In-Depth Reliability Characterization of NAND Flash based Solid State Drives in High Performance Computing Systems

Extended Abstract

Shuwen Liang, Zhi Qiao and Song Fu
Department of Computer Science and Engineering, University of North Texas
{shuwenliang,zhiqiao}@my.unt.edu,song.fu@unt.edu

Weisong Shi
Department of Computer Science, Wayne State University
weisong@wayne.edu

ABSTRACT

NAND flash based solid state drives (SSD) have been widely used in high performance computing (HPC) systems due to their better performance compared with the traditional hard disk drives. However, little is known about the reliability characteristics of SSDs in production systems. Existing works study the statistical distributions of SSD failures in the field. They do not go deep into SSD drives and investigate the unique error types and health dynamics that distinguish SSDs from hard disk drives. In this paper, we explore the SSD-specific SMART (Self-Monitoring, Analysis, and Reporting Technology) attributes to conduct an in-depth analysis of SSD reliability in a production datacenter. Our dataset contains half a million records with more than twenty attributes. We leverage machine learning technologies, specifically data clustering and correlation analysis methods, to discover groups of SSDs which have different health status and the relation among SSD-specific SMART attributes. Our results show that 1) Media wear affects the reliability of SSDs more than any other factors, and 2) SSDs transit from a health group to another which infers the reliability degradation of those drives. To the best of our knowledge, this is the first study investigating SSD-specific SMART data to characterize SSD reliability in a production environment.

1 INTRODUCTION

Solid state drives (SSDs) storage systems are receiving wide attention for high performance computing (HPC) and their deployment is steadily increasing due to their higher performance and lower power consumption compared with hard disk drives (HDDs) storage systems. While their deployment is increasing, the write endurance of SSDs still remains as one of the main concerns. As write and erase operations on an SSD wear it out gradually, after a certain number of operations, the SSD could fail and its data could be lost. Many studies have investigated the bit error failure behavior of multi-level cells (MLC) and single-level cells (SLC). They point out that the bit error rate of the flash memory increases with increased number of Program/Erase (P/E) cycles. These studies model the bit error rate as an exponential function of the number of P/E cycles the cell has gone through. There are also a number of recent studies that analyze the statistical distributions of SSD failures in the field. They also find that although flash drives offer lower field replacement rates than HDDs, they have a significantly higher rate of uncorrectable errors that can impact the user and the stored data.

The existing studies of SSD reliability are either at the circuit level (i.e., MLC and SLC) or at the system level using field data. They do not explore the rich set of performance and reliability related attributes provided by the SSD Self-Monitoring, Analysis, and Reporting Technology (SMART) at the drive level. In this paper, we perform an in-depth analysis of SSD reliability by using the SSD-specific SMART data collected from an active production datacenter. We investigate more than 20 SSD-specific performance and error related attributes from over half a million complete records. By using machine learning methods, we discover groups of SSD SMART records corresponding to different health status. We observe the transition of several SSDs from one group to another which represents the change of their health and the degradation of their reliability. Moreover, results from the correlation analysis show that some attributes, i.e., media wear which reports the number of cycles that the NAND media has undergone, affects the reliability of SSDs more than other attributes.

To the best of our knowledge, this is the first study that investigates SSD-specific SMART data to characterize SSD reliability in a production environment. Our analytic results provide a deeper understanding of SSD reliability and its dynamics. They can also enable system operators to develop countermeasures to extend SSD lifetime or protect the stored data for example by balancing I/O workload, predicting SSD failures, and proactive migration of data.

2 METHODOLOGY AND EXPERIMENTAL RESULTS

Our dataset is collected from a production datacenter in which SSD is organized as a data buffer between the main memory and the storage using HDD. The dataset contains six months of SMART data files pulled hourly from every SSD. We pre-process the raw SMART files and extract the values of SMART attributes. Half a million records of 20 attributes with raw values are produced. we first analyze the relation among these attributes, and then explore machine learning technologies to characterize SSD reliability.

2.1 SSD-Specific SMART Attributes

SSD SMART provides many attributes that are particularly designed for SSDs. We category the attributes into two groups, i.e., environmental factors (such as temperature and runtime) and I/O related factors, as shown in the following table.

2.2 Correlation among SSD Attributes

By analyzing the correlation among SSD SMART attributes, we obtain a better understanding of the influence among various factors and their criticalness for characterizing SSD reliability. Figure 1 shows the pair-wise correlation calculated by using Pearson's coefficient on the SSD SMART dataset.

Table 1: SSD SMART attributes.

