Data Analysis of Manufacturing Test Results for DRAM Module

White Paper

March 2018

Contents

Introduction

Quality, cost and production capacity are always the top priorities for manufacturing. It is important to optimize all three aspects simultaneously and continuously.

As a major hyperscale data center solution provider, Wiwynn provides Level 11 solutions and accumulates a huge amount of data through daily production testing procedure. After extensive experience in integrating and testing more than 20,000 racks over the past years, we are now working on data analysis to improve testing time and key component quality.

This whitepaper shows the preliminary statistic data regarding the quality of DIMMs, SSDs and HDDs. Furthermore, we will dive into DRAM module testing data and see how to optimize DIMM test process.

Quality Analysis of DIMMs, SSDs and HDDs by Vendor and by Capacity

DIMMs

We collected more than 3 million data for DIMMs through daily production testing process in 2017 and analyzed these data by vendor and by capacity. As shown in *Table1*, the defect rate of 8GB DIMMs is lower than 16GB and 32GB modules while vendor B has the highest defect rates for 8GB and 32GB.

Table 1: DIMM defect rates by capacity and by vendor

Capacity		Vendor A	Vendor B	Vendor C	Total
32GB	Test Qty	1M-2M	0.1M-0.5M	0.1M-0.5M	2M-3M
	Defect rate	0.12%	0.15%	0.06%	0.12%
16GB	Test Qty	0.1M-0.5M	N/A	<0.1M	0.1M-0.5M
	Defect rate	0.07%	N/A	0.08%	0.07%
8GB	Test Qty	0.1M-0.5M	0.1M-0.5M	0.1M-0.5M	0.5-1M
	Defect rate	0.06%	0.07%	0.05%	0.06%

SSDs

Defect data of more than 770k flash type storage are shown in this part, including M.2/U.2 and SAS/SATA SSDs. *Table2* shows that among the five vendors, E2 has the lowest defect rate.

Table 2: Flash storage defect rate of each vendor

	A2	B2	C2	D2	E2
Test Qty	>10K	>400K	>50K	>50K	>100K
Defect rate	0.08%	0.11%	0.10%	0.14%	0.05%

HDDs

Table3 shows the defect rate of each vendor for more than 2 million hard drives.

Table 3: Hard disk defect rate of each vendor

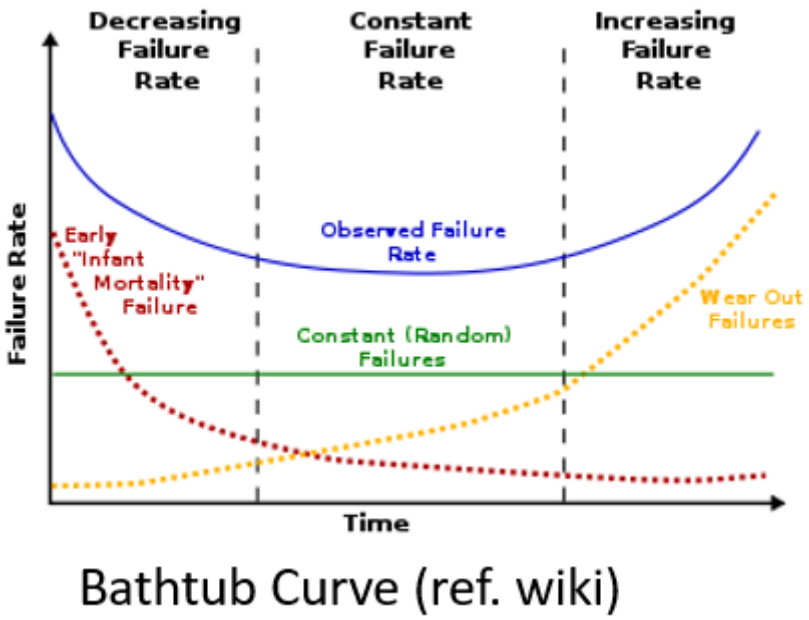
	A3	B3	C3
Test Qty	0.5M-1M	1M-2M	<0.1M
Defect rate	0.07%	0.042%	0.039%

Based on our experience, the reliability engineering of DIMMs has greater impact on testing process improvement but there is no golden testing time and criteria. To improve the manufacturing efficiency, the rest of this whitepaper will be focused on the relationship between the testing time and the criteria of defect DIMMs.

