



A Study of SSD Reliability in Large Scale Enterprise Storage Deployments

Stathis Maneas and Kaveh Mahdavian, *University of Toronto*;
Tim Emami, *NetApp*; Bianca Schroeder, *University of Toronto*

<https://www.usenix.org/conference/fast20/presentation/maneas>

This paper is included in the Proceedings of the
18th USENIX Conference on File and
Storage Technologies (FAST '20)

February 25–27, 2020 • Santa Clara, CA, USA

978-1-939133-12-0

Open access to the Proceedings of the
18th USENIX Conference on File and
Storage Technologies (FAST '20)
is sponsored by



A Study of SSD Reliability in Large Scale Enterprise Storage Deployments

Stathis Maneas
University of Toronto

Kaveh Mahdavian
University of Toronto

Tim Emami
NetApp

Bianca Schroeder
University of Toronto

Abstract

This paper presents the first large-scale field study of NAND-based SSDs in *enterprise* storage systems (in contrast to drives in distributed data center storage systems). The study is based on a very comprehensive set of field data, covering 1.4 million SSDs of a major storage vendor (NetApp). The drives comprise three different manufacturers, 18 different models, 12 different capacities, and all major flash technologies (SLC, cMLC, eMLC, 3D-TLC). The data allows us to study a large number of factors that were not studied in previous works, including the effect of firmware versions, the reliability of TLC NAND, and correlations between drives within a RAID system. This paper presents our analysis, along with a number of practical implications derived from it.

1 Introduction

System reliability is arguably one of the most important aspects of a storage system, and as such a large body of work exists on the topic of storage device reliability. Much of the older work is focused on hard disk drives (HDDs) [2, 26–28], but as more data is being stored on solid state drives (SSDs), the focus has recently shifted to the reliability of SSDs. In addition to a large amount of work on SSDs in lab conditions under controlled experiments [3, 5–11, 13, 18–21, 31, 32, 36], more recently, the first field studies reporting on SSD reliability in deployed production systems have appeared [22, 23, 29, 34]. These studies are based on data collected at data centers at Facebook, Microsoft, Google, and Alibaba, where drives are deployed as part of large distributed storage systems. However, we observe that there still are a number of critical gaps in the existing literature that this work is striving to bridge:

- There are no studies that focus on *enterprise storage systems*. The drives, workloads, and reliability mechanisms in these systems can differ significantly from those in cloud data centers. For example, the drives used in enterprise storage systems include high-end drives and reliability is ensured through (single, double or triple parity) RAID, instead of replication or distributed storage codes.

- We also observe that existing studies do not cover some of the most important characteristics of failures that are required for building realistic failure models, in order to compute metrics such as the mean time to data loss. This includes, for example, a breakdown of the reasons for drive replacements, including the scope of the underlying problem and the corresponding repair action (RAID reconstruction versus draining the drive), and most importantly, an understanding of the correlations between drives in the same RAID group.

In this paper, we work toward closing these gaps and provide the first field study of a large population of SSDs deployed in NetApp’s *enterprise storage systems*. Our study is based on telemetry data for a sample of the total NetApp SSD population over a period of 30 months. Specifically, our study’s SSD population comprises of almost 1.4 million drives and includes drives from three different manufacturers, 18 different models, 12 different capacities, and four different flash technologies, i.e., SLC, cMLC (*consumer-class*), eMLC (*enterprise-class*), and 3D-TLC. The data collected for these drives is very rich, and includes information on drive replacements (including reasons for replacements), bad blocks, usage, drive age, firmware versions, drive role (e.g., data, parity or spare), among a number of other things. This paper presents the results from our analysis with a focus to close the gaps in existing work.

2 Background

2.1 Description of the Systems

The basis of our study is telemetry data from a large population of NetApp storage systems deployed in the field. The systems, also referred to as *filers*, employ the WAFL file system [17] and NetApp’s Data ONTAP operating system [24], which uses software RAID to provide resiliency against drive failures. The RAID subsystem can be configured to use SSDs in a Raid-TEC [16] (triple Parity), RAID-DP [12] (double Parity), or RAID-4 [25] (single Parity) configuration. The SSDs within a RAID group are homogeneous (same manufacturer, model, and capacity); the drives’ deployment time can vary

(from a few months to several years), but most SSDs within a RAID group were deployed at the same time. Systems run on custom Fabric-Attached Storage (FAS) hardware and use drives manufactured by other companies. They serve data over the network using file-based protocols such as NFS and CIFS/SMB, and/or block-based protocols, such as iSCSI.

Filers vary widely in their hardware configurations, in terms of CPU, memory, and number of drives. They are divided into two groups: one that uses SSDs as an intermediate write-back caching layer on top of HDDs, and another consisting of flash-only systems (called All Flash FAS (AFF)).

2.2 Description of the Data

The majority of all NetApp systems in the field send weekly NetApp Active IQ[®] bundles (previously called AutoSupport), which track a very large set of system and device parameters, but do not contain copies of the customers' actual data. This information is collected and automatically analyzed for corrective action and for detecting potential issues.

Our study is based on mining this collection of NetApp Active IQ messages. More precisely, our data set consists of 10 snapshots, each of which is based on parsing the entire body of NetApp Active IQ support messages at 10 different points in time: Jan/Jun 2017, Jan/May/Aug/Dec 2018, and Feb/Mar/April/May 2019. Each snapshot contains monitoring data for every filer (and its drives) and consists of all those NetApp Active IQ messages that were collected before the end of the corresponding month. Moreover, the data set provides information on filers and their configuration, including information on its different RAID groups and the role of a drive within a RAID group (data, parity, or hot spare drive).

Finally, a separate data set contains an entry for each drive that was marked as *failed* during the course of our study. These drives are being *replaced* (typically by a hot spare) and sent for offline testing and diagnosis. In the remainder of the paper, we use the terms replacement and failure interchangeably. The data set also contains a reason *type* for the majority of SSD replacements, explaining why the drive was replaced.

3 Summary Statistics

In this section, we present baseline statistics on the characteristics of the drives in our population and summary statistics on various reliability metrics.

3.1 Drive characteristics and usage

The first six columns in Table 1 describe the key characteristics of the different drive families in the SSD population, including manufacturer and model (in anonymized form), capacity (ranging from 200GB to 15.3TB), interface (SAS versus SATA), flash technology (SLC, cMLC, eMLC, 3D-TLC), lithography and the model's program-erase (PE) cycle limit, i.e., the maximum number of PE cycles it is rated for (ranging from 10K to 100K). Each drive family contains a few thousand to hundred thousand SSDs. Finally, as shown in

Table 1, the population spans a large number of configurations that have been common in practice over the last years.

The next four columns in the table present some summary statistics on how the different drive models have been used, including the over-provisioning (OP) factor (i.e., what fraction of the drive is reserved as spare capacity mostly to enable drive-internal garbage collection), the date when the first drive of this model was deployed, the median number of years drives have been powered on, and the mean and median fraction of the drives' rated life that has been used (i.e., the number of PE cycles the drive has experienced as a percentage of its PE cycle limit, as reported by the drive).

3.2 Health metrics

The last three columns in Table 1 provide statistics on three different drive health and reliability metrics. Specifically:

- *Percentage of Spare Blocks Consumed*: Each drive reserves an area (which is equal to 2.5% of the total drive capacity for the SSDs in this study) for remapping the contents of blocks that the drive internally declares as *bad*, e.g., due to an excessive error count. The *Percentage of Spare Blocks Consumed* metric reports what percentage of this area has been consumed (population mean and median).
- *Number of Bad Sectors*: Data ONTAP keeps track of a drive's defect list, known as *g-list*. This list is populated with a new entry every time the operating system receives an *unrecoverable error* for a block. The mean and median length of this list are reported as the *Number of Bad Sectors*.
- *Annual Replacement Rate (ARR)*: We make use of the common *Annual Replacement Rate* metric, defined as number of device failures divided by numbers of device years.

