Beyond MTTDL: A Closed-Form RAID 6 Reliability Equation

JON G. ELERATH, Reliability Consulting Services
JIRI SCHINDLER, NetApp

7

We introduce a new closed-form equation for estimating the number of data-loss events for a redundant array of inexpensive disks in a RAID-6 configuration. The equation expresses operational failures, their restorations, latent (sector) defects, and disk media scrubbing by time-based distributions that can represent non-homogeneous Poisson processes. It uses two-parameter Weibull distributions that allows the distributions to take on many different shapes, modeling increasing, decreasing, or constant occurrence rates. This article focuses on the statistical basis of the equation. It also presents time-based distributions of the four processes based on an extensive analysis of field data collected over several years from 10,000s of commercially available systems with 100,000s of disk drives. Our results for RAID-6 groups of size 16 indicate that the closed-form expression yields much more accurate results compared to the MTTDL reliability equation and matching computationally-intensive Monte Carlo simulations.

Categories and Subject Descriptors: B.3.2 [**Memory Structures**]: Design Styles—*Mass storage*; B.4.4 [**Input/Output and Data Communications**]: Performance Analysis and Design Aids—*Formal models, simulation*; B.4.5 [**Input/Output and Data Communications**]: Reliability, Testing, and Fault-Tolerance—*Hardware reliability, redundant design*; C.4 [**Computer Systems Organization**]: Performance of Systems—*Design studios, fault tolerance*

General Terms: Design, Algorithms, Reliability, Measurement

Additional Key Words and Phrases: Storage systems, RAID reliability

ACM Reference Format:
Elerath, J. G. and Schindler, J. 2014. Beyond MTTDL: A closed-form RAID 6 reliability equation. *ACM Trans. Storage* 10, 2, Article 7 (March 2014), 21 pages.
DOI:http://dx.doi.org/10.1145/2577386

1. INTRODUCTION

Parity RAID [Patterson et al. 1988] is arguably the most prevalent cost-effective mechanism for protecting online data against hard disk drive (HDD) failures. Each group uses either one (for RAID-4 or RAID-5) or two (for RAID-6) additional parity disks for every D disks with user data. Increased disk capacities lead to longer rebuild times with longer windows of opportunity for potential data loss due to another failure. RAID-6 variants, such as EVENODD [Blaum et al. 1994] or RAID-DP [Corbett et al. 2004], can protect against simultaneous partial HDD failures, called latent media defects or latent sector errors [Bairavasundaram et al. 2007]. It is this scenario—a latent sector error occurring concurrently with an operational HDD failure—against which the D+2 configurations protect.

Authors' addresses: J. G. Elerath, Reliability Consulting Services, Placerville, CA 95667; email: jon.elerath@gmail.com; J. Schindler (corresponding author), NetApp, 495 East Java Drive, Sunnyvale, CA 94089; email: jiri.schindler@netapp.com.
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
© 2014 ACM 1553-3077/2014/03-ART7 $15.00
DOI:http://dx.doi.org/10.1145/2577386

When designing storage systems, the reliability of the next-generation HDD is always a big unknown. System architects must also consider factors such as the number of HDDs in a RAID group, the priority given to reconstructing data after a failed HDD, and the frequency of the scrubbing process to remove latent media defects. They need an intuitive, easy-to-use, and accurate method for assessing the relative and absolute impacts that each of these factors will have on overall reliability.

The original RAID reliability equation, expressed as mean-time-to-data loss (MTTDL) [Patterson et al. 1988], can provide some guidance. However, it assumes the Poisson process for modeling HDD failures and does not adequately capture the complexities of modern systems with proactive scanning and repair mechanisms. Additionally, for RAID-6, it generates unrealistically high reliability numbers [Greenan et al. 2010], leading to large discrepancies between the predicted reliability and that experienced in the field. Previous research attempted to formulate more accurate models. However, they are inaccurate [Greenan et al. 2010], impractical to use, and require complex formulations with Markov chains [Malhotra and Trivedi 1993] or time-consuming Monte Carlo simulations [Elerath and Pecht 2009]. Additionally, Markov chains require constant transition rates that do not adequately capture events observed in the field.

This work defines a new closed-form equation for assessing the reliability of $D{+}2$ parity RAID groups. The equation allows for expression of time-variant HDD operational failure rates and accurately models the contributions of restoration algorithms, latent sector errors, and media scrub processes to overall RAID reliability. The equation expresses the time-based distributions of these processes with two-parameter Weibull distributions [Nelson 1982], which can take on many different shapes, modeling increasing, decreasing, or constant occurrence rates. We derived the parameters of these distributions from data collected over several years from 100,000s of disk drives installed in 10,000s of enterprise-class NetApp FAS storage systems running DataONTAP [NetApp 2013]. Our results for RAID-6 groups of size 16 indicate that the closed-form expression yields much more accurate results compared to the MTTDL reliability equation and matches computationally intensive Monte Carlo simulations.

Our equation is easy to use and offers easy formulation in basic tools. We have programmed it as a spreadsheet as well as a Web-based calculator. We made our implementation of the Web-based equation calculator publicly available to everyone at `http://raideqn.netapp.com` to provide system designers, who are not reliability experts, with the ability to readily explore the design space for their RAID solution and understand its performance and reliability trade-offs. We believe that our formulation, as well as the Javascript implementation that we made available, can put to rest the MTTDL RAID reliability equation that has been shown to be woefully inadequate for 21st century HDD and RAID technology [Greenan et al. 2010].

The rest of this article is organized as follows. In Section 2, we review relevant prior work. In Section 3, we present the attributes of the RAID-6 reliability equation, followed by a discussion of the equation inputs in Section 4. We formulate the equation and explain each of its terms in Section 5 and compare its accuracy to both the Monte Carlo simulations and actual field data collected from 100,000s of HDDs. Section 7 describes our Web-based Javascript implementation of the closed-form equation, and Section 8 concludes.

2. PREVIOUS WORK

There are many methods for estimating reliability of RAID systems. They express reliability as the mean-time-to-data loss (MTTDL) for $D{+}1$ and $D{+}2$ configurations [Dholakia et al. 2006; EMC 2007; Gao et al. 2010; Oracle 2010; Paris et al. 2009; Serve The Home 2011], the number of data loss events over time [Elerath and Pecht 2009;

Rao et al. 2006], or use a normalized magnitude of data loss as a replacement for MTTDL.

The original works on RAID used MTTDL as the metric for measuring reliability for $D+1$ RAID configurations [Gibson and Patterson 1993; Patterson et al. 1988]. MTTDL is an extrapolation of a well-known steady-state availability analysis technique [Bazovsky et al. 1961; Kececioglu 1993]. However, it only applies when there is a single device with a constant failure rate, when a system is known (not just assumed) to have a constant failure rate, and when the restorations are known to occur at a constant rate. These assumptions were reasonable for the level of technology and our understanding of modeling repairable systems in 1988; however, since then, HDD technology has evolved, the design of RAID systems has grown in complexity, and we have now a much better understanding of HDD failure modes, failure mechanisms [Elerath 2009b], and time-to-failure distributions.