DIMM Quality with Correctable ECC Error

In recent years, studies on memory reliability and its relation to ECC error rates in datacenters revealed that the frequency of correctable ECC error can be regarded as an indicator for the quality of DIMMs. Some studies have identified that an unstable system has more than 100 correctable ECC errors in a week (Ref. 1 and Ref. 2).

Wiwynn referenced this result in designing our quality testing process. We set a zero-tolerance for uncorrectable ECC errors and a threshold for correctable ECC errors (E_{max}) within a given time(t_d) to verify the stability of DIMMs* in the early life cycle of DIMMs in bathtub curve (*Figure1*). This was applied to thousands of our server system racks and has proven to be an effective method**. From here on, ECC errors discussed only refer to correctable ones.**



Bathtub Curve (ref. wiki)

Figure 1: Bathtub Curve

After accumulating huge amounts of DIMM testing data, we further improved the testing time by analyzing the relationship between the numbers of correctable ECC errors and the testing time. We will discuss the distribution of time point for ECC error occurrence in different orders and analyze how to shorten the testing time while still keeping the catch rate (Catch Rate = Defect DIMMs judged by the new criteria / Defect DIMMs judged by the original criteria).

*We use Google's memory stressing tool - stressapptest (version 1.0.3) to test ECC DIMM.

First Correctable ECC Error Occurrence Time Point Distribution for DIMMs

To begin with, we catch the system event log (SEL) data to observe the timing of the first correctable ECC error among DIMMs with ECC errors from our sample data set. As shown in *Figure 2*, the *x* axis is the normalized time and the *y* axis is the count of DIMMs having their first ECC errors occurring at that time**. We normalized the total test time in field to 1000 time units for this analysis.**

From *Figure 2*, we found that the majority of the first ECC error occurs **within the 200 normalized time point** with a long tail going downward until the end of the testing.

In addition, the first ECC error time distribution fits in a continuous probability distribution curve with **Weibull** distribution. As shown in *Figure 3*, β is the time point (311 normalized time point) when 80% of DIMMs have their first ECC errors (accumulate from time = 0) while γ is the time point covering 90% of DIMMS with first ECC errors.

In spite of the long-tail, we estimated that 90% of the first DIMM ECC errors occurred within the 444 normalized time point. In other words, *90%* of the DIMMs with ECC errors burst their first ECC errors within the 444 normalized time point.

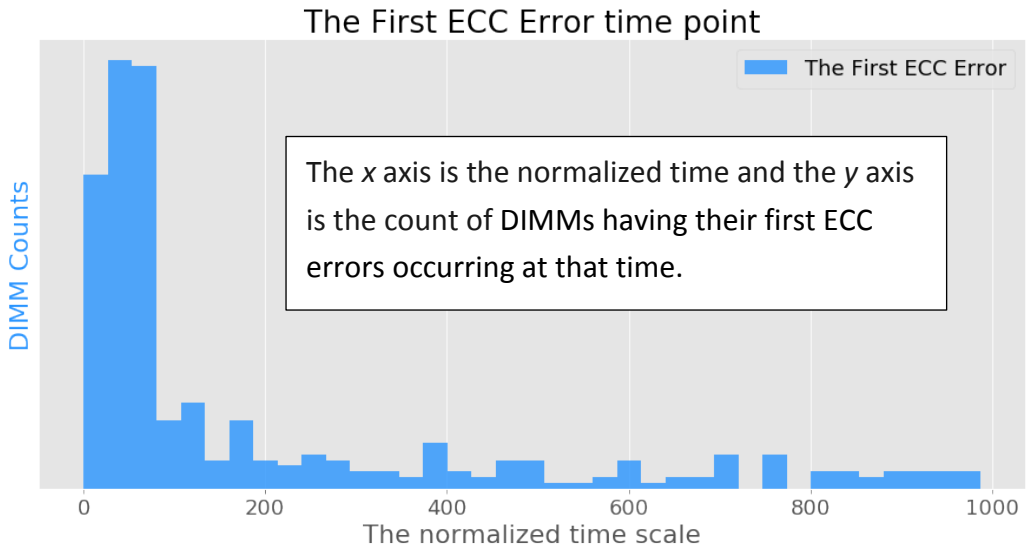


Figure 2: First ECC error occurrence time point distribution of DIMMs.