3.3 High-level observations

Below, we make a number of first observations based on Table 1, before we delve into a more detailed analysis of our data set in the remainder of this paper:

- The average ARR across the entire population is 0.22%, but rates vary widely depending on the drive model, from as little as 0.07% to nearly 1.2%. These numbers are significantly lower than numbers previously reported for data center drives. For example, even the worst model in our study (ARR of 1.2%) is at the low end of the range reported for SSDs in Google's data centers (range of 1-2.5% [29]). The rates are also significantly lower than common numbers reported for hard disk drives (i.e., 2-9% [26, 28]).
- Even for drive models that are very similar in their technical specifications (e.g., same manufacturer, flash technology, capacity, age), ARR can vary dramatically, e.g., 0.53% for II-G 15TB drives versus 1.13% for II-C 15.3TB drives.
- The spare area reserved for bad blocks is generously provisioned for the typical drive: even for drives that have been in the field for several years, the percentage of consumed spare blocks is on average less than 15%. Even the drives in the 99th and 99.9th percentile of consumed spare blocks have

Drive characteristics						Usage characteristics				Reliability metrics		
Manu- fact./ Model	Cap. (GB)	Inter- face	Flash Tech.	Lith.	PE Cycl.	OP	First Deploy- ment	Drive Power Years	Rated Life Used (%)	% of Blocks Cons.	Sp. Number of Bad Sectors	ARR (%) (all)
I - A	200	SAS	eMLC	2xnm	10K	28%	Apr '14	3.95	1.26 / 0	1.36 / 1	0.58 / 0	0.19
	400						Apr '14	3.93	0.52 / 0	1.54 / 1	2.96 / 0	0.16
	800						Mar '14	3.19	0.07 / 0	1.22 / 1	3.39 / 0	0.15
	1600						Mar '14	3.74	0.01 / 0	1.46 / 1	6.2 / 0	0.21
I - B	400	SAS	eMLC	1xnm	10K	28%	Dec '15	2.69	3.99 / 3	2.49 / 2	0.9 / 0	0.14
	800						Jan '16	2.58	1.87 / 2	3.7 / 4	4.62 / 0	0.17
	1600						Jan '16	2.55	1.68 / 2	3.55 / 3	6.01 / 0	0.23
I - C	400	SAS	eMLC	1xnm	10K	28%	Jan '17	1.7	7.48 / 7	1.83 / 2	1.86 / 0	0.12
	800						Jan '17	1.45	6.01 / 6	4.21 / 4	5.58 / 0	0.27
	1600						Mar '17	1.53	5.62 / 5	4.19 / 4	6.41 / 0	0.09
	3800						Jan '17	1.12	5.15 / 4	9.86 / 10	14.97 / 0	0.59
I - D	3800	SAS	eMLC	1xnm	10K	28%	Jul '17	1.01	4.01 / 3	10.01 / 10	12.34 / 0	0.23
II - A	3840	SAS	3D-TLC	V2	10K	7%	Dec '15	2.57	0.09 / 0	11.88 / 12	0.03 / 0	0.31
II - B	3800	SAS	3D-TLC	V2	10K	7%	Oct '16	1.77	0.01 / 0	12.01 / 12	0.36 / 0	0.13
II - C	8000 15300	SAS	3D-TLC	V3	10K	7%	Sep '17	1.06	0.01 / 0	12.38 / 12	0.03 / 0	0.69
							Sep '16	1.21	0.06 / 0	13.13 / 13	1.35 / 0	1.13
II - D	960 3800	SAS	3D-TLC	V3	10K	7%	Oct '16	1.54	0.06 / 0	16.05 / 15	0.03 / 0	0.12
							Oct '16	1.8	0.01 / 0	12	0.34 / 0	0.11
II - E	400	SAS	3D-TLC	V3	10K	28%	Dec '16	2.12	0.59 / 0	19.19 / 15	0.01 / 0	0.09
	800					28%	Jan '17	1.7	0.08 / 0	15.33 / 15	0 / 0	0.10
	3800					7%	Dec '16	2.02	0.05 / 0	12 / 12	0.26 / 0	0.13
II - F	400	SAS	3D-TLC	V2	10K	28%	Jan '16	2.52	0.86 / 0	12.31 / 12	0.01 / 0	0.48
	800						Feb '16	2.55	0.19 / 0	12.19 / 12	0.01 / 0	0.36
	1600						Jan '16	2.87	0.09 / 0	11.66 / 12	0.15 / 0	0.52
II - G	800	SAS	3D-TLC	V2	10K	28%	Apr '18	0.38	0.03 / 0	0 / 0	0 / 0	0.18
	960						Jan '18	0.5	0.11 / 0	0 / 0	0.03 / 0	0.18
	3800						Jan '18	2.89	0.09 / 0	11.64 / 12	0.15 / 0	0.28
	8000						May '18	0.45	0 / 0	0 / 0	0.25 / 0	0.37
	15000						May '18	0.46	0 / 0	0 / 0	0 / 0	0.53
II - H	800	SAS	cMLC	1xnm	10K	28%	Nov '14	3.61	1.34 / 0	7.49 / 7	1.36 / 0	0.10
II - I	200	SAS	eMLC	2xnm	30K	28%	Aug '09	4.83	5.24 / 2	6.7 / 6	0.31 / 0	0.07
	400						Dec '10	3.86	2.2 / 0	8.09 / 8	0.29 / 0	0.07
	800						Dec '10	4.66	0.69 / 0	6.94 / 7	4.53 / 0	0.11
II - J	400	SAS	eMLC	1xnm	10K	28%	May '15	3.37	2.32 / 0	6.9 / 7	0.08 / 0	0.18
	800						Jul '15	3.21	0.41 / 0	6.77 / 7	0.36 / 0	0.21
	1600						Jun '15	3.36	0.13 / 0	8.59 / 9	0.49 / 0	0.38
II - K	100	SATA	SLC	4xnm	100K	28%	Apr '12	0.43	1.4 / 1	3.78 / 4	3.62 / 0	0.14
II - L	100	SATA	SLC	3xnm	100K	28%	May '10	5.97	2.05 / 1	3.73 / 4	0.51 / 0	0.06
II - M	100	SATA	SLC	5xnm	100K	28%	Jul '10	4.23	1.63 / 1	14.82 / 13	0.18 / 0	0.11
III - A	200	SAS	eMLC	2xnm	30K	28%	Dec '12	5.89	3.08 / 1	0.02 / 0	6.18 / 0	0.07

Table 1: Summary statistics describing our population of drives. Whenever a column includes two values (separated by “/”), these correspond to the mean and median values of that population, respectively.

consumed only 17% and 33% of their spare blocks.

- The typical drive remains far from ever reaching its PE cycle limit. Even for models where most drives have been in the field for 2-3 years, less than two percent of the rated life is consumed on average. Even the drives in the 99th and 99.9th percentile of rated life used have consumed only 15% and 33% of their rated life, respectively. Hence, for the vast majority of drives, early death due to wear-out after prematurely reaching the PE cycle limit is unlikely.

4 Reasons for replacements

There are different reasons that can trigger the replacement of an SSD and also different sub-systems in the storage hierarchy which can detect issues that trigger the replacement of drives. For example, issues might be reported by the drive itself, the

storage layer, or the file system. Table 2 describes the different reason *types* that can trigger a drive replacement, along with their frequency, the recovery action taken by the system (i.e., copying out data from the drive to be replaced versus reconstructing the data using RAID parities), and the scope of the problem (i.e., risk of partial data loss, risk of complete drive loss, or no immediate problems). In our data set, the reason type is missing for 40% of all replacement events due to issues with the data collection pipeline. These issues are not related to the actual reason for the replacements. Hence, we can assume replacements with a missing reason type to be proportionately spread over the remaining categories. Therefore, the frequency of each replacement type is normalized to account for the missing data. We group the different reason types behind SSD replacements into four *categories*, labelled

