A Study of Soft Error Consequences in Hard Disk Drives

Timothy Tsai
Hitachi Global Storage Technologies
timothy.tsai@hgst.com

Nawanol Theera-Ampornpunt, Saurabh Bagchi
Purdue University
{*ntheeraa, sbagchi*}*@purdue.edu*

***Abstract*—Hard disk drives have multiple layers of fault tolerance mechanisms that protect against data loss. However, a few failures occasionally breach the entire set of mechanisms. To prevent such scenarios, we rely on failure prediction mechanisms to raise alarms with sufficient warning to allow the at-risk data to be copied to a safe location. A common failure prediction technique monitors the occurrence of soft errors and triggers an alarm when the soft error rate exceeds a specified threshold. This study uses data collected from over 50,000 customer deployed disk drives to evaluate the performance of a failure prediction algorithm that relies solely on soft errors to predict failures manifested as hard errors. The data analysis shows that soft errors alone cannot be used as a reliable predictor of hard errors. However, in those cases where soft errors do accurately predict hard errors, sufficient warning time exists for preventive actions.**

***Keywords*-hard disk drive; failure prediction; soft errors; hard errors; data mining;**

I. Introduction

As a storage device for persistent data that is often critical, a hard disk drive must provide a high level of reliability. Due to the inherent complexity of a device that incorporates mechanical, electronic, and other parts, a wide assortment of failure modes are possible and must be addressed [1]. When stored data is requested, the disk drive must be able to provide that data with complete fidelity and in a timely manner. To achieve these goals, the architecture and design of disk drives incorporate many fault avoidance and fault tolerance techniques. Multiple levels of cyclic redundancy codes and error correction codes protect against data loss by either correcting a small number of bit errors on the fly or by detecting greater corruption and initiating read retries. These read retries are based on a carefully engineered sequence of retry steps that are ordered to minimize the expected overall latency. Since most errors will be corrected in the initial steps, a typical distribution of steps at which errors are corrected will have a shape similar to that shown in Figure 1.

A small number of errors will be uncorrectable even after the drive exhausts the entire sequence of error recovery steps. These uncorrectable errors are often described as "hard" errors to distinguish them from the "soft" errors that are successfully corrected during one of the error recovery steps. Since the goal of a disk drive is to store data for retrieval at some future time, a hard read error is a device failure. To address such hard read errors, higher-level fault tolerance may be added at the hardware level (e.g., RAID [2]) or at the software level. However, not all systems can accept the added cost, complexity, or latency imposed by such higher-level fault tolerance.

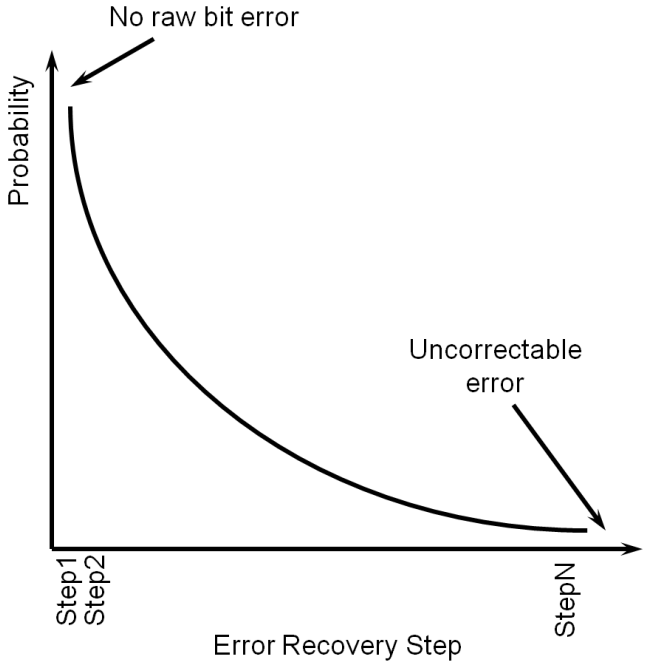

Figure 1: Typical distribution of successful error recovery steps

Failure prediction is another approach that attempts to mitigate the potential loss of data due to a hard read error by providing sufficient advance notice for emergency measures such as duplication of at-risk data. The key to failure prediction is the determination of failure predictors that are well correlated to imminent failure while providing sufficient time for preventive actions. Good predictors have high detection rate coupled with low false positive rates. In particular, the false positive rate must be minimized to limit the impact on performance since each false positive prediction incurs a nonessential preventive action, such as, remapping some sectors of the drive, which in some situations may involve lengthy diagnostics.