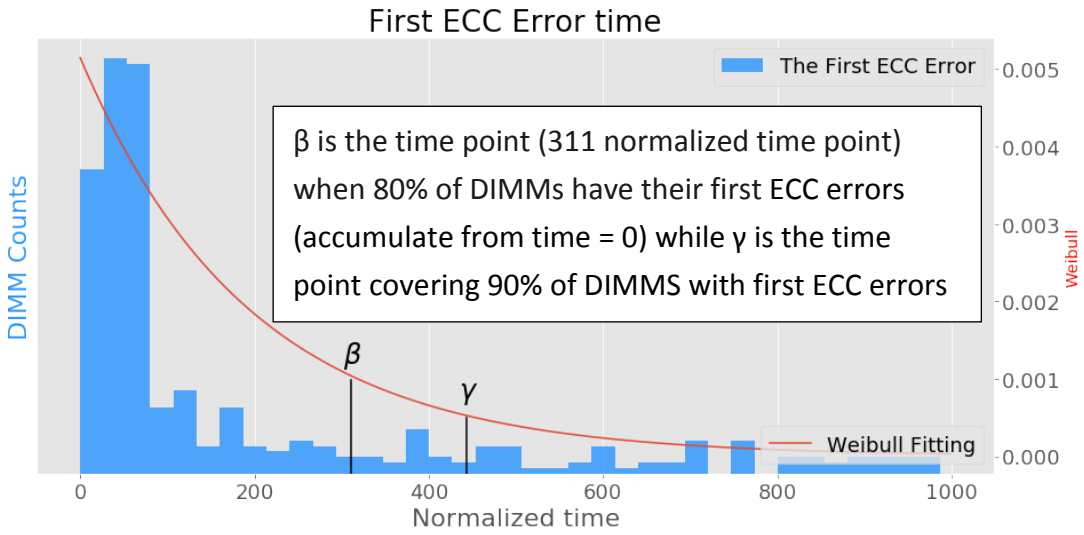


Figure 3: The Weibull distribution curve fitting on the ECC error distribution.

Characteristics of Different Correctable ECC Error Types

To further analyze the relationship between the ECC error frequency and the DIMM failure, we observed the ECC error occurrence patterns of the defective DIMMS after their first ECC errors occurred.

We collected the time interval data between the first ECC error occurrence time point and the E_{max}^{th} ECC error occurrence time point of each defective DIMM~~s~~ from SEL and plotted as in *Figure 4*.

We then used the average-linkage hierarchical clustering to analyze the error time segment and found that a big portion of DIMMs reach their E_{max}^{th} ECC error occurrence within 3.5 normalized time units after their first ECC error occurrence. We then categorized these DIMMs as **Spike** type, and use 3.5 normalized time units as the time bucket for classification. On the other hand, the other type, called **Sparse**, describes DIMMs with low ECC error occurrence frequency. It takes minutes or hours for Sparse DIMMs to have one ECC error occurrence.

In other words, Spike presents DIMMs with higher error occurrence frequency and can be identified as defective items earlier compared with Sparse which has lower error occurrence frequency.