Thompson [1981] and Ascher [1999] developed sound statistical bases and rules for modeling repairable systems and proved that the failure rate of a component is statistically different from the rate of occurrence of failure (ROCOF) for a repairable system. Nelson [2003] as well as Tobias and Trindade [2011] supported their findings using the mean cumulative failure function (MCF) for measuring the reliability of a population of fielded, repairable systems.

Ascher [1999] goes so far as to say that even if all the components in a system, HDDs in our case, have constant failure rates, there is no basis for assuming that the system will also have a constant failure rate. Several authors assert that "mean values" are not useful [Elerath and Pecht 2009; Greenan et al. 2010; Rao et al. 2006] even if they are statistically correct. Many studies have shown that HDD failure rates are not constant [Pinheiro et al. 2007; Shah and Elerath 2005; Schroeder and Gibson 2007], and logic dictates that restoration times are bounded on both ends of the distribution, which is not the case for exponential distribution.

Elerath and Pecht developed a time-sequential Monte Carlo simulation for RAID-4 [2009] to learn how the four underlying distributions impact the reliability of a RAID system. Elerath also developed an equation that fits the simulation [2009a], which defined the RAID reliability metric as the number of double-disk failures, or data-loss events, as a function of time. Our work builds on this model and extends it to RAID-6, the most common configuration in use or commercial parity-based RAID. To the best of our knowledge, our work is the first closed-form formulation of RAID reliability for a configuration protecting against double failures. We also extend previous work by extensive field data analysis to come up with realistic values of the closed-form equation parameters and put them into an easy-to-use Web-based format.

3. EQUATION ATTRIBUTES

A RAID reliability equation should (i) have a sound statistical basis, (ii) express a useful metric, (iii) accept time-dependent event rates, (iv) account for HDD capacity trends, and (v) have reasonable accuracy [Greenan et al. 2010]. All reliability equation formulations to date lack some of these properties to varying degrees. This section illustrates their shortcomings and motivates the need for a closed-form expression.

3.1. Statistical Basis

As discussed previously, many authors assume that if each HDD in the RAID group has a constant failure rate, then the times between failures for the RAID group are also constant [Gibson and Patterson 1993; Rao et al. 2006]. They use this to assert that

Table I. Failure Times and Times between Failures

Failure Number	Ordered Time to Failure	Time Between Failures
1	224	224
2	3,430	3,206
3	3,548	118
4	4,448	899
5	4,768	320
6	4,784	16
7	5,594	810
8	5,718	124
9	6,212	494
10	6,422	211

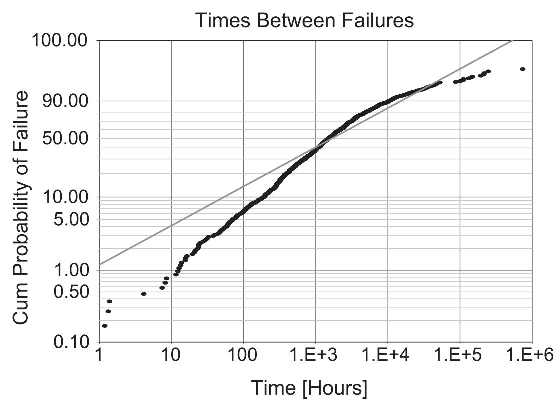

Fig. 1. Plot of times between failures.

repairable RAID group behavior can be predicted based on HDD failure distributions. However, a basic experiment shows this is not true [Ascher 1983].

Let's consider an exponential distribution for *time to failure* with a mean of 1,000,000 hours and 1,000 failure events. To generate the data, we use a Weibull distribution with $\beta = 1$, which degenerates to an exponential distribution. Table I lists ten of those 1,000 failure events with the shortest time to failure and the corresponding (observed) times between failures. However, when we plot these 1,000 times between observed failures using Weibull Plot [Nelson 1982], the result is not an exponential distribution, that is, a distribution with a constant failure rate.

If the observed times between failures were to follow an exponential distribution, they would fit a straight line. This line, shown in Figure 1, represents a two-parameter Weibull distribution with $\beta = 1$ and $\eta = 1{,}000{,}000$. Note that a Weibull distribution can be used to express time-dependent distributions with decreasing ($\beta < 1$), constant ($\beta = 1$), or increasing ($\beta > 1$) rates.

Thus, if 1,000 HDDs are all put into service at the same time and removed when failed, as would be done in a simple nonredundant system, the time between HDD failures would not be exponential, and the failure rate would not be constant. The same would be true for a system that contains redundancy, such as parity-based RAID; there is simply no basis to assume that the times between failures for a repairable system with redundancy would follow an exponential distribution and, therefore, will not have a constant failure rate.

This simple example shows that even if every component in a system has a constant failure rate, there is no statistical basis for the assertion that the rate of occurrence of failure for a repairable system is constant. Ascher [1983] illustrates how a repairable system can have an increasing or decreasing rate of occurrence of failure at the system level, even though the components in the system are all from populations with constant failure rates. Previous work developed the statistical basis for the concepts illustrated by our preceding example [Ascher 1999; Thompson 1981].

3.2. Metric

A mean value is in itself not useful for estimating RAID group reliability regardless of how the mean is determined. By definition, the mean, also called the expected value, is the integral of the probability density function (PDF) over the entire range for which it is defined. For time-bound functions in which t is the variable of interest and its minimum is 0, the equation is

$$E(t) = \int_0^\infty t \cdot f(t)dt. \tag{1}$$

Based on the mean value alone, it is not possible to tell if the failure rate is decreasing, constant, or increasing in time. If the mean time is, say, 1,000 years, but the useful mission life of a RAID system is 5 to 7 years before it is retired, there is no way of knowing how many failures there will be during this time span.