Two aspects of failure prediction are essential and must be carefully designed: the raw input data and the algorithm that transforms the raw data into alarms. For disk drives, the two main types of data are sensor data and event data. A disk drive contains many sensors to detect abnormal operating conditions. Examples of monitored conditions are temperature, physical shock and vibration, data channel signal-to-noise ratios. These sensors often provide direct feedback to the disk drive hardware and firmware but can also be incorporated into the algorithm that raises external alarms.

In addition to sensor data, the firmware can also generate event data based on events of interest, such as errors, retries, power-on, sector reallocation, etc. In contrast to sensor data which is limited by the set of physical sensors present in the system, the amount of event data is virtually unlimited since the firmware can be easily programmed to recognize additional events. Another important difference between sensor data and event data is the relative obviousness of how the data should be used to raise an alarm. The thresholds for abnormal sensor data are often somewhat intuitive. For example, if the temperature exceeds the design specification, then certainly an alarm should be raised. However, many innocuous events occur frequently (e.g., raw read error rates can be as high as 1 in 104 bits in a good drive), so it is not entirely clear how these events should be translated into alarms. *What is particularly lacking is an understanding of the relationship between events and resulting failure modes.*

The main goal of this study is to evaluate the performance of a failure prediction algorithm that relies solely on soft error events in predicting failures that are manifested as hard errors. We hope the results will aid the feature selection process when building a failure prediction algorithm that has a richer set of features at its disposal. The results show that many hard errors occur without a single preceding soft error, thus indicating a limit to the predictive ability of soft errors. However, in those cases where soft errors do precede hard errors, the data suggests the possibility of developing algorithms that raise alarms of imminent hard errors with sufficient advance warning to permit actions that prevent data loss.

One of the key contributions of this paper is the analysis of data collected by hard disk drive firmware which include knowledge of and details about events that are not available from data collected on the other side of the hard disk drive host interface. Several excellent studies have analyzed data collected by storage subsystems containing a large number of drives [3][4][5]. While hard error events are usually reported across the host interface, soft errors by default are not reported, since the sheer number of soft errors would significantly degrade performance. However, disk drives can be configured to report soft errors [6], although the set of reported soft errors is often a subset of the internally collected errors for several reasons. First,

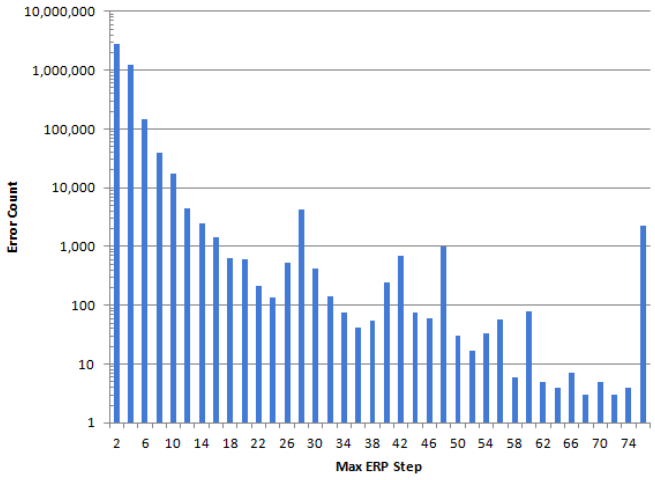

Figure 2: Histogram of error recovery steps that successfully corrected an error

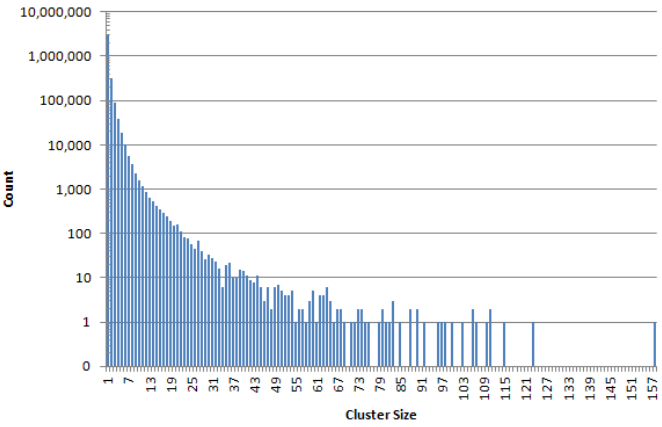

Figure 3: Histogram of sizes of clusters of errors associated with a single host command