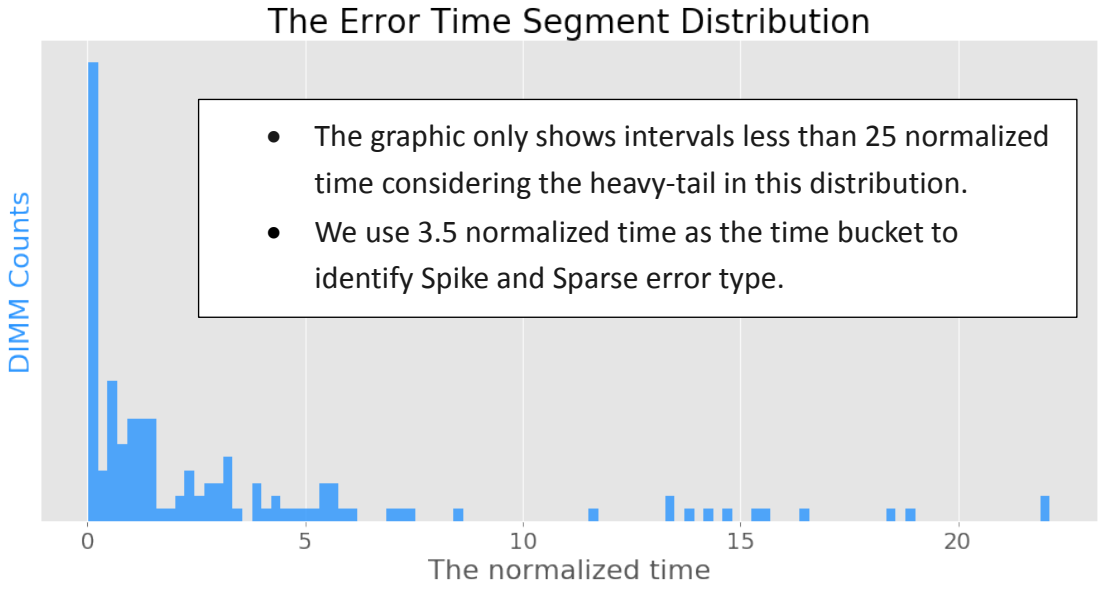


Figure 4: The error time segment distribution:
Distribution of time interval between the first ECC error occurrence time point and the Emax th ECC error occurrence time point of each defective DIMM

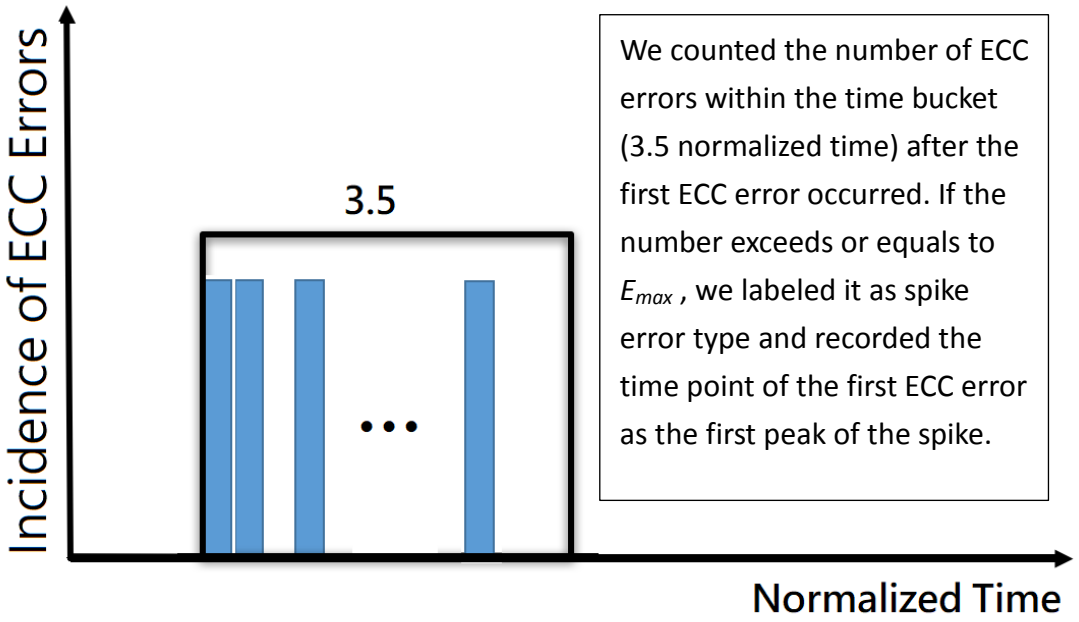


Figure 5: An illustration of the implementation of the error type labeling

As shown in *Table 4*, the percentages of Spike and Sparse are approximate half for each. We will then further explore the behavior of the DIMM Spike and Sparse types and try to estimate the defect coverage.