As illustrated earlier, we cannot assume a constant failure rate. An exponential distribution has a constant failure rate, but integrating the PDF from 0 to the mean value renders a 63.2% chance of failure by the MTTF.

$$F(MTTF) = \int_0^{MTTF} e^{-t/MTTF}dt = 1 - e^{-1} = 0.632. \tag{2}$$

This means 63.2% of the failures will occur before the mean value. If the time-to-failure has a decreasing failure rate, however, the percentage of HDDs that fail during the first 5–7 years may be much greater than the percentage calculated assuming a constant failure rate. We can calculate MTTF for a Weibull distribution,

$$MTTF = \gamma + n \cdot \Gamma\left(\frac{1}{\beta}\right) + 1, \tag{3}$$

where $\Gamma(z)$ is the gamma function of z, η is the characteristic life, β is the shape parameter, and γ is the location parameter of the distribution. When $\beta = 0.5$, which implies a decreasing failure rate, then

$$\Gamma\left(\frac{1}{0.5} + 1\right) \;=\; \Gamma(3) \;= (3-1)! \;=\; 2. \tag{4}$$

If $\gamma = 0$ and $\eta = 876{,}000$ hours (10 years), then we get MTTF = 876,000*2 = 1,752,000 hours. Some authors assume that the failure rate is the inverse of the mean and estimate the number of failures by taking that failure rate and multiplying by the number of hours in the mission. For this example and a mission time of 61,320 hours (7 years), this erroneous calculation estimates only 3.5% of failures. If, on the other hand, we use the Weibull distribution with the preceding parameters, then 23.2% of the systems are expected to fail—a value that is 6.5 times larger!

3.3. Time-Dependent Input Distributions

HDD failure rates are often not constant [Malhotra and Trivedi 1993], even though assuming exponential distributions may make reliability computations easier. However, distributions with time-dependent or non-constant failure rates are a critical factor in deriving a closed-form RAID equation with good accuracy.

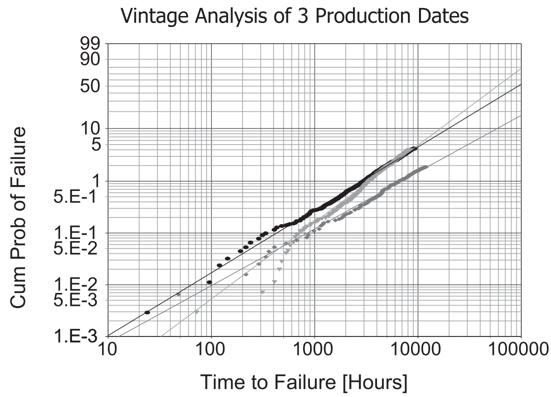

Fig. 2. Effects of vintage on HDD reliability. Reprinted with permission from Elerath and Pecht [2009].

Consider the Weibull probability plots in Figure 2 [Elerath and Pecht 2009]. They show field data for one HDD model from a single supplier grouped into several different vintages based on production date. The earlier vintages have a constant failure rate, while the later vintages have increasing failure rates ($\beta > 1$). RAID designers need to realize that different vintages will affect overall reliability.

3.4. HDD Capacity

The most common factor contributing to data loss includes a latent media defect. So-called grown media defects or foreign particles can cause sectors to be unreadable or contain corrupted data. As HDD capacity continues to grow at the annual compound growth rate of ~30%, the opportunity for corrupted data increases and so does the length of time necessary to scrub the media—a process that periodically reads every sector and remaps corrupted ones to spare ones in other areas of the media.

RAID reliability equations expressing MTTDL do not account for capacity in terms of either scrub time or defect frequency. Our equation, on the other hand, accounts for HDD capacity as it considers both the likelihood of latent media defects and the rate of repairing them through periodic media scrubbing.

3.5. Accuracy

An equation for driving HDD selection and RAID architecture must be compared to field data and shown to be consistent. Elerath and Pecht [2009] showed that the MTTDL equation for RAID-4 can be off by 3,000 times or more and concluded that the MTTDL metric is too inaccurate to be useful or practical when considering the design of a RAID system. Elerath [2009a] developed an alternate model with errors less than 10%. Our work achieves a similar level of accuracy for a more complex $D+2$ RAID-6 configuration (see Section 6).

4. RAID-6 EQUATION PRELIMINARIES

There is no exact statistical basis for a closed-form equation that models m-out-of-n failures with repair. It is not possible to begin with fundamentals such as PDFs and derive a closed-form equation, as Ascher [1983] has already noted. However, he also concurred that, if the final shape is determined by experiment, field data, or simulation, then an equation that fits the "known correct" system response is a reasonable way to model that system [Ascher 2010].

Our approach follows that process. First, we model and analyze a RAID-6 system using a Monte Carlo (MC) simulator. Elerath and Pecht [2009] have shown this to be an accurate albeit very time-consuming process; it takes between 14 seconds and 18 hours to produce a result for a single set of inputs (i.e., parameters of the four time-dependent distributions). Next, we formulate our equation and compare its results to the field data gathered from tens of thousands of DataONTAP storage systems.

4.1. Assumptions

We first list some of the assumptions we make about the operation of a RAID system in order to formulate our closed-form RAID-6 reliability equation.

4.1.1. RAID System Operation. RAID-6 systems operate 24 hours a day, 365 days a year. After writing data to a HDD, they do not typically verify the content immediately, because the resulting slowdown would be too high. Instead, they discover and correct errors the next time the data is accessed or scrubbed. All HDDs come from the same population and follow the same time-to-failure distribution.

4.1.2. HDD Failures and Latent Media Defects. An operational failure is a failure in which no data on an HDD can be accessed. A latent defect is an undiscovered data corruption that can be corrected by reading the stripe data from the remaining HDDs and reconstructed via parity information. Data can become corrupted either during the write process or after being successfully written due to media contamination.

4.1.3. Restoring Failures. When an operational failure occurs, a storage system incorporates a spare HDD into the RAID group and reconstructs the missing data. The time required for this process depends on the foreground activity, the capacity of the HDD, and the data transfer rate of the HDD. The time required to restore an operational failure has a minimum value that is greater than zero. Similarly, the rate has a finite upper bound, since the restore operation eventually finishes. An exponential distribution, on the other hand, has an unbounded upper value.

Scrubbing is a background activity that proactively reads data from all HDD blocks. It checks the data against the checksum verifier information, typically collocated with the data on an adjacent sector. If the check fails, the system recovers the corrupted data by reading data and parity from the other disks in the RAID stripe and rewrites them to the same physical location on the HDD. If the media is permanently defective, the HDD firmware transparently remaps the affected data to a new location within spare sectors.

The time required to scrub an entire HDD is a random variable that depends on the HDD capacity and the rate of media scrubbing. If not scrubbed, latent defects will persist (accumulate) over the life of the HDD. Since scrubbing eliminates a partial HDD failure due to a latent defect, it reduces the probability of simultaneously having a latent defect during a RAID reconstruction after an operational failure.

4.1.4. Data-Loss Event Due to RAID Group Failure. When a single stripe has corrupted data across at least three HDDs, the RAID-6 group cannot recover that stripe, so data is lost. This constitutes a RAID group failure. Data loss can occur from two catastrophic HDD failures and one latent defect, or three catastrophic failures. The probability of two or three latent defects in the same RAID stripe is too small to be considered.