in order to minimize the overhead, soft errors are reported if corrected at an error recovery step above a specified threshold. For example, Figure 2 shows the histogram of error recovery steps that successfully corrected an error for one of the drive populations studied in this paper. Note that the last bar includes all hard errors, hence the spike of the count. A typical threshold setting for that drive population (threshold=20) resulted in less than one percent of all internal errors reported to the host because errors are much more likely to be corrected earlier than the 20th error recovery step. Second, although multiple errors (hard or soft) may occur in response to a single host command, only one error, usually the one corrected at the highest error recovery step, is reported. Some media defects, such as scratches along a track, can result in multiple errors. For such multiple-error commands, all errors are recorded internally, but only one is reported to the host. These multiple-error

commands are fairly frequent. Figure 3 shows the histogram of the sizes of error clusters that are associated with a single host command. About one quarter of all soft errors were not reported to the host due to being 2nd, 3rd, etc. error resulting from a single host command. *Thus, internal data collected by a disk drive provides significant amount of information not observed by the host.* Furthermore, even though error recovery step information can be communicated over the host interface, to the best of our knowledge, no previous work has utilized this information. Nevertheless, host-based data collection does have one advantage over drive-based data – the disk drive typically has a large but limited buffer space to record errors, in contrast to the virtually unlimited storage available to the host.

The paper is organized as follows. Section II describes data source and analysis methodology. Section III contains the results of the study. Section IV discusses related work. Finally, a discussion of the impact of the results and potential avenues for future research are given in Section V.

II. Methodology

A. Challenges in Data Collection

As the quality of data analysis is limited by the quality of the underlying data, the data collection process is very important. In addition, because the overall reliability of disk drives is a function of not only the manufactured unit but also the workload over its lifetime and the attendant environmental conditions, data sampled from customer deployed drives is much more meaningful in terms of providing insight into the expected operation of that population of drives. In contrast to manufacturing test and qualification test data, which is easily obtained, collection of data from customer drives is fraught with many challenges.

First, the data analyst usually does not have direct access to customer drives. Thus, these drives must include a mechanism for recording the relevant data, and there must be a logistical procedure for dumping the recorded data and transporting that data to the data analyst. If the data analyst is the drive manufacturer, then several intervening parties (storage systems manufacturers, integrators, IT contractors, etc.) may need to cooperate to transfer that data. The storage system administrators must be willing to perform the data collection and initiate the transfer to the appropriate parties. The customers must be willing to accept the hopefully minimal impact on performance, and provision storage for the collected data. Depending on the level of detail contained in the collected data, the drive population size, and the periodicity of data collection, the amount of storage can be quite significant.

Second, customer privacy concerns must be addressed. This involves obtaining legal consent as well as assurances that no customer data is included in the collected data. The data must also be anonymized so that each party in the chain can only identify the adjacent parties, e.g., the IT service provider and the storage server provider know each other as customer and vendor, but the hard disk drive manufacturer does not know the identity of the IT service provider, nor of the end client whose workloads are going to execute on the hard disk drives and at whose site the data gathering is to happen.

