to low bus margin it is not generally known, however we do know which margin parameter was at risk. More importantly perhaps, the data shows us those issues related to bus margin versus the other possible causes (Internal Memory Controller Error, Internal DRAM Error, or Soft Error).

## III. DATA COLLECTION METHODOLOGY

Figure 5 shows an example of the data collection methodology where a large number of server systems (235,736) were margined during the production test process and then correlated to failures at the customers' site.

### A. Factory Data Collection

Bus margin data used in this study was captured in the system manufacturing and test process at a large server system manufacturer. Using built in test features, the margin data was collected at multiple points throughout the test and stored to a database for future analysis before the system was packaged and shipped to the end customer for installation. Margin data was collected, however, it was not used as a production pass or fail screen. In some cases, other system level tests failed and the memory modules or CPU were replaced but in those cases, the margin data was recaptured after the system was repaired. Only the final data for as-shipped configurations were used for this analysis. Consequently, the margin data in this study represents the actual margin data of the systems as they were shipped out of the factory.

### B. End User DIMM Replacements

Tracking of end user DIMM issues was accomplished through an analysis of service call data for systems manufacturer over a period of 360 days. All systems which required a DIMM replacement in the field were identified and correlated back to the margin data collected for that system in the factory. Note that although these service calls may have included replacement of other system components in addition to memory, such as motherboard or CPU, in all cases, the DIMM was replaced.

| SystemID | End of Line Bus Margin |
|---|---|
| 1 | 130% |
| 2 | 125% |
| 3 | 440% |
| 4 | 220% |
| … | |
| 235736 | 350% |

| Fail Unit System ID | Days in Service |
|---|---|
| X | 120 |
| Y | 200 |
| Z | 58 |

Figure 5. Data collection overview. Failing systems are cross referenced to original "end-of-line" bus margin for analysis.

## IV. DATA ANALYSIS

### A. Factory Margin Distribution

Significant insight can be obtained by analyzing the distribution of observed DDR bus margin across the resulting high volume factory dataset. The resulting data included 6 different server board designs across many different DIMM configurations. Since the data was collected in the factory environment as a study versus a screen, the margins measured represent exactly what the end customer would experience. In other words, the margins were simply measured and logged – a low margin case was shipped 'as is' to the end customer.

Therefore, it is interesting to consider the number of systems that fall below the minimum margin expectation (guardband). The data in Figure 6 shows that only 15 cases out of 235,736 systems were below the guardband which equals a system per million (SPM) of 64. This indicated that about 64 systems out of a million, or 0.0064%, may be at risk of experiencing a bus related error if margins remain the same over time.



 ISBN: 978-1-61208-168-7 

| Case | CPU Voltage Margin | DIMM Voltage Margin | CPU Timing Margin | DIMM Timing Margin |
|---|---|---|---|---|
| 1 | Good | Low | Good | Good |
| 2 | Low | Good | Good | Good |
| 3 | Good | Good | Low | Good |
| 4 | Good | Low | Good | Good |
| 5 | Good | Low | Good | Good |
| 6 | Good | Good | Low | Low |
| 7 | Good | Low | Good | Good |
| 8 | Good | Low | Good | Good |
| 9 | Good | Low | Good | Good |
| 10 | Good | Low | Good | Good |
| 11 | Good | Low | Good | Good |
| 12 | Low | Good | Good | Good |
| 13 | Good | Low | Good | Good |
| 14 | Good | Low | Good | Good |
| 15 | Good | Low | Good | Good |

Figure 6. Cases falling below guardband

Further analysis of the data, shown in Figure 7, indicates that of the 15 systems below guardband, 8 of those were the same DIMM part number/type. These 8 DIMMs were produced in a limited DIMM manufacturing date range of 6 weeks. In fact, 6 of the 8 were in a 3 week period. This particular DIMM represented only 1.6% of the population of DIMMs yet accounted for more than half the low margin cases. This strongly suggests a manufacturing deviation or test hole in the DIMM manufacturing and test process leading to a bus marginality situation. If you were to remove this sub-population of "defective" DIMMs, the effective ratio of systems at risk of bus errors would drop to 0.0030% or 30 SPM.

| Case | Low Margin Description | DIMM Vendor | DIMM Part Number | DIMM Manuf Date |
|---|---|---|---|---|
| 1 | Low voltage margin at DIMM | A | 2 | 2610 |
| 2 | Low voltage margin at CPU | A | 3 | 2210 |
| 3 | Low timing margin at CPU | B | 4 | 5109 |
| 4 | Low voltage margin at DIMM | C | 5 | 410 |
| 5 | Low voltage margin at DIMM | B | 1 | 710 |
| 6 | Low timing margin at CPU & DIMM | B | 6 | 410 |
| 7 | Low voltage margin at DIMM | B | 1 | 1010 |
| 8 | Low voltage margin at DIMM | D | 7 | 5209 |
| 9 | Low voltage margin at DIMM | B | 1 | 1010 |
| 10 | Low voltage margin at DIMM | B | 1 | 1210 |
| 11 | Low voltage margin at DIMM | B | 1 | 1310 |
| 12 | Low voltage margin at CPU | A | 8 | 910 |
| 13 | Low voltage margin at DIMM | B | 1 | 1210 |
| 14 | Low voltage margin at DIMM | B | 1 | 1210 |
| 15 | Low voltage margin at DIMM | B | 1 | 1210 |

Figure 7. Low Margin Cases by DIMM Information. Highlighted rows are the same part number and date code range of week 7-12, 2010 indicating a DIMM manufacturing excursion.