4.2. Equation Inputs

We use four random variables as inputs to calculate the reliability of a RAID group. Three of the four variables—time-to-operational failure, T_{Op}, time-to-restore the operational failure, T_{Rest}, and time-to-scrub, T_{Scrub}—have time-dependent occurrence

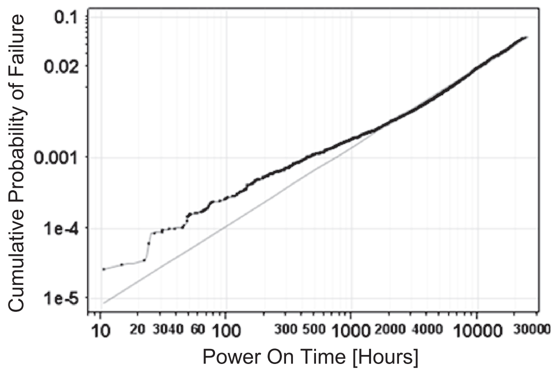

Fig. 3. Time-to-operational failure.

rates. The occurrence rate for latent defects, T_{Ld}, is constant. The following paragraphs detail these variables and their parameters.

4.2.1. Operational Failure Distribution. When the HDD is incapable of reading or writing, an operational failure occurs. This failure is the most obvious and most readily detected and constitutes the most common type of information collected in the field. Some systems also keep track of the number of hours of field operation for every failed and non-failed HDD and thus, with correct analysis, allow us to determine the time-to-failure. We use data collected from commercial storage systems with which we can determine these distributions and use them in our simulation and formulation of the closed-form equation.

Figure 3 shows the distribution for one SATA model. Over 283,000 HDDs of this model were in the field, and 9,065 failed during the observed 2.5 years. For this data, the shape parameter, β, is 1.13 and the characteristic life, η, is 320,810 hours. The annualized return rate was measured at 4.4% and increasing. Section 6 shows Weibull distribution parameters for all four input parameters for other disk models.

4.2.2. Restore Distribution. Generally, data for the time-to-restore an operational failure is more difficult to collect in the field, because it requires the RAID controller to log the information about when the error occurred and when the RAID reconstruction started and finished. In the original MTTDL equation, the mean time-to-repair included the time to physically replace the failed HDD.

The DataONTAP systems from which we obtained our data include spare HDDs. The system can assign a new drive from the spare disk pool and immediately start a restore operation. The logs include the information about how long the reconstruction took. Figure 4 shows the time-to-repair distribution for one HDD model; 90% of all drives of this model are restored within 38 hours. The distribution is also a function of how busy the systems are, which affects the rate with which the RAID repair proceeds. Naturally, systems with light activity will complete repairs sooner.

4.2.3. Latent Defect Generation. For latent defect errors, we must assume a constant rate of occurrence. Unless the RAID system can read data just after writing, which is too costly, it is not possible to know whether the corruption occurred upon writing or some time later. All we know is that the data was corrupted when read. Thus, we express the latent defect occurrence rate as a function of total data read and written over time in service.

The NetApp DataONTAP system collects the following information about each HDD in the system: power-on hours (POH), reported media errors at the sector level (ME),

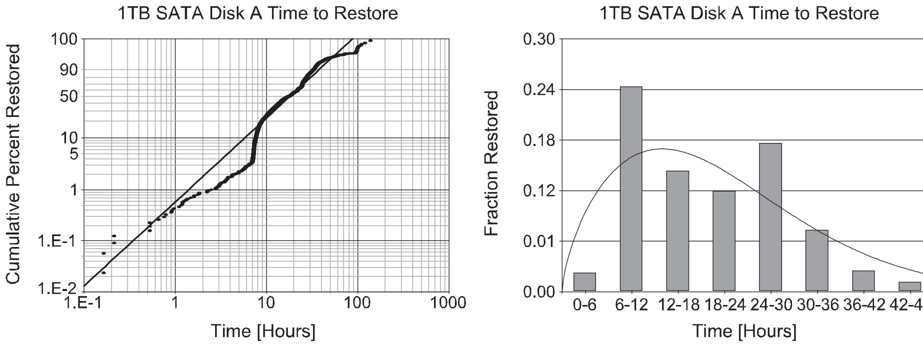

Fig. 4. Time-to-restore distribution for a single SATA HDD model.

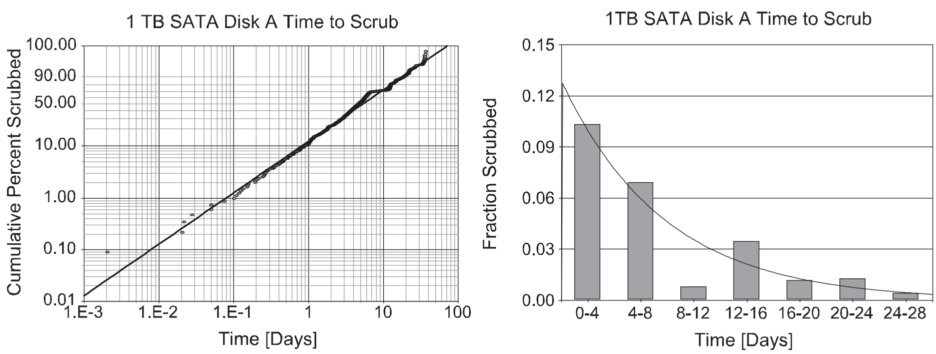

Fig. 5. Time-to-scrub distribution for a single SATA HDD model.

total bytes written (BW), and total bytes read and written (BRW). From this data, we express the mean time-to-latent-defect, T_{Ld}, by dividing the TOTAL_POH value by the observed TOTAL_ME. The TOTAL_value for each variable is the sum of individual values collected for each HDD by the RAID system.

An alternative, but less accurate, method for calculating T_{Ld} is to use the HDD manufacturer's bit-error rate (BER) and knowing the amount of data read or written per unit of time. For example, if the manufacturer's specification states one uncorrectable error in 10^{16} bits and we assume an average of 5MB accessed per second,

$$T_{Ld} = \frac{10^{16}\text{ bits}}{\text{UCE}} \times \frac{\text{sec}}{5{,}000{,}000\,\text{Bytes}} = 69{,}444 \text{ hours}.$$

This equation can be use when field data is not available, and we can estimate the average load on the HDD. However, we observed that manufacturer-specified values of BER are often very different from field data.

4.2.4. Scrub Distribution. Disk media scrubbing is performed in the background as a lower-priority process to the user-initiated reading and writing processes. The length of time required for scrubbing depends on the priority set by the RAID controller and the capacity of the HDD. Figure 5 shows the scrubbing time distribution derived from our field data and system logs for a single HDD model. Interestingly, β for this distribution is 1.0. Other HDDs in our field data had $\beta > 1$. For this model, 90% are scrubbed within 17 days.