Table I: Description of Collected Data

Population	#Logs	#Drives	Max Power-on Hours
Field	110,520	57,154	9,142
Qual1	51,897	1,200	2,025
Qual2	8,015	894	1,203
Qual3	15,278	983	1,045

B. Data Source and Characteristics

We cooperated with a large storage system manufacturer to help with data collection and transport. Table I describes the collected data that will be used in this study. The data includes four populations containing the same model of disk drives of the same brand. All drives are 15krpm, enterprise class drives, with FC-AL and SAS host interfaces and either 4 or 8 heads. The 'Field' population was collected by a single large storage system manufacturer and represents installed drives at multiple customer sites. This population consists of 57,154 drives. The total number of data logs collected was 110,520, as data was collected multiple times for some drives. The data collected covered drive operation from October 28, 2007 to November 13, 2008, a period of approximately one year. As seen in Figure 4, most of the drives had power-on hours (the total amount of time a drive has been powered on, taking into account the fact that a drive may be powered on and off depending on load, though this is rarely done in enterprise settings) of less than 4 months (which is roughly 2,900 hours). Fortunately, this amount of usage was already sufficient to allow some hard errors to occur.

In addition to the data for the large customer field population, data was also collected for three additional populations of drives that underwent qualification testing. The test duration and environment are similar, but the test workloads are different. The characteristics for each of these populations are given in Table I.

When data is collected from a particular drive, the error event log stored on the drive is saved. This error event log is a collection of errors that occur, including *all* soft and hard errors. However, since the available storage on each drive for the error event log is finite, the set of errors must necessarily be limited in two ways. First, the error events are saved to a ring buffer of size N entries that only saves the last N error events. If data is dumped from the same drive multiple times, the multiple sets of N error events will be merged to form one error event log of more than N entries. The exact time the data is dumped from a drive depends

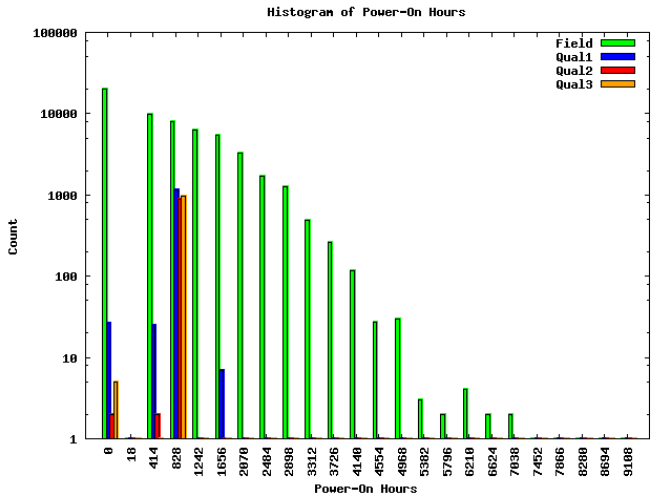

Figure 4: Histogram of power-on hours for all drives

on site-specific parameters and therefore no formal rule can be presented here for when in the operation log dumps are taken. Roughly half of the drives had their data dumped only once, while the arithmetic mean for how many times a data dump is taken for a disk drive is 1.93. This event of taking a data dump is not a periodic event and happens in the granularity of a few months. Second, as seen from Figure 1, the numbers of errors handled at the earliest error recovery steps are much greater than at later steps. Thus, an arbitrary cutoff for error recovery steps is implemented in the firmware, such that errors that are corrected at a recovery steps less than the cutoff are not stored in the error event log. In our case, soft errors that were corrected at error recovery step 1 or 2 were not stored on the drives' log.

C. Study Objectives

The issues to be addressed by this study involve the relationship between soft errors and hard errors. In particular, the key question is how well occurrences of soft errors can be used to predict subsequent hard errors. In this study, we consider the events where the drive signals an uncorrectable error in response to a host read, verify, or write command as hard errors. Because a disk drive contains many mechanical, electronic, and other parts, a varied set of failure modes is possible. In many of these failure modes, the failures are related to problems with a specific read/write head or media defects related to a specific read/write head. In addition, when a soft or hard error occurs, the drive's firmware does not provide information about the nature of the error (in particular, whether it affects multiple heads). Thus, this study focuses on analysis of soft errors on a per head basis. For each drive head, we only consider events up to and including the first hard error, because the study focuses on the predictive ability of soft errors in forecasting hard errors. It may at first thought appear that if a drive has a hard error, it will be taken out of commission and will therefore not have any chance of seeing any subsequent event. However, this is not necessarily true, as it depends on the policy of the administrator. For example, a hard write error generally results in a remapped sector but no loss of data. Usually a drive will not be decommissioned in that scenario. Also, for enterprise systems, usually there is some level of RAID, so a single hard error is usually not sufficient to decommission a drive.