Category	Type	Pct.	ARR (%)	Description	Recovery Action	Scope
A	SCSI Error	32.78%	0.055	The SCSI layer detects a hardware error reported by the SSD, that is severe enough that <i>immediate</i> replacement of the drive and reconstruction of the data is triggered. For example, these errors could be due to ECC errors originating from the drive's DRAM that prevent it from functioning properly.	RAID Reconstr.	Full
	Unresponsive Drive	0.60%	0.001	The drive has completely <i>failed</i> and become unresponsive.		
B	Lost Writes	13.54%	0.023	A lost write is detected when the contents of a 4K WAFL block (read from the SSD) are inconsistent based on its <i>signature</i> , which includes attributes and version number. Since there are many potential causes with the same symptom, a heuristic is used to decide whether to fail the disk or not. If multiple such errors occur within one SSD and no errors within any other SSD, then the former SSD is marked as <i>failed</i> .	RAID Reconstr.	Partial
C	Aborted Commands	13.56%	0.023	This error is generated due to an aborted command and is reported either by the SSD itself or the Storage Layer. For instance, this error can occur when the host sends some write commands to the device, but the actual data never reach the device due to an issue on the host or due to connection issues.	RAID Reconstr.	Partial
	Disk Ownership I/O Errors	3.27%	0.005	This error is related to the sub-system responsible that keeps track of which node owns a disk. In case an error occurs during the communication with this sub-system, then the SSD is immediately marked as <i>failed</i> .		
	Command Timeouts	1.81%	0.003	SSDs internally keep track of timers and also the Storage Layer maintains its own timers for every command sent to each SSD. This error indicates that the operation could not be completed within the allotted time even after retries.		
D	Predictive Failures	12.78%	0.021	The SSD reports this error based on a pattern of recovered errors that have occurred internally using its own thresholds and criteria, as specified by the corresponding manufacturer.	Disk Copy	Zero
	Threshold Exceeded	12.73%	0.020	The Storage Health Monitor sub-system keeps track of different parameters for each SSD and in case a threshold (e.g., on the number of media errors) is exceeded, the SSD is proactively replaced.		
	Recommended Failures	8.93%	0.015	This error is reported by the system and indicates that the drive should be replaced in the near future. This failure type is less strict and less urgent than Threshold Exceeded failures.		

Table 2: Description of reason types that can trigger a drive replacement. Disk copy operations are performed only where possible, i.e., a spare disk must be available; otherwise, the data of the replaced drive is constructed via RAID reconstruction.

from A to D, based on their severity.

The most benign¹ category is category D, which relates to replacements that were triggered by logic either inside the drive or at higher levels in the system, which predicts future drive failure, for example based on previous errors, timeouts, and a drive's SMART statistics [33].

The most severe category is category A, which comprises those situations where drives become completely unresponsive, or where the SCSI layer detects a drive problem severe enough to trigger immediate replacement of the drive and RAID reconstruction of the data stored in it.

Category B refers to drive replacements that are taking place when the system suspects the drive to have *lost a write*, e.g., because it did not perform the write at all, wrote it to a wrong location, or otherwise corrupted the write. The root cause could be a firmware bug in the drive, although other layers in the storage stack could be responsible as well. As there are many potential causes, a heuristic is used to decide whether to trigger a replacement or not; specifically, if multiple such errors occur within one SSD and no errors within any other SSD, then the former SSD will be replaced.

Finally, in category C most of its reasons for replacements are related to commands that were aborted or timed out.

¹We call them "benign" as the drive was still operational before getting replaced. Also, recovery is minimal (disk copy versus RAID reconstruction).

When examining the frequency at which individual replacement reason types are reported, we observe that the single most common reason type are SCSI errors, which are responsible for ~33% of all replacements and are unfortunately also one of the most severe reason types. The other severe reason for drive replacements, i.e., a drive becoming completely unresponsive, is reported for only 0.60% of all replacements.

Fortunately, one third of all drive replacements are merely preventative (category D) using predictions of future drive failures and are hence unlikely to have severe impact on system reliability. A detailed investigation of predictive replacements (not covered in the table due to space reasons) reveals that the most common trigger behind a preventative replacement is exceeding the threshold of consecutive timeouts.

The two remaining categories are roughly equally common and both have the potential of partial data loss if RAID reconstruction of the affected data should turn out unsuccessful. The first category (C) refers to aborted and timed out commands, and makes up ~19% of all reason types. The other category (B) refers to lost writes. This is an interesting category, since it is somewhat less clear whether it is the drive or other layers in the stack that are to blame for the lost write.

We will come back to the different reason types for replacements at various places in the remainder of the paper, when we will, for example, consider how the frequency of different

reason types behind replacements varies depending on drive capacity, lithography, age, or firmware version.

Finding 1: One third of replacements are associated with one of the most severe reason types (i.e., SCSI errors), but on the other hand, one third of drive replacements are merely preventative based on predictions.

5 Factors impacting replacement rates

In this section, we evaluate how different factors impact the *annual replacement rate* of the SSDs in our data set. We conduct our analysis on eMLC and 3D-TLC SSDs, and exclude cMLC and SLC drives due to insufficient data.

5.1 Usage and Age

Usage, and the wear-out of flash cells that comes with it, is well known to affect the reliability of flash-based SSDs; drives are guaranteed to remain functional for only a certain number of PE cycles. In our data set, SLC drives have a PE cycles limit of 100K, whereas the limit of most cMLC, eMLC, and 3D-TLC drives is equal to 10K cycles, with the exception of a few eMLC drive families with a 30K PE cycles limit.

Each drive reports the number of PE cycles it has experienced as a percentage of its PE cycle limit (the “*rated life used*” metric, recall Section 3.1), allowing us to study how usage affects replacement rates. Unfortunately, the rated life used is only reported as a truncated integer and a significant fraction of drives report a zero for this metric, indicating less than 1% of their rated life has been used. Therefore, our first step is a comparison of the ARR of drives that report less than 1% versus more than 1% of their rated life used. The results for eMLC and 3D-TLC drives are shown in Figure 1, which includes both overall replacement rates (“All”), and rates broken down by their replacement category (A to D). Throughout our paper, error bars refer to 95th percentile confidence intervals and we exclude two outlier models, i.e., II-C and I-C, with unusually high replacement rates to not obscure trends (except for graphs involving individual drive families).

We also perform statistical tests and calculate *p-values* to confirm our hypotheses (where applicable). For each test case, we perform a two-sample *z-test* [1]. Since our analysis is based on replacement rates, we need to calculate and compare the replacement rates of the two groups in each test. For each group, we create 1,000 random samples of replacement rates; in each sample, the replacement rate is measured based on a randomly chosen set of 1,000 SSDs from the corresponding group. Finally, we perform a *z-test* on the two sets of samples and report the calculated *p-value* associated with the test.

Figure 1 provides evidence for effects of infant mortality. For example, for eMLC drives, the drives with less than 1% rated life used are more likely (1.25X) to be replaced than those with more than 1% of rated life used (the estimated mean replacement rates of the two populations are 0.168 and 0.126 respectively, whereas the corresponding *p-value* is equal to $6.3211e-45$). When further breaking results down by

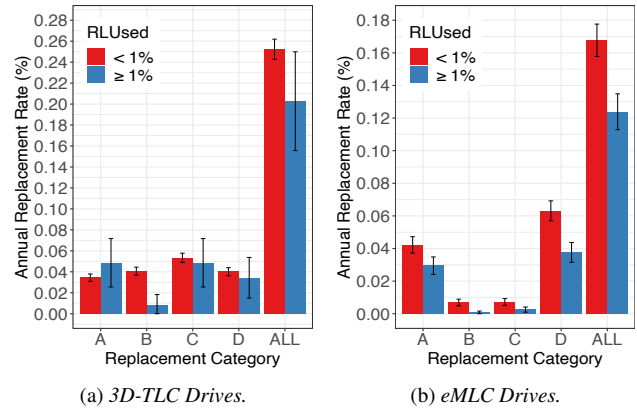


Figure 1: Annual replacement rate per flash type based on the drives’ “rated-life-used” percentage.

reason category, we find that drives with less usage consistently experience higher replacement rates for all categories. Making conclusive claims for the 3D-TLC drives is harder due to limited data on drives above 1% of rated life used, resulting in wide confidence intervals. However, where we have enough data, observations are similar to those for eMLC drives, e.g., we see a significant drop in lost writes for drives above 1% of rated life used.

We also looked separately at drives that are extensively used (more than 50% of their PE cycles) and their typical reasons for replacement. We see the trend of decreasing rates of lost writes continues here, as we don’t observe a single case related to *lost writes* among these drives. One possible explanation is that lost writes might be related to firmware bugs, and as firmware gets updated to improved versions over the course of a drive’s life, rate of lost writes drops. We take a closer look at firmware versions in Section 5.5. It’s also possible that issues leading to lost writes typically become evident early in a drive’s life and the drive gets replaced before it makes it to more than 1% or 50% of its rated life.