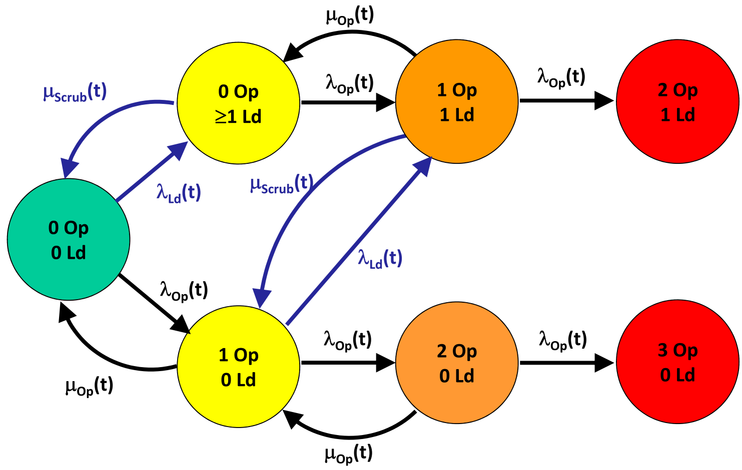

Fig. 6. Allowable transition diagram (not a Markov diagram) for time-dependent failures and restorations. Failure processes are denoted by $\lambda(t)$, and restoration processes ar denoted by $\mu(t)$. Note that this is not a Markov chain, as the state transitions are time-dependent variables. For clarity, we omit the D+2, D+1, and D multipliers in front of the state transition variables.

5. THE RAID-6 EQUATION

We now formulate the RAID-6 reliability equation. Since our goal is a closed-form formulation that models time-varying distributions yet allows for simple expression in readily available tools, we do not consider models based on generalized semi-Markov processes. Such formulations are complex and thus not convenient for an easy-to-run analysis.

We start by considering the various states for a RAID-6 group. Our equation computes the expected number of data-loss events, expressed as the number of simultaneous partial or whole triple-disk failures.

5.1. RAID-6 States

A RAID system is continuously succumbing to HDD failures and latent defects and having these restored and scrubbed. Each time a failure or defect occurs, the RAID group enters into a degraded mode. This process continues throughout time until an operational failure occurs when the RAID group is already in a degraded mode by having an operational failure and a latent defect, or two operational failures.

Figure 6 shows the state transitions for a RAID-6 group that our equation considers. Note, however, that the diagram is not a Markov chain, because many of the transitions do not have constant rates. Similarly, it leaves out some of the would-be transitions, as we describe next. The initial state is at the far left, with no operational failures and no latent defects. The next state has either one or more latent defects (upper path), or one operational failure (lower path). A second operational or latent-defect failure event produces a transition to another state that we call degraded mode. From these two states, yet another operational failure results in either of the two data-loss states on the far right. This logic forms the number of times we expect to see three disks that concurrently have data corrupted or unavailable for access due to Ld or Op events. We term these two terminal states the expected number of triple-disk failures, $N_{TDF}(t)$, which is a function of time in service.

Given two types of failure events, Ld and Op, and three concurrent failures, there are in theory eight possible permutations of events that could lead to a data-loss event. However, assuming that the likelihood of two Ld failures occurring concurrently in the same stripe is negligible and leveraging the fact that having two simultaneous *Ld* failures in different stripes does not cause data loss, we do not consider permutations with two simultaneous *Ld* events: *Op-Ld-Ld*, *Ld-Op-Ld*, *Ld-Ld-Op* and *Ld-Ld-Ld*.

Recall that RAID arrays reconstruct data by computing parity across a stripe of data (typically 4–256KB). Similarly, we do not consider the *Op-Op-Ld* sequence, which corresponds to a scenario of undergoing a RAID reconstruction on account of having experienced two simultaneous *Op* failures already and only then developing a latent defect on the two disks to which we are writing anew. Hence our diagram does not include a transition from the "2 Op 0 Ld" state in the bottom to the "2 Op 1 Ld" terminal state. Note that we do model the sequences of whole or partial disk failure events *Op-Ld-Op*, *Ld-Op-Op*, where having an *Ld* failure already while experiencing two simultaneous *Op* failures leads to a data-loss event. The bottom row of the state transitions captures the eighth and final possible permutation *Op-Op-Op*.

5.2. Terms of the Equation

Given the states and sequence of failure events just described, our RAID-6 reliability equation includes a product of two terms: a term for being in a degraded mode (two concurrent failures) and a term for the expected number of additional HDD failures given the group is in degraded mode. The overall closed-form RAID-6 reliability equation expresses the expected number of RAID-6 data loss events as

$$N_{TDF}(t) = (DM_1 + DM_2) \cdot D \cdot H(t), \tag{5}$$

where *DM1* is the probability of being in "degraded mode 1", *DM2* is the probability of being in "degraded mode 2", D is the number of data disks in the RAID group, and $H(t)$ is the time-dependent hazard function for the HDDs. Each of these terms is explained next.

5.3. Pseudo-Characteristic Life

Our RAID-6 equation allows us to calculate the number of expected data-loss events, expressed as a function of time, for many different HDD time-to-failure distributions. We use the Weibull distribution to represent the time-to-operational failure, because it can model decreasing, constant, and increasing failure rates. The Weibull cumulative density function (CDF) is

$$F_{Weibull}(t) = e^{-\left(\frac{t}{\eta}\right)^{\beta}}, \tag{6}$$

where t is time, β is the parameter controlling the shape of the distribution, and η is the characteristic life, which is always the time at which 63.2% units have failed. When $\beta = 1$, the Weibull distribution degenerates into the exponential distribution in which the characteristic life is the mean time before failure (MTBF).

The degraded mode terms cannot factor in the characteristic life, η, unless the shape, β, is unity. Therefore we must calculate an "equivalent" characteristic life that results in the same cumulative probability of failure, $F(t) = 63.2\%$, as for the time-dependent CDF at the same point in time. Basically, this amounts to approximating a curve with a straight line. We term this approximation the pseudo-characteristic life, η_{pseudo}, for the Weibull distribution and calculate it

$$\eta_{pseudo} = \frac{t}{\left(\frac{t}{\eta}\right)^{\beta}} = \frac{\eta^{\beta}}{t^{\beta-1}}. \tag{7}$$

Figure 7 shows the pseudo-characteristic life for four different profiles of 1-, 5-, 7-, and 10-year mission time, respectively. Note that even though the approximations

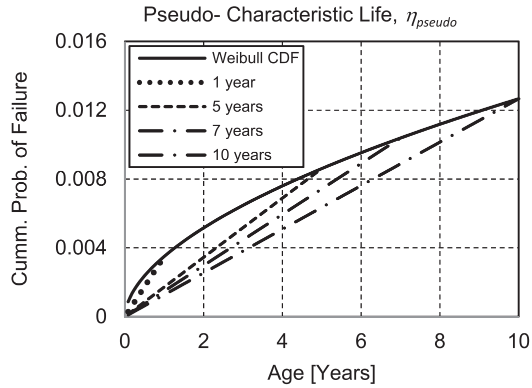

Fig. 7. Pseudo-characteristic life approximation to the true characteristic life.

are straight lines, they do account for non-constant failure rates, because the η_{pseudo} expression depends on the shape parameter β, which captures the time-dependent nature of the distribution. Stated differently, it models situations where failures are more or less likely to occur towards the end of the expected mission time.