III. Results

There are two main questions that will be studied based on the collected data: (1) Do soft errors precede hard errors? (2) If soft errors do precede hard errors, how much advance warning time do these soft errors give before the ensuing hard error occurs?

A. Do Soft Errors Precede Hard Errors?

Table II provides some answers to these questions based on data from the combination of all four populations described in Section II-B. 157 out of 387,840 heads experienced at least one hard error. A greater number of heads experienced at least one soft error. The column 'ERP Step Cutoff' specifies the minimum error recovery step for a soft error to be considered. The column "#Heads with SE" shows the number of heads with at least one soft error that is corrected at the particular error recovery step or higher. Since the error recovery steps are always traversed in exact sequential order, the numbers in this column are always nonincreasing.

Insight for the first question about whether soft errors precede hard errors is given by the fifth column "#Heads HE preceded by SE (% of all HE)", which shows the number of heads with hard errors that were preceded by soft errors at a given error recovery step. *The table shows that about one-third of the hard errors were preceded by soft errors* (38.9% for error recovery step 3 and 30.6% for error recovery step 12). This is an important number because it indicates a hard limit on the effectiveness of the use of soft errors to predict hard errors. A hard error that is not preceded by a soft error cannot be predicted by any algorithm that relies solely on soft error. In other words, most hard errors cannot be predicted by soft errors. Thus, the false negative rate for any hard error predictor based on soft errors is expected to be very high.

In addition to recall (true positive rate), a good predictor must also maximize precision. The column "%Heads HE | SE" from Table II shows the precision for a given error recovery step, i.e., the percentage of the heads with soft errors (from the fourth table column) that also experience an eventual hard error. In other words, it is equal to the number in the fifth column divided by the number in the fourth column. As expected, as the error recovery step cutoff

Table II: Soft/Hard Error Counts from All Populations

ERP Step Cutoff	Total Heads	#Heads with HE	#Heads with SE	#Heads HE preceded by SE (% of all HE)	%Heads HE | SE
3	387,840	157	18,932	61 (38.9%)	0.32%
4	387,840	157	11,142	58 (36.9%)	0.52%
5	387,840	157	5,711	57 (36.3%)	1.00%
6	387,840	157	2,496	53 (33.8%)	2.12%
7	387,840	157	1,634	52 (33.1%)	3.18%
8	387,840	157	1,426	50 (31.8%)	3.51%
9	387,840	157	1,276	49 (31.2%)	3.84%
10	387,840	157	1,133	49 (31.2%)	4.33%
11	387,840	157	1,050	49 (31.2%)	4.67%
12	387,840	157	910	48 (30.6%)	5.28%

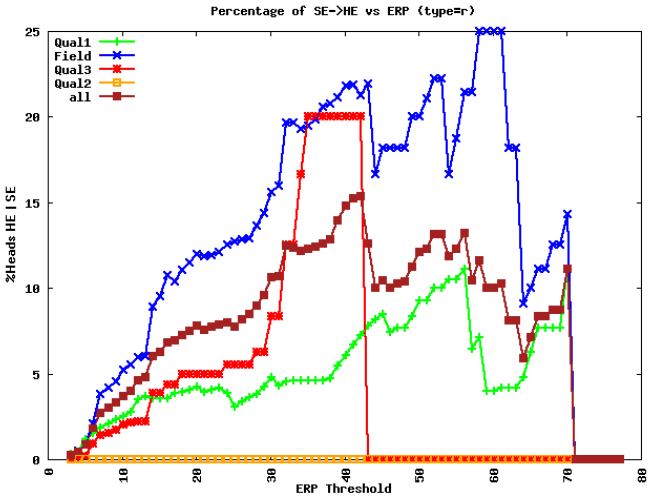

Figure 5: Percentage of drive heads which encounter a soft error at a specific error recovery step (ERP) before encountering a hard error. Only soft read errors are considered.