Another interesting observation is that the heavily used drives are more likely to be replaced due to *predictive failures* compared to the overall population. This could mean that issues leading to predictive failures are only exposed after some significant usage (e.g., hardware problems that cause media errors and bad blocks, which then trigger failure prediction, require thoroughly exercising the NAND). Alternatively, it could mean that after more drive usage more data is available on the drive’s health status, which improves predictions.

We also look at replacement rates as function of a drive’s age measured by its total months in the field. Figure 2 (top) shows the conditional probability of a drive being replaced in a given month of its life. i.e., the probability that the drive will fail in month x given that it has survived month $x-1$.

We observe an unexpectedly long period of infant mortality with a shape that differs from the common “bathtub” model often used in reliability theory. The bathtub model assumes a short initial period of high failure rates, which then quickly

drops [14, 15, 28, 35]. Instead, we observe for both 3D-TLC and eMLC drives, a long period (12–15 months) of increasing failure rates, followed by a lengthy period (another 6–12 months) of slowly decreasing failure rates, before rates finally stabilize. That means that, given typical drive lifetimes of 5 years, drives spend 20–40% of their life in infant mortality.

We wondered whether these unexpected results might just be an artifact of our heterogeneous population, since each line in Figure 2 (top) is computed over a population comprising different drive families with different drive ages and characteristics. We therefore plotted in Figure 2 (bottom) the same probabilities, but this time only over a subset of drive families with similar characteristics (e.g., age and lithography). Again, we observe the same trends, and in fact in some aspects even slightly more pronounced: the duration of the two phases is similar in length and for 3D-TLC drives, the ratio of the peak failure rate to the lowest rate is even larger (a factor of 2.5X).

It might be surprising at first that we do not observe an increase in ARR for drives towards the end of their life. The reason is that the majority of drives, even those deployed for several years, do not experience a large number of PE cycles. Their fraction even in the population of older drives is too small to drive up the overall ARR.

Finding 2: We observe a very drawn-out period of infant mortality, which can last more than a year and see failure rates 2–3X larger than later in life.

5.2 Flash and drive type

The drive models in our study differ in the type of flash they are based on, i.e., in how many bits are encoded in a single flash cell. For instance, Single Level Cell (SLC) drives encode only one bit per cell, while Multi-Level Cell (MLC) drives encode two bits in one cell for higher data density and thus a lower total cost, but potentially higher propensity to errors. The most recent generation of flash is based on Triple Level Cell (3D-TLC) flash with three bits per cell.

The last column in Table 1 allows a comparison of ARRs across flash types. A cursory study of the numbers indicates generally higher replacement rates for 3D-TLC devices compared to the other flash types. Also, we observe that 3D-TLC drives have consumed 10–15X more of their spare blocks.

For a more nuanced comparison between 3D-TLC and eMLC we turn to Figures 1 and 4, which also take usage and lithography into account. Figure 1 indicates that ARRs for 3D-TLC drives are around 1.5X higher than for eMLC drives, when comparing similar levels of usage. Figure 4 paints a more complex picture. While V2 3D-TLC drives have a significantly higher replacement rate than any of the other groups, the V3 3D-TLC drives are actually comparable to 2xnm eMLC drives, and in fact have lower ARR than the 1xnm eMLC drives. So, lithography might play a larger role than flash type alone (we take a closer look at in Section 5.4).

We are interested in differences between the enterprise-class eMLC drives and consumer-class cMLC drives. Unfor-

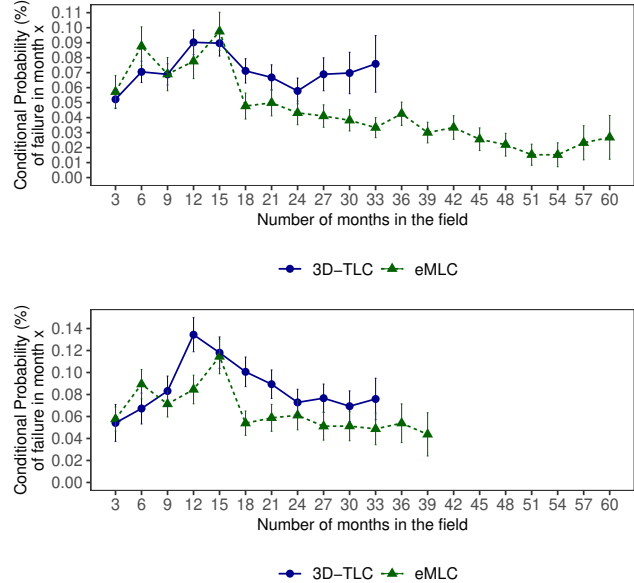


Figure 2: Conditional probability of failure based on a drive’s age (number of months in the field) for all drive families (top) and a subset of them (bottom), i.e., II-A, II-B, and II-F for 3D-TLC drives and I-B, I-C, I-D, and II-J for eMLC drives.

tunately our data set contains only one family of cMLC drives (II-H). Interestingly, we find that this one family of cMLC drives reports much lower replacement rates than eMLC families of similar age and capacity (i.e., II-H drives vs II-J drives). Narayanan et al. [23] report replacement rates between 0.5–1% for their *consumer* class MLC drives, with the exception of a single *enterprise* class model, whose replacement rate is equal to 0.1%; however, the authors in [23] consider only fail-stop failures. In our study, we consider different types of failures and thus, the reported replacement rates would have been even smaller had we considered only fail-stop failures.

Finally, we observe that SLC models are not generally more reliable than eMLC models that are comparable in age and capacity. For example, when we look at the ARR column of Table 1, we observe that SLC models have similar replacement rates to two eMLC models with comparable capacities, i.e., II-I and III-A drives (their difference is small but still statistically significant, i.e., the estimated mean replacement rates of the two populations are 0.112 and 0.091 respectively, with a p-value equal to 5.0841e-22). This is consistent with the results in a field study based on drives in Google’s data centers [29], which does not find SLC drives to have consistently lower replacement rates than MLC drives either. Considering that the lithography between SLC and MLC drives can be identical, their main difference is the way cells are programmed internally, suggesting that controller reliability can be a dominant factor.

Finding 3: Overall, the highest replacement rates in our study are associated with 3D-TLC SSDs. However, no single flash type has noticeably higher replacement rates than the

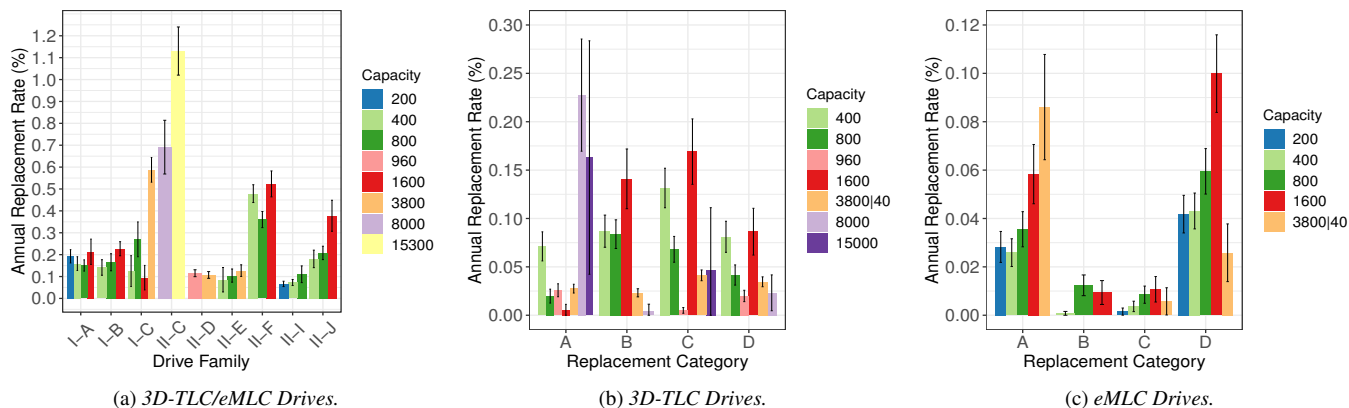


Figure 3: Figure 3a shows the annual replacement rates for the drive families shipped with multiple capacities. Figures 3b and 3c show replacement rates for different capacities broken down by their replacement category, for 3D-TLC drives and eMLC drives, respectively. In these figures, the 3800GB and 3840GB capacities have been consolidated.

other flash types studied in this work, indicating that other factors, such as capacity or lithography, can have a bigger impact on reliability.