5.4. Degraded Mode

Degraded Mode-1 exists when both an operational failure and a latent defect occur concurrently. Since we do not know which event occurs first, we approximate this probability. Therefore, we calculate the average of the contributions from both orders of occurrence to get

$$DM_1 = \frac{DM_{Op-Ld} + DM_{Ld-Op}}{2}. \tag{8}$$

DM_{Op-Ld} is the probability of being in the degraded mode due to an operational failure followed by a latent defect. DM_{Ld-Op} is the probability of having a latent defect followed by an operational failure. The fundamental statistical basis for each factor for both of these probabilities comes from the concept that, in the long run, the probability of being in a failed state can be calculated using the characteristic lives for time-to-failure and time-to-repair as

$$DM_{Op-Ld} = \left[1 - \left(\frac{\eta_{Op}}{\eta_{Op} + \eta_{Rest}}\right)^{D+2}\right] \cdot \left[1 - \left(\frac{\eta_{Ld}}{\eta_{Ld} + \eta_{Scrub}}\right)^{D+1}\right], \tag{9}$$

and

$$DM_{Ld-Op} = \left[1 - \left(\frac{\eta_{Ld}}{\eta_{Ld} + \eta_{Scrub}}\right)^{D+2}\right] \cdot \left[1 - \left(\frac{\eta_{Op}}{\eta_{Op} + \eta_{Rest}}\right)^{D+1}\right], \tag{10}$$

where η_{Op} is the pseudo-characteristic life for operational failures, η_{Ld} the characteristic life of latent defect distribution, η_{Scrub} the characteristic life of the scrub distribution, and η_{Rest} the characteristic life for the restoration distribution.

Degraded Mode-2 exists when two operational failures occur concurrently. We can express it as

$$DM_2 = \left[1 - \left(\frac{\eta_{Op}}{\eta_{Op} + \eta_{Rest}}\right)^{D+2}\right] \cdot \left[1 - \left(\frac{\eta_{Op}}{\eta_{Op} + \eta_{Rest}}\right)^{D+1}\right]. \tag{11}$$

Order does not influence the probability, since both probabilities are based on the same underlying distributions for failure and restoration.

5.5. Number of Operational Failures

The last term of our RAID-6 equation is the number of operational failures that occur as a function of time. As stated previously [Nelson 1990], for any distribution, the cumulative hazard function, $H(t)$, can be calculated by

$$H(t) = -\ln\left[R(t)\right], \tag{12}$$

where $R(t)$ is the reliability function. For the Weibull distribution in the RAID-6 notation, the cumulative hazard function is

$$H(t) = \left(\frac{t}{\eta_{Op}}\right)^{\beta}. \tag{13}$$

The cumulative hazard function is not a probability, that is, it does not have an upper bound and can exceed unity. For the equation, the number of HDDs in the RAID group also impacts the number of expected operational failures, which is the reason for multiplying $H(t)$ by D, the number of remaining operational HDDs.

5.6. Modeling More Detailed Features

The goal of our work is not to formulate the most detailed and thus likely complex-to-express and possibly complicated-to-use model. Instead, we strive to create a formulation that allows easy manipulation akin to the original MTTDL equation, but one that is much more accurate.

There are many RAID system features and possibly several HDD failure modes that our formulation does not account for. For example, we do not model environmental effects, such as humidity or temperature, that would likely manifest as correlated failures. However, we do allow for modeling correlated failures due to manufacturing variabilities; our equation allows us to change the operational failure Weibull distribution parameters and rerun the analysis for different setups with different characteristics.

Similarly, our formulation allows the expression of I/O load-dependent variables and their effects on overall RAID system reliability. The bit-error rate BER is load-dependent, and the Weibull parameters for scrubbing and operational failure repairs can be adjusted accordingly to model scenarios, where the rate of completions for these repair processes depends on the level of foreground workload. In providing the parameters from our analysis, we grouped all systems together simply to provide a single distribution for these processes.

6. EVALUATION

In evaluating the prediction accuracy of our closed-form RAID-6 reliability equation, we consider three complementary methods. First, we compare the results from the equation to those from almost 100 Monte Carlo simulations. Then, we compare the results of the equation to field data and show they are consistent. Finally, we compare the results from the equation to those obtained by the MTTDL method and illustrate the superiority of the equation over the MTTDL.

6.1. MC Simulation Comparison

We demonstrate the accuracy of our closed-form equation by first comparing its results to Monte Carlo (MC) simulation and then through comparison to our field data

Table II. Parameters for MC Simulation Cases

		Operational					Latent				
		Op. Fr. Dist.		Op. Fr. Rstr.			Defect Dist.		Scrub Dist.		
	Case	η	β	γ	η	β	η	β	γ	η	β
Decreasing Failure Rate	16–2	58,400	0.85	12	36	3	9,259	1	12	84	3
	16–6	58,400	0.85	12	36	3	46,295	1	0	48	3
	16–10	58,400	0.85	12	12	2	9,259	1	12	336	3
	16–14	58,400	0.85	12	12	2	46,295	1	12	84	3
	16–18	58,400	0.85	12	12	2	185,180	1	0	48	3
	16–22	58,400	0.85	0	12	3	46,295	1	12	336	3
	16–26	58,400	0.85	0	12	3	185,180	1	12	84	3
Constant Failure Rate	16–30	876,000	1	12	36	3	9,259	1	0	48	3
	16–34	876,000	1	12	36	3	185,180	1	12	336	3
	16–38	876,000	1	12	12	2	9,259	1	12	84	3
	16–42	876,000	1	12	12	2	46,295	1	0	48	3
	16–46	876,000	1	0	12	3	9,259	1	12	336	3
	16–50	876,000	1	0	12	3	46,295	1	12	84	3
	16–54	876,000	1	0	12	3	185,180	1	0	48	3
Increasing Failure Rate	16–56	876,000	1.2	12	36	3	9,259	1	12	84	3
	16–60	876,000	1.2	12	36	3	46,295	1	0	48	3
	16–64	876,000	1.2	12	12	2	9,259	1	12	336	3
	16–68	876,000	1.2	12	12	2	46,295	1	12	84	3
	16–72	876,000	1.2	12	12	2	185,180	1	0	48	3
	16–76	876,000	1.2	0	12	3	46,295	1	12	336	3
	16–80	876,000	1.2	0	12	3	185,180	1	12	84	3

collected from thousands of deployed systems with RAID-6 configuration of group size $D + 2 = 16$.