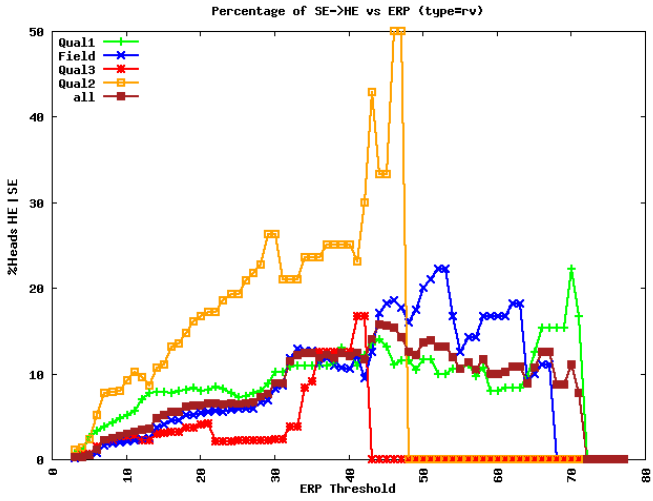

Figure 6: Percentage of drive heads which encounter a soft error at a specific error recovery step (ERP) before encountering a hard error. Both soft read and soft verify errors are considered.

increases, precision increases, since soft errors corrected at early steps are unlikely to be associated with a hard error.

Table II only lists a subset of error recovery steps due to space limitations. Precision for all error recovery steps are shown in Figure 5. Results from each population are also shown separately in addition to results from the combined population.

As expected, precision generally increases as the error recovery step increases. However, at some recovery steps, precision decreases due to the elimination of some hard errors that are only preceded by soft errors at lower recovery steps. This shows that soft errors are correlated to hard errors. However, regardless of the recovery step used, precision remains low, at no more than 25%, which means that at least 75% of raised alarms are false alarms.

The preceding results only considered soft read errors. Sometimes soft errors occur for verify operations, and these soft errors could potentially also be used as predictors for eventual hard errors. Verify operation is one where a read is attempted of a sector and it is checked if an error occurs. Often an error in the verify operation can be corrected by a code, such as, a Reed-Solomon code. The result of the read is discarded and only the error status of the sector being read is taken into further account. Figure 6 is similar to Figure 5 except that both soft read and soft verify errors are considered. In one qualification population (Qual2) precision improves significantly, but the other populations do worse at certain recovery steps.

B. *Amount of Advance Warning*

The amount of advance warning that a soft error provides for an ensuing hard error is heavily dependent on the specific algorithm used. For example, if a soft error rate coupled with a threshold is used as the prediction algorithm, then the amount of advance warning depends on the exact threshold, not to mention the rate calculation parameters, such as weighting of each error and the size of window

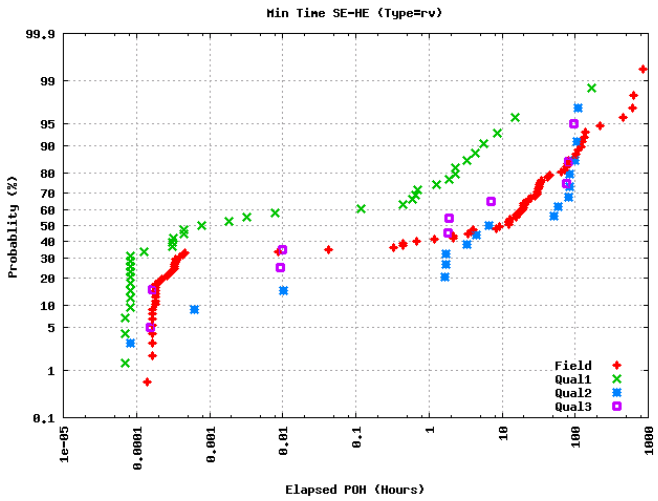

Figure 7: Cumulative distribution of durations between the last soft error and the hard error. Both soft read and soft verify errors are considered.

for determining which errors to include. However, we can gain some insight into this issue by looking at the duration between the last soft error and the hard error as it provides the lower bound of the amount of advance warning for algorithms that solely rely on soft errors. We call this value T_{min}. Note that hard errors that are not preceded by a soft error are not included in this analysis.