Table 4. Percentages of DIMM Spike and Sparse types

ECC error type	Spike	Sparse
Percentage	52.55%	47.45%

Spike Type Assessment

Generally speaking, we can quickly identify Spike type DIMMs as defective once the first ECC error occurred since the E_{max}^{th} ECC error will then occur within 3.5 normalized time units. When we go back to check the first ECC error occurrence time point distribution of Spike type DIMMS, we can find it fits in the Weibull distribution and use it to evaluate possible testing time reduction.

As shown in *Figure 6*, the normalized time point *β* (normalized time=491) covers 80% of the Spike DIMMs with the first ECC error occurrence while 90% are covered at normalized time point *γ* (normalized time=702). After normalized time point *β*, the

spike incidence starts to converge. However, it takes a long time to increase the coverage considering the long tail pattern.

We can then build a model to evaluate the relationship between the coverage and the testing time to strike a balanced time point for both effective testing time and reasonable defect coverage.

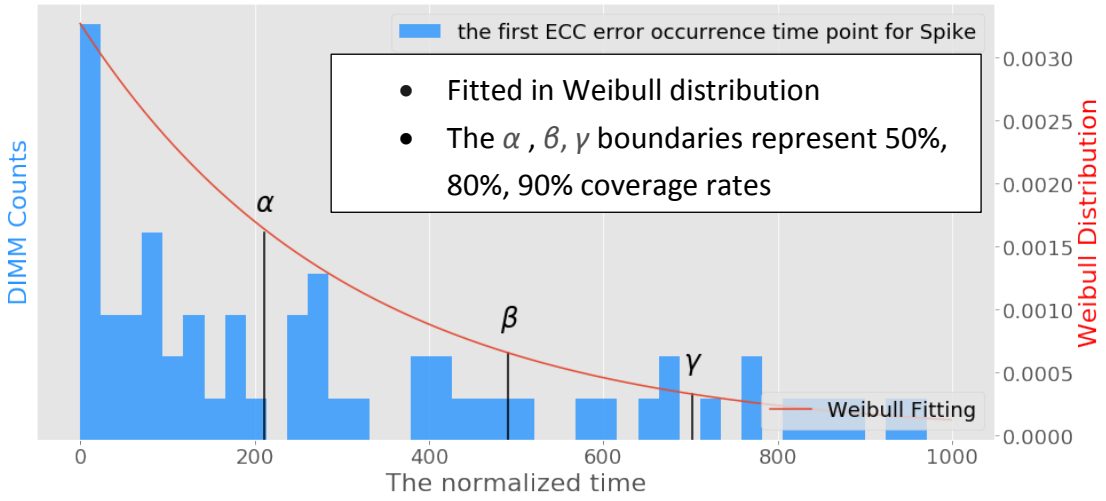


Figure 6: The distribution of the first ECC error occurrence time point for Spike type DIMMs

Sparse Type Assessment

Sparse is the type of defective DIMM with low ECC error occurrence frequency after the first ECC error occurred. It takes longer for this type to reach the ECC error threshold.

In this part, we will examine Sparse type DIMMs with the distribution of the first ECC error occurrence time point as well as the E_{max}^{th} ECC error occurrence time point distribution. The distribution pattern and the comparison with Spike type DIMMs will be discussed.

As shown in *Figure 7*, 80% of the first ECC errors of Sparse type DIMMs occurred within the normalized time point 423, which is approximately the same as Spike type DIMMs.