6.1.1. Synthetic Weibull Distribution Parameters.. We adapted the Monte Carlo simulation developed previously for RAID-4 by Elerath and Pecht [2009]. We created 81 different combinations of the four input distributions using a broad range of values for the Weibull distribution parameters β and η, and ran each case through the MC simulation program. This evaluation was the most time-consuming part of our work. For brevity, we show here the comparison to our equation results against a subset of 21 representative MC simulations.

Table II lists the input parameters for our 21 simulated cases. There are seven comparisons with decreasing hazard rates, seven with constant rates, and seven with increasing rates. For each of the three groups, we use synthetic albeit representative values for the Weibull distribution parameters. For example, we consider the values of 12 and 36 hours for the η_{Rest} parameter of the time-to-restore the operational failure distribution, T_{Rest}, and 48, 84, and 336 hours (i.e., two days, half a week, and two weeks) for the η_{Scrub} parameter of the time-to-scrub distribution, T_{Scrub}. The other 60 cases not shown here included three different restoration, defect, and different scrub distributions; they produced similar results.

Table II also includes the Weibull location parameter, γ, which is an offset in time. For example, if a single disk cannot be restored in less than 12 hours, then the offset would be 12 hours. Operational failures and latent defects do not have minimum times before which a failure cannot occur, so for these distributions, $\gamma = 0$.

Figure 8 shows the accuracy of our closed-form equation. Rather than listing the expected number of triple failures in absolute terms, we express the accuracy as relative

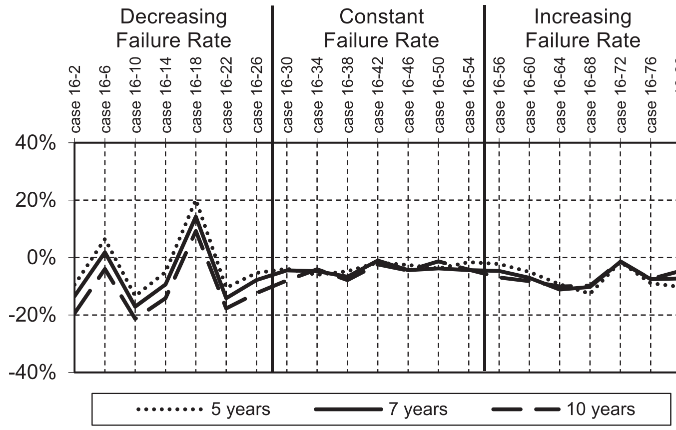

Fig. 8. Relative error to MC simulations.

error, or the difference, between the results obtained from the MC simulation and from our equation for each of the 21 cases. The graph shows three lines, each representing a system with mission time of 5, 7, and 10 years respectively.

Even though the comparisons are discrete points, we connect them with lines to better illustrate that the trends are similar for each of the mission time series. For all data points the relative error, calculated as the difference between equation value and simulation value divided by the simulation value, is always less than 20%, in many cases less than 10%. Since we used a wide range of input distributions across the 81 simulation cases, we believe the error will generally be less than +/– 20%.

The closed-form equation yields significant time savings over the MC simulation. Our Javascript implementation, described in Section 7, calculates results with no human-perceivable delay (approximately 0.1ms). A single MC simulation, on the other hand, ran between 14 seconds and 18 hours, depending on the parameters of the input distributions. The more reliable configurations required $1\text{x}10^9$ evaluations, each of which consisted of multiple samplings of data for each distribution.

6.1.2. HDD Field Data Analysis Parameters. As mentioned previously, we analyzed field data collected from 10,000s of commercial enterprise-class storage systems. The data includes detailed RAID configuration information, including the time-in-operation for every disk in the system, logs with occurrences of operational disk failures, latent sector errors, and times for each RAID repair and media scrub operation. The most prevalent configuration in the deployed systems is a variant of RAID-6 with 16 disks ($D = 14$). Table III shows the Weibull distribution parameters we obtained from our analysis of three representative disk models—two near-line SATA 1TB models and one enterprise-class FC 288GB model—that have been in the field for several years. Figure 4 and Figure 5 plot some of these distributions for the 1TB SATA Disk A.

The distributions for each disk model are very different. In particular, the 1TB SATA disk A model has much higher operational failure rates than the “more reliable” disk B. Similarly, the T_{Rest} and T_{Scrub} distributions for the two different 1TB SATA HDDs are different, because for our analysis, we first divided the sample population of our 10,000s of systems into different groups based on the disk drive models they contained and only then extracted the respective distribution parameters, as described in Section 4. Given that both T_{Rest} and T_{Scrub} have higher values for the characteristic

Table III. Parameters for MC Simulation Cases

		SATA Disk A	SATA Disk B	FC/SCSI Disk C
T_{Op}	η_{Op}	302,016	4,833,522	1,058,364
	β_{Op}	1.13	0.576	0.721
T_{Rest}	η_{Rest}	22.7	20.25	6.75
	β_{Rest}	1.65	1.15	1.4
T_{Scrub}	η_{Scrub}	186	160	124
	β_{Scrub}	1	0.97	2.1
T_{Ld}		12,325	42,857	50,254
Capacity		1 TB	1 TB	288 GB
ARR		4.4%	0.2%	0.4%
Time in field		3 yrs	3 yrs	5 yrs

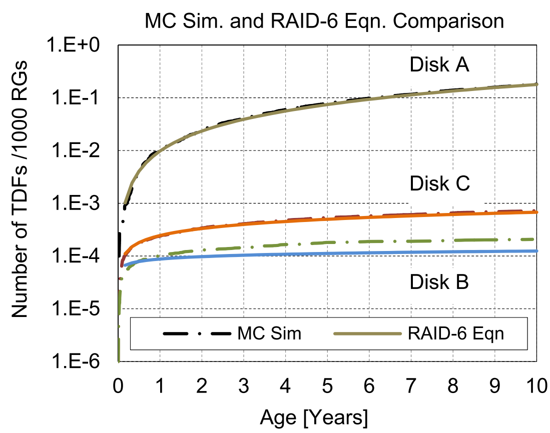

Fig. 9. Comparison of Monte Carlo simulations configured with field-obtained HDD parameters and the closed-form RAID-6 equation for the three disk models. For disks A and C, the two lines coincide. For disk B, Monte Carlo simulation predicts a slightly higher number of triple-disk failures (TDFs). The dash lines labeled “MC Sim” are results from the Monte Carlo simulation. The y-axis is log base 10 rather than linear.

life η, we can only speculate that the systems with SATA disk A experienced generally higher loads and thus took longer to finish the media scrubbing and RAID restoration processes. We did not correlate this information from the load observed by these systems.