Figure 7 shows the cumulative distribution for T_{min} for all four populations from Table I. The x-axis is given in hours and is logarithmic in scale. Each data point represents one hard error. The distributions in Figure 7 are bimodal. About one-third of all hard errors occur within one second after a soft error. However, a large portion of hard errors have T_{min} of greater than one hour, sometimes reaching several days. In the "worst" of the four populations, 30% of hard errors have a T_{min} of one hour or greater, and in the "best", this number is 80%. This is encouraging because it allows preventive actions to finish before the hard errors occur.

IV. Related Work

There are several published studies on factors that influence disk drive failures and the failure patterns based on data collected from large-scale real-world storage systems. However, most of these studies focus on the *correlation* between failure rate and a particular parameter such as age or temperature. While this is useful for reducing failure rate of storage systems as well as deeper understanding of disk drive failures, the parameters with strong correlation to failures are not necessarily effective failure predictors. For example, suppose that in a drive population, 80% of drives operate at ≤ 40°C on average, and 1% of these drives eventually fail (within the window where the data were collected). On the other hand, within the remaining 20% which operate at > 40°C on average, 5% eventually fail. It would be said that temperature is strongly correlated with failures, as the failure rate is five times higher in the group that operates at higher temperature. However, if temperature was used as a failure predictor, it would achieve recall of 55.6% and precision of only 5%. This level of accuracy is generally not acceptable.

Pinheiro *et al.* [5] analyze more than 100,000 disk drives consisting of varied models and configurations. The analyzed factors are SMART parameters. SMART is a standard for monitoring disk drive parameters, which is thought to be useful for predicting some failures. Each drive manufacturer defines a set of attributes to expose under the SMART standard and selects threshold values which attributes should not go out of, under normal operation. Attribute values can range from 1 to 253 (1 representing the worst case and 253 representing the best). Depending on the manufacturer, a value of 100 or 200 will often be chosen as the "normal" value. While some factors are found to be strongly correlated with failure rate, an attempt to build a failure predictor for individual drives based on these factors is unsuccessful. More than half of the failed drives have zero counts in all of the factors strong correlated with failures. Even when other SMART parameters are used along with temperature data, 36% of the failed drives still have no indicator of failures at all.

Schroeder and Gibson [7] analyzes about 70,000 disks. The study focuses on the failure rate and patterns compared to drive specification and common assumptions. The results show that failure rate increases with age, including the early part of the drive's lifecycle. The analysis of failure distribution shows high temporal correlation between successive failures. Comparison with drives' datasheet MTTF shows that actual failure rates are much higher than specified.

Bairavasundaram *et al.* [6] uses data collected from 1.53 million disks to analyze the patterns of latent sector errors (hard errors) for both nearline and enterprise class disks. The study finds moderate degree of spatial correlation and high degree of temporal correlation between multiple latent sector errors. The correlation between latent sector errors (hard errors) and recovered errors (soft errors) are also analyzed. However, the analysis also includes soft errors that occur after the first hard error, as the goal was not to use soft errors to predict hard errors. Furthermore, soft errors used in the study come from the information reported to host after each command. This implies that they only include the soft errors that exceed a specific error recovery step threshold, which account for less than 1% of all soft errors internally logged. In addition, multiple errors that occur in response to a single host command are reported as a single error. Due to these reasons, the effectiveness of soft errors as a predictor of hard errors cannot be interpreted from the results.

Murray *et al.* [8] compare several machine learning methods in predicting failures in hard drives. The data used in the study came from 369 drives, 178 of which were failed drives.

However, it is important to note that the good drives came from a reliability test, run in a controlled environment by the manufacturer, while the failed drives were returned drives from actual users. The drive attributes used are SMART parameters sampled at two-hour interval. The problem is framed in the multiple-instance framework as each sample from a drive forms an instance and they correspond to a single drive which has a class label. Several machine learning algorithms were used to build a classifier and compared. The algorithm that achieves the best performance is SVM, with detection rate of 50.6% and no false alarms. However, since the source of good drives is different from the source of failed drives, the machine learning algorithms may be capturing the differences that are caused by different usage and environment rather than factors that are really indicative of an impending failure.