ATTRIBUTE ID	ATTRIBUTE NAME	DESCRIPTION
Environmental Attributes		
9	POH	Power On Hour
12	PCC	Power Cycle Count
194	TC	Temperature Celsius
I/O Related Attributes		
5	RSC	Reallocated Sector Count
166	MWEC	Min Write/Erase Count
167	MBB	Min Bad Block/Die
168	MEC	Max Erase Count
169	TBB	Total Bad Block
171	PFC	Program Fail Count
172	EFC	Erase Fail Count
173	AWEC	Average Write/Erase Count
174	UPLC	Unexpected Power Lost Count
187	RU	Reported Uncorrect
212	SPE	SATA PHY Error
230	PWEC	Percentage Write/Erase Count
232	PARS	Percentage Avaliable Reserved Space
233	TNWG	Total NAND Write(GB)
241	TWG	Total Write(GB)
242	TRG	Total Read(GB)
243	Unknown	N/A

Figure 1: Correlations among SSD SMART attributes.

From Figure 1, we have the following findings. 1) Write and erase operations have a strong correlation between each other. 2) I/O operations are not distributed evenly in the system. 3) Environmental attributes do not directly affect SSD health. 4) Although wear bearing SSDs are supposed to have a higher number of Total Back Blocks (TBB), no evidence shows that I/O operations lead to increase of back blocks in SSDs.

In Figure 1, we also observe that those attributes that are related to write and erase operations, such as MEC, WEC, PWEC, TNWG and TWG, have a positive correlation higher than 0.9 between each other. In addition, MBB and TBB has a correlation of 0.8, the same as PCC and UPLC.

2.3 SSD Reliability Characteristics

I/O operations, including read, write and erase, can influence the health status of SSDs. We analyze the categories of SSD health and their possible transitions by using data clustering methods. Specifically, we explore K-Means clustering on all SMART records in the dataset. Experimental results show that the SSDs are grouped into five clusters.

More importantly, the health of an SSD does not stay in one group forever. It can transit from one group to another as the wear level changes. In some cases, such transitions happen more than once (A maximum of three transitions is observed in our study). Over 83% of SSDs experience the health status transition, which we call *reliability degradation*. The following table presents the relative size of each health cluster and the frequency of reliability degradation.

Table 2: Groups of SSD health and transitions.

CATEGORY	CLUSTER	PERCENTAGE of SSDs in CATEGORY(%)
Cluster	Cluster 0	2.0%
	Cluster 1	1.3%
	Cluster 2	2.0%
	Cluster 3	10.7%
	Cluster 4	0.7%
Cluster Transition	Cluster 2->4	17.3%
	Cluster 4->0	27.3%
	Cluster 2->4->0	38.0%
	Cluster 2->4->0->3	0.7%

To analyze these SSD wear levels, we investigate each cluster produced by K-Means clustering. We find that 1) SSDs in Cluster 1 experience I/O intensive operations. The number of read operations is the highest, and the number of write and erase operations are also relatively high. 2) SSDs in Cluster 3 experience the highest number of write and erase operations, while the number of read operation is the average. (3)Clusters 0, 2 and 4 include the majority of SSDs which experience the average number of I/O operations. However, SSDs in the three clusters have reliability degradation following the same transition pattern, i.e., Cluster 2 -> Cluster 4 -> Cluster 0. Along this transition, the value of TBB decreases while the number of write and erase operations increases.

Based on the preceding findings, we infer that SSDs in Clusters 1 and 3 have a worse health status than other SSDs which are still in good shape. Those good SSDs will experience reliability degradation, i.e., transition to Clusters 1 and 3, as more I/O operations and cycles cause wear to those drives.

3 CONCLUSIONS

We study SSD SMART data collected from an active productive datacenter. We observe that SSD has many unique attributes compared with HDD. By analyzing these SSD-specific attributes, we find that they are highly useful for characterizing and modeling the health status SSDs health status even in real time. Our analytic results show that the volume of I/O operations and cycles has a major influence on the wear level of SSD. Write and erase related attributes display a strong correlation. Read operation is relatively independent and is not evenly distributed in among the drives. We observe many health status transitions in our analysis.

ACKNOWLEDGMENT

This work is supported in part by NSF grants CCF-1563728 and CCF-1563750.

REFERENCES

[1] Song Huang, Song Fu, Quan Zhang, and Weisong Shi. 2015. Characterizing disk failures with quantified disk degradation signatures: An early experience. In *Proceeding of IISWC*. IEEE, 150–159.