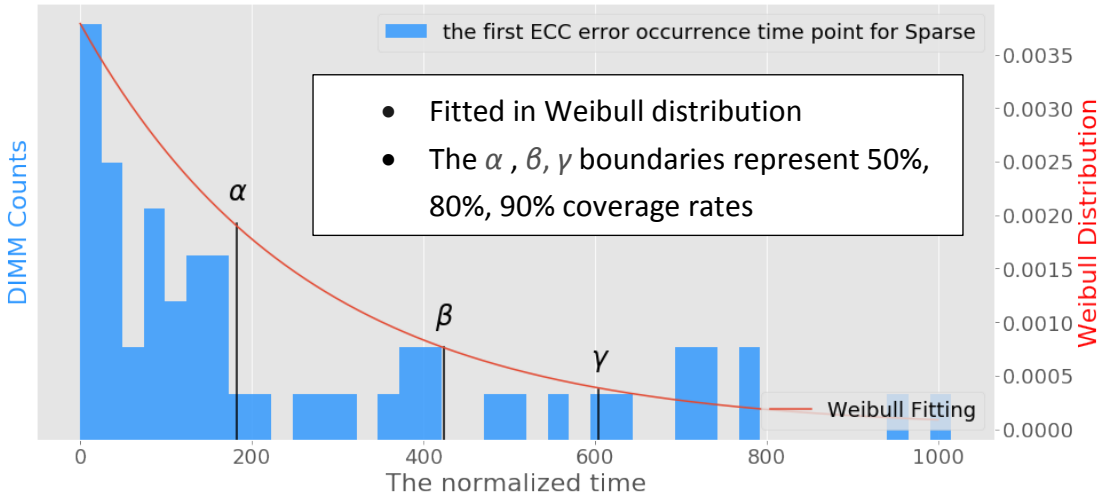


Figure 7: The first ECC error occurrence time point distribution for sparse type

From *Figure8*, we can see that 80% of Sparse type DIMMs reached the E_{max}^{th} ECC error by normalized time point 560. It is around 137 normalized time or 24.5% more compared with Spike type DIMMs. Since Sparse type DIMMs need more time to reach the E_{max}^{th} ECC error, some defective DIMMs may be missed if we shorten testing time. Therefore, the threshold should be raised to keep the filtering rate. However, another side-effect, named **false alarm**, would occur along with the raising threshold. Some non-defective DIMMs may be recognized as defects while we implement stricter ECC error threshold. We will discuss the related behaviors when adjusting the error threshold in the following chapter.

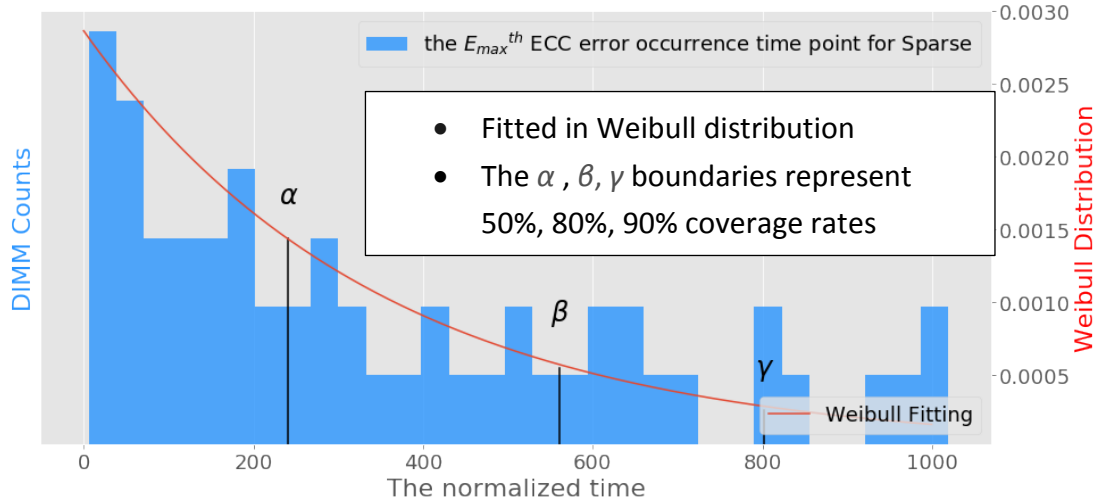


Figure 8 The E_{max}^{th} ECC error occurrence time point distribution for Sparse type

Discussion

Since stress test of DIMMs accounts for a significant portion of manufacturing overhead, it is important to optimize the process with less time and more reasonable filtering index. It is straightforward to just cut testing time and raise the threshold as compensation. However, the benefits of less test time will be offset by less defect coverage and increased false alarms. Therefore, we design a cost-response model based on exponential growth model with consideration of cost of various testing time, RMA, false alarms, operational cost for datacenters and customers' reputation loss.