5.3 Capacity

The drives in our data set range in capacity from 100GB to 15.3TB and most drive families include drives of different capacities, which allows us to study the effect of drive capacity on replacement rates. One would expect the rate of failures that are due to underlying hardware issues (such as failure of NAND cells or DRAM) would grow with capacity. On the other hand, failures that are related to firmware bugs are not likely to be strongly correlated with capacity (all else being equal). In the remainder of this section, we test our hypothesis of NAND related problems increasing with capacity.

First, we turn to the reported numbers on bad sectors in Table 1. We observe that consistently within each drive family, the total number of bad sectors continuously increases with capacity. For example, for model I-C, the average number of bad sectors per drive is growing from 1.9 to 5.6 to 6.4 to 14.97 for capacities of 400, 800, 1600 and 3800 GB, respectively. Moreover, the percentage of drives with a non-zero count of bad blocks continuously increases with capacity.

Figure 3a explores how overall ARR change with capacity by plotting the ARR of different drive families, broken down by capacity. We make a slightly more nuanced observation for ARR, compared to the bad sector count. For smaller capacities, in the range of 200 to 1600 GB, the ARR shows no clear relationship with capacity. It might be that for these smaller capacities replacements are dominated by reasons other than issues with the underlying NAND. The trend starts to change around 1600GB, as for four out of the five families that have 1600GB drives, those drives have the highest ARR. And for larger capacities, there is a clear trend for increasing ARR. The 15TB drives always have higher ARR than the other drives in the same family. The 3800GB and 8000GB

drives always have higher ARR than the drives less than 3800GB within the same family.

We also looked for differences in the reasons for replacement between smaller and larger capacity drives and made an interesting observation: for the largest capacity drives, the rate of predictive failures is lower than for smaller capacity drives. In contrast, the most severe failure reason, i.e., an unresponsive drive, occurs at a much higher rate for the larger capacity drives than for the smaller capacity drives. Figures 3b and 3c illustrate this observation, as they break down the ARR by capacity and replacement category for different flash technologies. Among the eMLC drives, the 3800GB and 3840GB capacities and among the 3D-TLC drives, the 8TB and 15TB capacities have very high rates of replacement due to an unresponsive drive, compared to smaller capacities. They also have a lower rate of replacements due to predictive failures. This means that the replacement rate associated with high capacity drives is not only bigger, but also has potentially more severe consequences. Another potential implication is that failures of large capacity drives are either harder to predict or the prediction algorithms have not been optimized for them. It may be possible that the severe failures and unpredictability of such failures is an artifact of the larger DRAM footprint associated with large flash capacity, rather than the flash capacity itself. Potential for such impact could be mitigated by upcoming architectures such as Zoned Storage (ZNS) [4, 30] that obviate the need for large Flash Translation Layer (FTL) tables in DRAM and consequently reducing the DRAM footprint.

Finding 4: Drives with very large capacities not only see a higher replacement rate overall, but also see more severe failures and fewer of the (more benign) predictive failures.

5.4 Lithography

Lithography has been shown to be highly correlated with a drive’s raw bit error rate (RBER); models with smaller lithog-

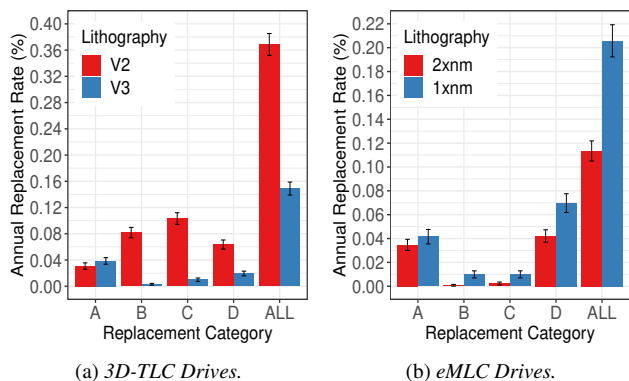


Figure 4: Annual replacement rate per flash type and lithography broken down by replacement category.

raphy report higher RBERS according to a study based on data center drives [29], but not necessarily higher replacement rates. We explore what these trends look like for the drives in enterprise storage systems. To separate the effect of lithography from flash type (i.e., SLC, cMLC, eMLC, 3D-TLC), we perform the analysis separately for each flash type.

The bar graph in Figure 4 (right) shows the ARR for eMLC drives separated into 2xnm and 1xnm lithographies broken down by failure category, also including one bar for replacements of all categories. We observe that the higher density 1xnm drives experience almost twice the replacement rate of 2xnm drives (the p-value is equal to 4.2365e-120). Also, replacement rates for each of the individual reason categories are higher for 1xnm drives than for 2xnm, with the only exception of reason category A, which corresponds to unresponsive drives. Finally, we also observe that the 1xnm drives also have, on average, consumed a larger percentage of spare blocks (an indicator of developing bad blocks) and developed a larger number of bad sectors, despite the fact that they are generally younger than the 2xnm drives.

In contrast to eMLC drives, the 3D-TLC drives see higher replacement rates for the lower density V2 drives, which internally have fewer layers than V3 (the corresponding z-test returns a p-value equal to 2.7624e-275). When breaking replacement rates down by reason category, we observe that consistently with the results for eMLC drives, the only category that is not affected by lithography is category A, which corresponds to unresponsive drives. Regarding the percentage of spare blocks consumed, we observe comparable values between V2 and V3 drives (if we exclude the II-G family, which is much younger than the others).

Finally, for SLC drives, we do not see a clear trend for replacement rates as a function of lithography; however, we also have limited data, with only one drive model in each lithography for SLC drives.

Finding 5: *In contrast to previous work, higher density drives do not always see higher replacement rates. In fact, we observe that, although higher density eMLC drives have higher replacement rates, this trend is reversed for 3D-TLC.*

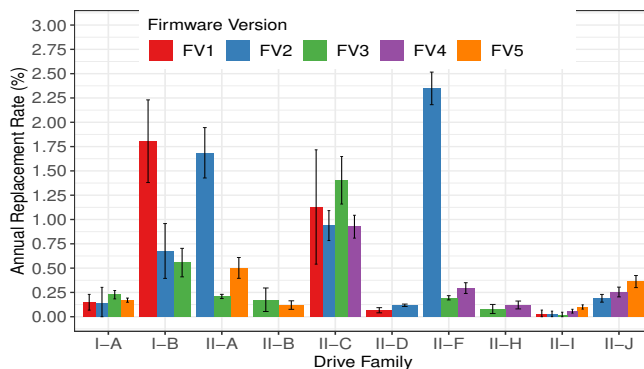


Figure 5: Annual replacement rates per firmware version.

5.5 Firmware Version

Given that bugs in a drive’s firmware can lead to drive errors or in the worst case to an unresponsive drive, we are interested to see whether different firmware versions are associated with a different ARR. Each drive model/family in our study experiences different firmware versions over time. We name the first firmware version of a model FV1, the next one FV2, and so on. An individual drive’s firmware might be updated to a new version, but we observe that the majority of drives (70%) appear under the same firmware version in all data snapshots.

Figure 5 shows the ARR associated with different firmware versions for each drive model. Considering that firmware varies across drive families and manufacturers, it only makes sense to compare the ARR of different firmware versions within the same drive family. To avoid other confounding factors, in particular age and usage, the graph in the figure only includes drives with rated life used of less than 1% (the majority of drives). We have also analyzed the data in different ways, for example by including only drives that appear consistently under the same firmware version in all data snapshots, observing similar results.

We find that drive’s firmware version can have a tremendous impact on reliability. In particular, the earliest versions can have an order of magnitude higher ARR than later versions. This effect is most notable for families I-B (more than factor 2X decrease in ARR from FV1 to FV2), II-A (factor 8X decrease from FV2 to FV3) and II-F (more than 10X decrease from FV2 to FV3). The corresponding z-tests return extremely small p-values and thus, confirm our results.