Consider this low percentage of systems at risk of bus errors (0.0030%) compared to either the 30% of systems experiencing memory errors in one large scale study [1] or a more commonly expected rate of 10%-15%. Note in the referenced study [1] that these systems experiencing errors have a median number of errors between 25 and 611 per year. Given the random nature of Soft Errors, there is strong evidence that these are instead related to one of the other three sources. The margin data also suggests that relatively few systems should experience bus errors which would indicate that the bulk of end user memory errors are likely not Soft or Bus Errors, but instead either Internal Memory Controller or Internal DRAM Errors. This of course assumes that the populations of systems from the two studies were similar. We'll explore this from another angle by looking at service call data for DIMM replacements.

### B. *End User DIMM Replacement Analysis*

The prior analysis was done against systems that had low margin and were at risk to fail. Next we will consider systems that actually did experience some form of memory error in the field. In this analysis, we will study systems that required a DIMM module replacement at the end customer installation.

The systems which required DIMM replacement were cross-referenced back to the bus margin data collected for that specific system when it was tested at the factory. The bus margin for these systems was then compared against the minimum bus margin guardband expectation. The data in Figure 8 shows that only a small proportion of the systems requiring DIMM replacements contained low margins at the time the system was shipped from the factory. Only 0.16% of the systems were below the margin guardband and in fact, even if you double the guardband, this would still be less than 1% of systems.

Clearly, the margins on the memory bus are a minor factor driving field DIMM replacements for the population of systems under study. What is driving these replacements then? Unfortunately, detailed failure analysis was not possible for the failures returned from customer sites, but we can use the data we have to draw some important conclusions. The factory bus margin data indicates that both the CPU and DRAMs have sufficient voltage and timing margin to ensure robust data transfers. Assuming that bus margins have not degraded over time, there is a strong indication that signal integrity issues are not a major factor in memory failure rates. Considering that the sub-population of systems requiring a DIMM replacement included 32 unique DIMM part numbers across 5 different vendors, margin degradation due to aging seems unlikely. While it might be reasonable to assume that a particular DRAM design might have degradation problems due to aging, it is



 ISBN: 978-1-61208-168-7 

very unlikely that this is a widespread problem across so many different part numbers and DRAM suppliers. What about margin degradation due to aging of the CPU? The fact that the DIMM is being replaced and thereby resolving the issue contradicts this theory and indicates it is not the CPU.

Given that these DIMM failures don't correlate to low bus margins as measured in the factory and it seems unlikely that DRAM I/O degradation is a widespread problem, it is assumed that these DIMM replacements are largely driven by internal DRAM issues.

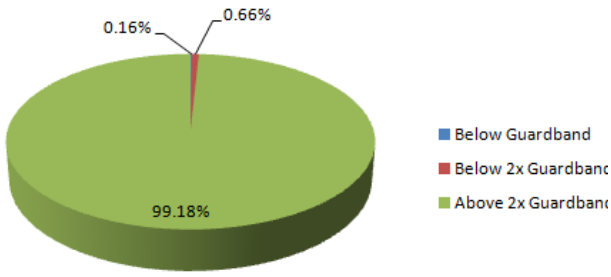


Figure 8. Chart shows the percentage of systems with low bus margin. This is for all field memory failures within our data set.

## V. CONCLUSIONS

Through a large scale study of almost a quarter of a million systems and over a million DIMMs, we have found that memory bus margin is a minor contributor to memory subsystem issues which drive end user DIMM replacements. The key data supporting this conclusion:

- Prior studies indicate that somewhere between 5 and 30% of systems experience memory issues, yet high volume margin data collected at the system assembly factory suggests that only about 0.0064% of systems would be susceptible to experiencing problems due to bus margin.
- Only 0.16% of the systems that required a DIMM replacement showed low bus margin in the factory study.
- Bus margin degradation over time is unlikely given that the population of DIMM replacements includes a large variety of different DIMMs from 5 different DRAM manufacturers, eliminating any systemic problems.

It should be noted that this data was from a specific CPU family and set of product design requirements. The data suggests that the systems are well designed from a bus integrity point of view. It is possible that other products may have inferior bus designs, higher memory error rates, and higher proportion of those memory errors attributable to bus marginality. Also, further aging studies including contact corrosion and degradation are under investigation to better understand how bus margins may change over several years. However, for a well-designed system, the data shows that bus marginality is a very small contributor to overall memory subsystem health.

## REFERENCES

[1] B. Schroeder, E. Pinheiro, and W. Weber. "DRAM Errors in the Wild: A Large-Scale Field Study", *SIGMETRICS/Performance '09*, June 15-19, 2009, Seattle, WA, USA

[2] T. C. May and M. H. Woods, "Alpha-Particle-Induced Soft Errors in Dynamic Memories", IEEE Transactions on Electron Devices 26, No. 1, 2-9, 1979

[3] J. F. Ziegler and W. A. Lanford, "Effect of Cosmic Rays on Computer Memories", Science 206, No. 4420, 776-788, 1979

[4] J. F. Ziegler and W. A. Lanford, "The Effect of Sea Level Cosmic Rays on Electronic Devices", IEEE International Solid-State Circuits Conference, 1980

[5] S. Mukherjee. "Computer Glitches from Radiation: A Problem with Multiple Solutions", Microprocessor Report, May 19, 2008

[6] T. J. Dell, "System RAS implications of DRAM soft errors", IBM Journal of Research and Development, Vol. 52, No.3, May 2008

[7] X. Li, M. C. Huang, K. Shen, and L. Chu, "An Empirical Study of Memory Hardware Errors in A Server Farm", HotDep Workshop, 2007

[8] "Soft Errors in Electronic Memory – A White Paper", Tezzaron Semiconductor, January 5, 2004, Naperville, IL, USA



 ISBN: 978-1-61208-168-7 