From the analysis of ECC error occurrence time distribution for Spike and Sparse type DIMMs, we found that we can keep considerable catch rate for unstable DIMMs (Catch Rate = Defect DIMMs judged by the new criteria / Defect DIMMs judged by the original criteria) in the early testing stage. We can then adjust threshold to reduce the decrease of catch rate. Both Spike and Sparse type DIMMs can be applied in this new model.

We analyzed various combinations of test time and ECC error threshold with our cost model. It shows that these two factors are inverse with non-linear relationship. Furthermore, we found some combinations that could keep catch rate at 90% and 80% from direct verification data and predictions based on Weibull distribution as shown in *Table5.* With further comparison of the direct verification data and predictions, we found that data from the distribution model is more conservative with 90% catch rate while the result for 80% catch rate are similar among actual data and prediction.

For cross verification, we prepared another hold-out dataset of 216,808 32GB DIMMs from our factory in different time periods. The testing time and the ECC error threshold to keep 90% and 80% catch rates are shown in *Table 5.* We can see that the cross verification result is even closer to the distribution model. Moreover, the verification of training dataset, hold-out dataset and distribution model all show that it requires less than 0.25 t_d testing time to reach 50% catch rate. The model proved that it can be duplicated in the field testing process.

Furthermore, we observed that Sparse type DIMMs are more sensitive to changes in the ECC error threshold. We will adjust the parameters and model with more RAM data coming from customers.

Table 5: Testing time and ECC Error threshold combination by catch rate

	Catch rate	Testing time	ECC Error threshold	False alarm rate
Direct verification data	90%	0.75 t_d	0.5 E_{max}	0.016%
	80%	0.5 t_d	0.33 E_{max}	0.028%
Cross verification data	90%	0.68 t_d	0.5 E_{max}	0.036%
	80%	0.51 t_d	0.5 E_{max}	0.025%
Prediction based on Weibull distribution	90%	0.7 t_d	0.5 E_{max}	NA
	80%	0.49 t_d	0.5 E_{max}	NA

Conclusion

Based on the study described above, we can simplify test process and also assure quality.

We implement predictive analytics by analyzing event logs generated from the manufacturing process. We can reduce the number of required tests and find the best effective stress test time for different parts and brands. For example, we can predict what type or brand of DIMM modules are more reliable and what is the minimum hours required to verify the quality based on ECC error detection on the same criteria.

With the optimized test process, we can improve cost and capacity but still keep high quality level.

Terminology and Reference

Terminology

- Testing Time, ECC Error Threshold : (t_d , E_{max})
- Defect DIMMs Judged by The Original Criteria (t_d , E_{max}): D_o
- Not Defect DIMMs Judged by The Original Criteria (t_d , E_{max}): N_o
- **Catch Rate =** (Defect DIMMs Judged by New Criteria and Existed in D_o) / D_o
- **False Alarm Rate =** (Defect DIMMs Judged by New Criteria and Existed in N_o) / N_o

Reference

1. Meza, Justin, et al. "Revisiting memory errors in large-scale production data centers: Analysis and modeling of new trends from the field." *Dependable Systems and Networks (DSN), 2015 45th Annual IEEE/IFIP International Conference on*. IEEE, 2015.

2. Sridharan, Vilas, and Dean Liberty. "A study of DRAM failures in the field." *High Performance Computing, Networking, Storage and Analysis (SC), 2012 International Conference for*. IEEE, 2012.

About Wiwynn

Wiwynn® is an innovative cloud IT infrastructure provider of high quality computing and storage products, plus rack solutions for leading data centers. We aggressively invest in next generation technologies for workload optimization and best TCO (Total Cost of Ownership). As an OCP (Open Compute Project) solution provider and platinum member, Wiwynn® actively participates in advanced computing and storage system designs while constantly implementing the benefits of OCP into traditional data centers.

For more information, please visit http://www.wiwynn.com/english or contact sales@wiwynn.com
Follow Wiwynn® on Facebook and Linkedin for the latest news and market trends.