We note that the effect where earlier firmware versions have higher replacement rates persists even if we only include drives whose firmware has never changed in our data snapshots, e.g., we compare drives that spend their entire lives in FV1 and compare them to drives who only saw FV2. This provides confirmation that the effect is actually due to firmware versions, and not due to infant mortality, where the earlier version is used at an earlier time of a drive’s life and the later version during a later point in life.

A likely explanation is that later firmware versions include bug fixes and improvements over earlier versions. This expla-

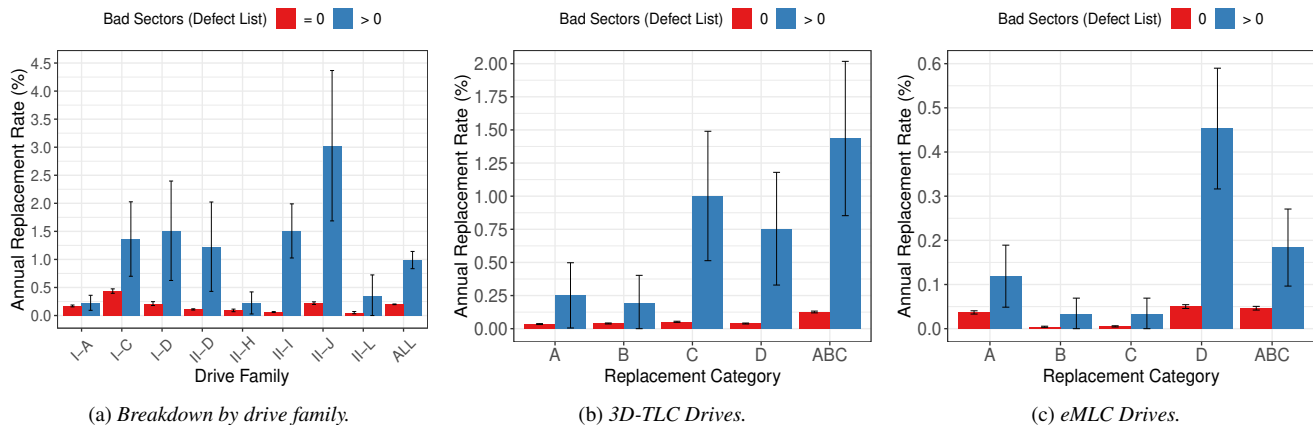


Figure 6: Annual replacement rates per flash type based on the drives' bad sectors count. Figure 6a breaks results down by drive family and Figures 6b and 6c by replacement category.

nation is further supported by our observation that the failures that decrease the most when moving from FV1 to later versions (e.g., for the two families with the highest decrease in ARR, II-A and II-F) are failures in categories B and C (lost writes and timeouts), both of which could be caused by firmware problems.

Interestingly, we also observe cases where the ARR increases with increasing version numbers, albeit not as frequently. One example is family II-J, where FV5 has a significantly higher ARR than FV2 or FV4. The difference between FV2 and FV5 is more than factor of 10X when considering only drives that do not change firmware version (graph omitted for lack of space). One possible explanation is that as the firmware code base evolves, it becomes more complex and the new code also introduces new bugs.

Finding 6: Earlier firmware versions can be correlated with significantly higher replacement rates, emphasizing the importance of firmware updates.

5.6 All Flash FAS (AFF) Systems

We also looked at whether a drive's type of usage, i.e., either as part of an AFF system or as part of a caching layer, affects its replacement rate. We find no indication that within a drive family, replacement rates vary as a function of type of usage.

5.7 Device Role

We also studied whether drives within a RAID group have different replacement rates, depending on their role in the RAID group (i.e., data and parity), but found no indication of statistically significant differences. This might indicate that WAFL is effective at balancing load across drives and minimizing the number of parity updates.

5.8 Over-provisioning

We also looked at the amount of over-provisioning (OP) as a factor, but find no clear correlation between the amount of over-provisioned space and ARR. One reason might be that the typical drive in our population is far from reaching its en-

durance limit. Therefore, the potential endurance-increasing effects of over-provisioning do not become relevant.

5.9 Number of bad blocks

In this section, we are exploring the relationship between a drive developing bad blocks and replacement rates. We consider two different metrics associated with bad blocks.

The first metric is the length of the *g-list*, also referred to as defect list, which is maintained by Data ONTAP and contains an entry for every block generated an *unrecoverable* error upon access. Since the *g-list* is empty for a large fraction of drives (99.04%), we distinguish between drives with an empty and a non-empty *g-list*, and plot their ARR separately. Figure 6a shows the results broken down by drive family.

We observe that drives that have experienced at least one unrecoverable error (i.e., they have a non-empty *g-list*) have significantly higher replacement rates. Part of this observation might just be an artifact of predictive drive replacements (category D), as predictions might be based on the length of the *g-list*. We therefore plot in Figures 6b and 6c the same rates, but broken down by replacement category.

Not surprisingly, we see that there is a strong correlation between a non-empty *g-list* and predictive failures (category D); however, the more interesting observation is that also for the other replacement reasons, there is a correlation between having a non-empty *g-list* and the drive being replaced. That means developing unrecoverable errors is indicative of a variety of future issues a drive might develop.

The second factor we consider is the number of consumed spare blocks inside each individual SSD. While we omit full results due to lack of space, we note that again we observe similar correlations.

Finding 7: SSDs with a non-empty defect list have a higher chance of getting replaced, not only due to predictive failures, but also due to other replacement reasons as well.

Finding 8: SSDs that make greater use of their over-provisioned space are quite likely to be replaced in the future.

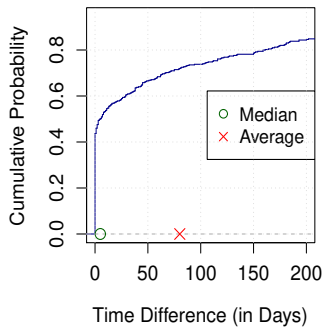
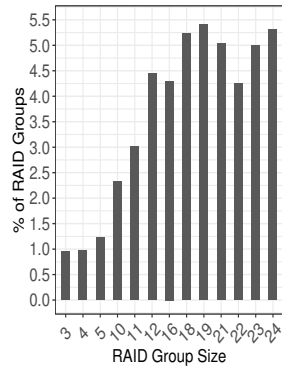
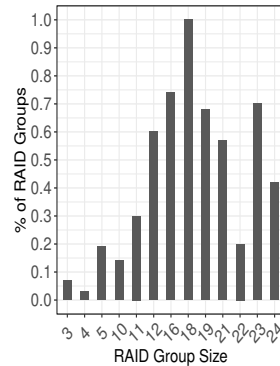


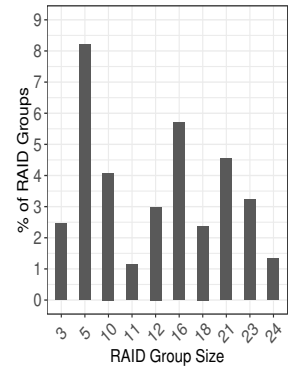
Figure 7: Time difference between successive replacements within RAID groups.



(a) RAID groups that experience at least 1 replacement.



(b) RAID groups that experience multiple replacements.



(c) RAID groups with replacement that experience at least 1 follow-up replacement (within 1 week).

Figure 8: Statistics on replacements within RAID groups.

6 Correlations between drive failures

A key question when deriving reliability estimates, e.g., for different RAID configurations, is how failures of drives within the same RAID group are correlated.

As a first measure of correlation, we explore the probability that a RAID group will experience a drive replacement following a prior drive replacement. More precisely, we start by computing the empirical probability that a RAID group will experience a drive replacement in a random week; this probability is equal to 0.0504%. Then, we compute the probability that a RAID group will experience another drive replacement within a week following a previous drive replacement. The probability is equal to 9.39%, that is, more than a factor of 180X increase compared to the probability that a drive replacement will occur within a random week. We speculate on a few possible reasons that could explain this. First, RAID reconstruction imposes an additional load to the other drives of the group and exposes latent errors, as these drives must be fully scanned to reconstruct the data of the failed drive. Second, shared environmental issues (e.g., overheating, power surge), could affect multiple drives from the same group simultaneously, as they are all placed within the same filer.