In summary, while several well-done studies have looked at the problem of reliability of disk drives, our current work is distinguished from them in two ways. First, it uses data that is available only to the firmware within the hard disk drive and not visible to the software stack outside of it. Second, it looks at the issue of predictability of hard errors and thus is concerned with events that occur preceding hard errors. This issue of predictability is of practical concern because a correct (accurate as well as precise) prediction means higher level software can take mitigation actions in practice and shield the end user from a visible failure.

V. Conclusion

Some error prediction algorithms use soft error rates to anticipate the future occurrence of hard errors. Such algorithms are predicated on the idea that the underlying physical causes of soft errors increase in intensity to the point of eventually causing a hard error. This study analyzes a set of hard disk drive populations, including a large customer field population, for insight into the relationship between soft errors and hard errors. The results are mixed in terms of providing support for the basis underlying such hard error prediction algorithms.

The first conclusion is that soft errors cannot be used to predict the majority of hard errors. Only about one-third of all hard errors are preceded by a soft error, and the remaining cannot be predicted by any soft error-based algorithm. We find soft errors to be correlated to hard errors. However, when used as the sole predictor of hard errors, the precision is at most 25%, meaning at least 75% of raised alarms are false alarms. With the consideration of additional data, such as data from sensors and various event data, in conjunction with soft errors, it may prove to be more successful. This should be one takeaway from this paper and should motivate investigation into enriching the feature set in trying failure prediction for hard errors.

The second conclusion is that for instances where prediction is successful, the prediction often yields sufficient time to initiate preventive actions. Often the amount of early warning time exceeds several hours. This amount of time is important because even successful prediction without time for emergency actions is useless. The results suggest a potential for a useful prediction algorithm that consider additional data, beyond soft errors.

It should be noted that although more extensive data, especially those from drives that have undergone longer periods of operation would certainly be desirable, studies involving actual customer deployed data are difficult to conduct due to the great challenges in obtaining such data. Thus, a large part of the value of this study derives from the large and varied size of the population and the use of these drives in real operational conditions and with real workloads.

Acknowledgment

As data is the most crucial element of data analysis, thanks are due the many companies and individuals that collected and transported the many disk drive data dumps used in this study. Thanks are also due to Michael Minckoff for his help in facilitating the availability of the collected data, to Don Gillis and Ed Kral for their collaboration in reliability analysis, and to Zvonimir Bandic, Wendy Chung, and Peter Baumgart for their support of the project.

References

[1] J. Elerath, "Hard disk drives: The good, the bad and the ugly!" *Queue*, vol. 5, pp. 28–37, September 2007.

[2] D. A. Patterson, G. Gibson, and R. H. Katz, "A case for redundant arrays of inexpensive disks (raid)," in *Proceedings of the 1988 ACM SIGMOD international conference on Management of data*, ser. SIGMOD '88. New York, NY, USA: ACM, 1988, pp. 109–116.

[3] P. M. Chen, E. K. Lee, G. A. Gibson, R. H. Katz, and D. A. Patterson, "Raid: high-performance, reliable secondary storage," *ACM Comput. Surv.*, vol. 26, pp. 145–185, June 1994.

[4] W. Jiang, C. Hu, Y. Zhou, and A. Kanevsky, "Are disks the dominant contributor for storage failures?: A comprehensive study of storage subsystem failure characteristics," *Trans. Storage*, vol. 4, pp. 7:1–7:25, November 2008.

[5] E. Pinheiro, W.-D. Weber, and L. A. Barroso, "Failure trends in a large disk drive population," in *Proceedings of the 5th USENIX conference on File and Storage Technologies*. Berkeley, CA, USA: USENIX Association, 2007, p. 2.

[6] L. N. Bairavasundaram, G. R. Goodson, S. Pasupathy, and J. Schindler, "An analysis of latent sector errors in disk drives," in *Proceedings of the 2007 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, ser. SIGMETRICS '07. New York, NY, USA: ACM, 2007, pp. 289–300.

[7] 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 *Proceedings of the 5th USENIX conference on File and Storage Technologies*. Berkeley, CA, USA: USENIX Association, 2007.

[8] J. F. Murray, G. F. Hughes, and K. Kreutz-Delgado, “Machine learning methods for predicting failures in hard drives: A multiple-instance application,” *J. Mach. Learn. Res.*, vol. 6, pp. 783–816, December 2005.