For a more detailed understanding of correlations, we consider all RAID groups that have experienced more than one drive replacement over the course of our observation period and plot in Figure 7, the time between consecutive drive replacements within the same RAID group. We observe that very commonly, the second drive replacement follows the preceding one within a short time interval. For example, 46% of consecutive replacements take place at most one day after the previous replacement, while 52% of all consecutive replacements take place within a week of the previous replacement.

Another important question in RAID reliability modelling is how the chance of multiple failures grows as the number of drives in the RAID group increases. Figure 8a presents, for the most common RAID group sizes, the percentage of

RAID groups of that size that experienced at least one drive replacement. As one would expect, larger RAID groups have a higher chance of experiencing a drive replacement; yet, the effect of a RAID group’s size on the replacement rates saturates for RAID groups comprising more than 18 drives.

However, we make an interesting observation in Figure 8b, when we look at the percentage of RAID groups that have experienced at least two drive replacements (potential double failure): this percentage is not clearly correlated with RAID group size, except for maybe very small RAID groups of three or four drives. The largest RAID group sizes do not have a higher rate of double (or multiple) failures.

The reason becomes clear when we look at the conditional probability that a RAID group will experience a replacement, *given that it has already experienced another replacement*, in Figure 8c. More precisely, for each RAID group size, we consider the RAID groups that had at least one drive replacement and compute what percentage of them had at least one more replacement within a week. Interestingly, we observe there is no clear trend that larger RAID group sizes have a larger chance of one drive replacement being followed by more replacements. Note that, as already mentioned, the chance of experiencing a drive failure grows with the size of the RAID group (Figure 8b); however, the chance of *correlated failures* does not show a direct relationship with the group’s size.

Finding 9: While large RAID groups have a larger number of drive replacements, we find no evidence that the rate of multiple failures per group (which is what can create potential for data loss) is correlated with RAID group size. The reason seems to be that the likelihood of a follow-up failure after a first failure is not correlated with RAID group size.

7 Related Work

Four recent field studies have looked at failure characteristics of SSDs in data centers at Facebook, Microsoft, Google, and Alibaba, respectively [22, 23, 29, 34]. Our work is different

in that we focus on enterprise storage systems, rather than distributed data center storage and is the first to report on TLC drives, large capacity drives (8 TB and 15 TB), and several models with 10xnm lithographies. Moreover, our study considers a large number of factors that were not studied in previous work, such as the effect of firmware versions and failure correlations within a RAID group.

Where we report statistics that were also considered in previous work, we have included a comparison in the relevant sections of our paper. In other cases, a direct comparison with failure rates reported in prior work is not meaningful. For example, the Facebook [22] and Microsoft [23] studies focus on uncorrectable errors and fail-stop events respectively, which are different from the drive replacements considered in our study. Furthermore, fail-stop events do not always lead to drive replacements, and other events that might lead to replacements are not included in the rates reported in [23]. Similarly, while the study of drives at Alibaba [34] includes a breakdown of reason for replacement, their taxonomy is different, with categories that do not map to ours. Moreover, their work does not report on rates of failures (only the relative frequency of reasons).

8 Lessons learned

- Our observations emphasize the importance of firmware updates, as earlier firmware versions can be correlated with significantly higher failure rates (§5.5). Yet, we observe that 70% of drives in our study remain at the same firmware version throughout the length of our study. Consequently, we encourage enterprise storage vendors to make firmware upgrades as easy and painless as possible, so that customers apply the upgrades without worries about stability issues.
- A question that often comes up when configuring RAID groups is how the size of a group, in terms of number of drives, will affect its reliability. After all, intuitively, more drives create more potential for failures. Our observations show that larger RAID groups might not be as bad as often thought. While large RAID groups have a higher number of drive replacements, we have no evidence that the rate of multiple failures per group (which is what creates potential for data loss) is correlated with RAID group size (§6).
- Our results highlight the occurrence of temporally correlated failures within the same RAID group (§6). This observation indicates that single parity RAID configurations (e.g., RAID-5), might be susceptible to data loss, and realistic data loss analysis certainly has to consider correlated failures.
- Drives with very large capacities experience higher failure rates overall and see more severe failures (§5.3). The higher failure rate could stem from the larger amount of NAND and dies on the drives, emphasizing the importance of a drive and its system being able to handle a partial drive failure, such as a die failure. NetApp is working toward this direction by carving out the lost capacity of a dead die from the OP area.
- Our observation regarding the smaller rate of predictive

failures for larger capacities (§5.3) also brings up the question whether large capacity drives require different types of failure predictors and potentially more input from the drive on its internal issues (e.g., a bad die or issues with DRAM).

- There is renewed concern around NAND-SSDs reliability with the introduction of QLC NAND, whose PE cycle limit is significantly lower than current TLC NAND. Based on our data, we predict that for the vast majority of enterprise users, a move towards QLC's PE cycle limits poses no risks, as 99% of systems use at most 15% of the rated life of their drives.
- There has been a fear that the limited PE cycles of NAND SSDs can create a threat to data reliability in the later part of a RAID system's life due to correlated wear-out failures, as the drives in a RAID group age at the same rate. Instead, we observe that correlated failures due to infant mortality are likely to be a bigger threat. For example, for the 3D-TLC drives in our study, the failure rate at the peak of infant mortality is 2.5X larger than later in life (§5.1).
- We observe unexpected behavior for failure rates as a function of age (§5.1). In contrast to the “bathtub” shape assumed by classical reliability models, we observe no signs of failure rate increases at end of life and also a very drawn-out period of infant mortality, which can last more than a year and see failure rates 2-3X larger than later in life. This brings up the question what could be done to reduce these effects. One might consider, for example, an extended, more intense burn-in period before deployment, where drives are subjected to longer periods of high read and write loads. Given the low consumption of PE cycles that drives see in the field (99% of drives do not even use up 1% of their PE cycle limit), there seems to be room to sacrifice some PE cycles in the burn-in process. More detailed recommendations would require a more thorough understanding of the relationship between PE cycles and failure rates; we are currently working on collecting such data.
- When choosing among drive types/models, our results indicate that from a reliability point of view, flash type (i.e., eMLC versus 3D-TLC) seems to play a smaller role than lithography (i.e., 1xnm versus 2xnm eMLC) or capacity (§5.2–5.4).

9 Acknowledgements

We would like to acknowledge several people at NetApp for their contributions to this work; Rodney Dekoning, Saumyabrata Bandyopadhyay, and Anita Jindal for their early support and encouragement, Aziz Htite, who helped cross-validate our data and assumptions along the way. The internal reviewers within the ATG, ONTAP WAFL, and RAID groups, whose careful feedback made this a better paper. A very special thank you to Biren Fondekar's Active IQ team in Bangalore; Asha Gangolli, Kavitha Degavinti, and finally Vinay N. who spent countless late nights on the phone with us, as we cleaned and curated the foundational data sets of this paper. We also thank our reviewers and our shepherd, Devesh Tiwari, for their detailed feedback and valuable suggestions.

References

- [1] Alan T. Arnholt and Ben Evans. *BSDA: Basic Statistics and Data Analysis*. 2017. R package version 1.2.0.
- [2] Lakshmi N. Bairavasundaram, Garth R. Goodson, Shankar Pasupathy, and Jiri 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 (SIGMETRICS '07)*, pages 289–300, 2007.
- [3] Hanmant P Belgal, Nick Righos, Ivan Kalastirsky, Jeff J Peterson, Robert Shiner, and Neal Mielke. A new reliability model for post-cycling charge retention of flash memories. In *Proceedings of the 40th Annual International Reliability Physics Symposium*, pages 7–20. IEEE, 2002.
- [4] Matias Bjørling. From Open-Channel SSDs to Zoned Namespaces. In *Linux Storage and Filesystems Conference (Vault 19)*, 2019.
- [5] Simona Boboila and Peter Desnoyers. Write Endurance in Flash Drives: Measurements and Analysis. In *Proceedings of the 8th USENIX Conference on File and Storage Technologies (FAST '10)*, pages 115–128. USENIX Association, 2010.
- [6] Adam Brand, Ken Wu, Sam Pan, and David Chin. Novel read disturb failure mechanism induced by FLASH cycling. In *Proceedings of the 31st Annual International Reliability Physics Symposium*, pages 127–132. IEEE, 1993.
- [7] Yu Cai, Erich F Haratsch, Onur Mutlu, and Ken Mai. Error patterns in MLC NAND flash memory: Measurement, Characterization, and Analysis. In *Proceedings of the Conference on Design, Automation and Test in Europe*, pages 521–526. EDA Consortium, 2012.
- [8] Yu Cai, Yixin Luo, Erich F Haratsch, Ken Mai, and Onur Mutlu. Data retention in MLC NAND flash memory: Characterization, optimization, and recovery. In *21st International Symposium on High Performance Computer Architecture (HPCA)*, pages 551–563. IEEE, 2015.
- [9] Yu Cai, Onur Mutlu, Erich F Haratsch, and Ken Mai. Program interference in MLC NAND flash memory: Characterization, modeling, and mitigation. In *31st International Conference on Computer Design (ICCD)*, pages 123–130. IEEE, 2013.
- [10] Yu Cai, Gulay Yalcin, Onur Mutlu, Erich F Haratsch, Adrian Cristal, Osman S Unsal, and Ken Mai. Flash correct-and-refresh: Retention-aware error management for increased flash memory lifetime. In *30th International Conference on Computer Design (ICCD)*, pages 94–101. IEEE, 2012.
- [11] Paolo Cappelletti, Roberto Bez, Daniele Cantarelli, and Lorenzo Fratin. Failure mechanisms of Flash cell in program/erase cycling. In *Proceedings of the IEEE International Electron Devices Meeting*, pages 291–294. IEEE, 1994.
- [12] Peter Corbett, Bob English, Atul Goel, Tomislav Gracanac, Steven Kleiman, James Leong, and Sunitha Sankar. Row-diagonal parity for double disk failure correction. In *Proceedings of the 3rd USENIX Conference on File and Storage Technologies (FAST '05)*, pages 1–14. USENIX Association, 2004.
- [13] Robin Degraeve, F Schuler, Ben Kaczer, Martino Lorenzini, Dirk Wellekens, Paul Hendrickx, Michiel van Duuren, GJM Dormans, Jan Van Houdt, L Haspeslagh, et al. Analytical percolation model for predicting anomalous charge loss in flash memories. *IEEE Transactions on Electron Devices*, 51(9):1392–1400, 2004.
- [14] Jon G Elerath. AFR: problems of definition, calculation and measurement in a commercial environment. In *Annual Reliability and Maintainability Symposium. 2000 Proceedings. International Symposium on Product Quality and Integrity (Cat. No. 00CH37055)*, pages 71–76. IEEE, 2000.
- [15] Jon G Elerath. Specifying reliability in the disk drive industry: No more MTBF's. In *Annual Reliability and Maintainability Symposium. 2000 Proceedings. International Symposium on Product Quality and Integrity (Cat. No. 00CH37055)*, pages 194–199. IEEE, 2000.
- [16] Atul Goel and Peter Corbett. RAID triple parity. *ACM SIGOPS Operating Systems Review*, 46(3):41–49, 2012.
- [17] Dave Hitz, James Lau, and Michael A Malcolm. File System Design for an NFS File Server Appliance. In *USENIX Winter*, volume 94, 1994.
- [18] S Hur, J Lee, M Park, J Choi, K Park, K Kim, and K Kim. Effective program inhibition beyond 90nm NAND flash memories. *Proc. NVSM*, pages 44–45, 2004.
- [19] Seok Jin Joo, Hea Jong Yang, Keum Hwan Noh, Hee Gee Lee, Won Sik Woo, Joo Yeop Lee, Min Kyu Lee, Won Yol Choi, Kyoung Pil Hwang, Hyoung Seok Kim, et al. Abnormal disturbance mechanism of sub-100 nm NAND flash memory. *Japanese journal of applied physics*, 45(8R):6210, 2006.
- [20] Myoungsoo Jung and Mahmut Kandemir. Revisiting Widely Held SSD Expectations and Rethinking System-level Implications. In *Proceedings of the 2013 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS '13)*, pages 203–216, 2013.

- [21] Jae-Duk Lee, Chi-Kyung Lee, Myung-Won Lee, Han-Soo Kim, Kyu-Charn Park, and Won-Seong Lee. A new programming disturbance phenomenon in NAND flash memory by source/drain hot-electrons generated by GIDL current. In *Non-Volatile Semiconductor Memory Workshop, 2006. IEEE NVSMW 2006. 21st*, pages 31–33. IEEE, 2006.
- [22] Justin Meza, Qiang Wu, Sanjev Kumar, and Onur Mutlu. A Large-Scale Study of Flash Memory Failures in the Field. In *Proceedings of the 2015 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS '15)*, pages 177–190, 2015.
- [23] Iyswarya Narayanan, Di Wang, Myeongjae Jeon, Bikash Sharma, Laura Caulfield, Anand Sivasubramaniam, Ben Cutler, Jie Liu, Badriddine Khessib, and Kushagra Vaid. SSD Failures in Datacenters: What? When? And Why? In *Proceedings of the 9th ACM International on Systems and Storage Conference (SYSTOR '16)*, pages 7:1–7:11, 2016.
- [24] NetApp Inc. Data ONTAP 9. <http://www.netapp.com/us/products/platform-os/ontap/>.
- [25] David A. Patterson, Garth Gibson, and Randy H. Katz. A Case for Redundant Arrays of Inexpensive Disks (RAID). In *Proceedings of the 1988 ACM SIGMOD International Conference on Management of Data, SIGMOD '88*, page 109–116, New York, NY, USA, 1988. Association for Computing Machinery (ACM).
- [26] Eduardo Pinheiro, Wolf-Dietrich Weber, and Luiz André Barroso. Failure Trends in a Large Disk Drive Population. In *Proceedings of the 5th USENIX Conference on File and Storage Technologies (FAST '07)*, volume 7, pages 17–23, 2007.
- [27] Bianca Schroeder, Sotirios Damouras, and Phillipa Gill. Understanding Latent Sector Errors and How to Protect Against Them. *ACM Transactions on storage (TOS)*, 6(3):9:1–9:23, September 2010.
- [28] Bianca Schroeder and Garth 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 (FAST '07)*, volume 7, pages 1–16, San Jose, CA, 2007.
- [29] Bianca Schroeder, Raghav Lagisetty, and Arif Merchant. Flash Reliability in Production: The Expected and the Unexpected. In *Proceedings of the 14th USENIX Conference on File and Storage Technologies (FAST '16)*, pages 67–80, Santa Clara, CA, 2016. USENIX Association.
- [30] Zoned Storage. NVMe Zoned Namespaces. <https://zonedstorage.io/introduction/zns/>. Accessed: 2019-09-21.
- [31] Kang-Deog Suh, Byung-Hoon Suh, Young-Ho Lim, Jin-Ki Kim, Young-Joon Choi, Yong-Nam Koh, Sung-Soo Lee, Suk-Chon Kwon, Byung-Soon Choi, Jin-Sun Yum, et al. A 3.3 V 32 Mb NAND flash memory with incremental step pulse programming scheme. *IEEE Journal of Solid-State Circuits*, 30(11):1149–1156, 1995.
- [32] Hung-Wei Tseng, Laura Grupp, and Steven Swanson. Understanding the Impact of Power Loss on Flash Memory. In *Proceedings of the 48th Design Automation Conference (DAC '11)*, pages 35–40, San Diego, CA, 2011.
- [33] Wikipedia. S.M.A.R.T. <https://en.wikipedia.org/wiki/S.M.A.R.T.> Accessed: 2019-09-12.
- [34] Erci Xu, Mai Zheng, Feng Qin, Yikang Xu, and Jiesheng Wu. Lessons and actions: What we learned from 10k ssd-related storage system failures. In *2019 USENIX Annual Technical Conference (USENIX ATC 19)*, pages 961–976, Renton, WA, July 2019. USENIX Association.
- [35] Jimmy Yang and Feng-Bin Sun. A comprehensive review of hard-disk drive reliability. In *Annual Reliability and Maintainability Symposium. 1999 Proceedings (Cat. No. 99CH36283)*, pages 403–409. IEEE, 1999.
- [36] Mai Zheng, Joseph Tucek, Feng Qin, and Mark Lillibridge. Understanding the Robustness of SSDs Under Power Fault. In *Proceedings of the 11th USENIX Conference on File and Storage Technologies (FAST '13)*, pages 271–284, San Jose, CA, 2013. USENIX Association.