We used these field-derived parameters for the four distributions as inputs to MC simulations. Figure 9 compares the simulation results with our closed-form equation calculations for the three HDD families. Since the inputs were from real HDDs, we show the results in absolute terms, that is, we plot the expected number of data-loss events due to three failures as a function of time for 1,000 RAID-6 groups with 16 disks ($D = 14$). We chose 1,000 groups to model a deployment typical for a single data center (in this case, with 16,000 disks).

As before, the simulation results compare favorably to the results from our closed-form equation. The greatest error of 43% occurred for disk B, which is the most reliable HDD of the three models with an annualized return rate (ARR) of 0.2%. The error for disk A, which has the highest ARR (4% and increasing), is 1.4%, while for disk C, which has an ARR of 0.4%, it is 6.4%.

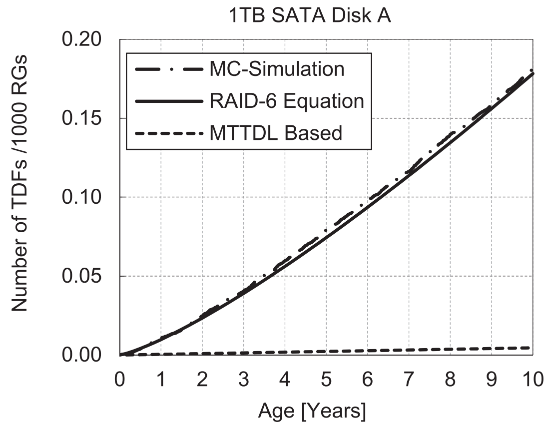

Fig. 10. Comparison of the Monte Carlo simulation, the MTTDL equation, and the closed-form RAID-6 equation results. MTTDL predicts much lover values compared to the other two methods.

6.2. Comparison to MTTDL

We now explore the question whether our equation is "better" than MTTDL for predicting the expected number of data-loss events, expressed as triple-disk failures (TDF). The equation for expressing MTTDL is

$$MTTDL = \frac{MTBF^3}{(D+2)(D+1)(D)MTTR^2}, \tag{14}$$

where *MTBF* is the mean time between failures for a single HDD, *MTTR* is the mean time to restore data on one HDD from parity information, and D is the number of data disks in the group.

Using the inverse of the MTTDL as a rate of occurrence of failure for the RAID group, which is erroneous, but done regularly, and multiplying by the time in the field, results in a linearly increasing function, as shown in Figure 10. The predicted number of TDFs in ten years is so small for the MTTDL that it appears to be on the horizontal axis. Earlier, the true number of TDFs for SATA disk A was greater than the equation predicted. The MTTDL is far lower than the equation, so if the equation is in error, it is still far better than the MTTDL.

6.3. Field Data Comparison

The previous section compared our equation to MC simulation. We now compare the equation-predicted results to field data of observed triple-disk failures. Generally, such comparisons are very difficult to make; RAID-6 is so reliable that few data-loss events occur, even though we have data from a large population of deployed systems. RAID-6 groups composed of SATA disk A and FC disk C showed only six data-loss events in the combined 15,000 groups of size 13, 14, and 16. Recall that disk A and disk C have been in the field for 3 and 5 years, respectively.

In our data, RAID groups of size 16 with disk A had only one TDF in over three years, while our closed-form equation predicts 0.12 TDFs for the same RAID group size and HDD population. This discrepancy occurs because we are dealing with the "tail of the distribution", which often is not representative of the true distribution. With only a single TDF, we expect poor correlation to the equation. However, if RAID-6 were not as reliable as predicted by the MC simulation and the close-form equation, there would be far more data-loss events in the field.

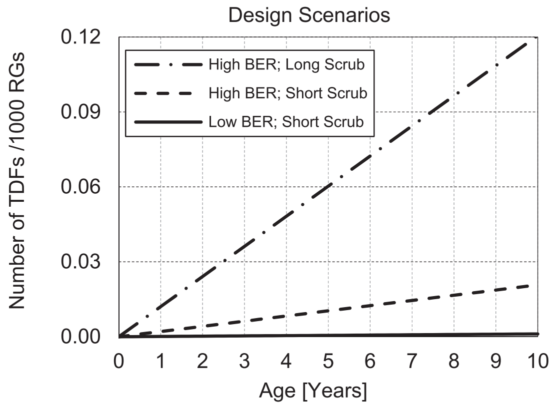

Fig. 11. Effects of latent defect rate, expressed by the bit-error rate (BER) and scrub time on the reliability of 1,000 RAID-6 groups (RG), expressed as triple-disk failures.

6.4. Sensitivity Analysis

There are several important what-if analyses that concern storage system designers as well as reliability engineers. They relate, among others, to the effects of scrubbing and latent sector error rates on the overall storage system reliability.

We illustrate these effects by three different sets of assumptions. In all three cases, the time-to-operational-failure distributions are the same and are representative of today's highly reliable HDDs (η = 876,000 and β = 1.0). Similarly, we assume the time-to-restore an operational failure is about 12 hours ($\eta = 12$, $\beta = 3.0$). The analysis differences are in the rate of latent sector errors and the time-to-restore those errors. We discuss the results next, referring to the data in Figure 11.

6.4.1. Effects of Latent Defects. The latent defect rates are often different from one HDD design, generation, or manufacturer to the next. An example of a recent major design change is the change from longitudinal to perpendicular recording. A system designer may not know the true rate or the effects of these changes until large quantities of HDDs are deployed in the field and the observed failure rates can be analyzed. With our equation, we can easily understand the effect of potentially high defect rates, or bit-error rates (BER), before the system is deployed and any potential data loss occurs.

To illustrate how system designers can use our equation, we chose a high error rate of $1.98\text{x}10^{-11}$ per byte accessed and a low error rate of $9.9\text{x}10^{-13}$ per byte accessed. For the first analysis, for example, we assume a long time to complete a scrub of about 348 hours. Based on the results from our equation, we expect to see about 0.12 TDFs (triple-disk failures—partial or whole) per 1,000 RAID groups of size 16 over a ten-year span. This is shown as the top curve in Figure 11.

6.4.2. Effects of Scrub Time. We can also see the effects of time-to-complete the scrub. Decreasing it from 348 hours to 12 hours reduces the expected number of TDFs by a factor of 6, from 0.12 to 0.02 over the ten-year span. This is shown in Figure 11 as the middle curve.

6.4.3. Combining BER and Scrub Time. Lastly, the bottom curve in Figure 11 shows the combined effects of lower defect rate with a shorter time-to-scrub. Over the ten-year span, we expect to see 0.0012 TDFs, a 1,000x reduction in the expected number of TDFs compared to the baseline.

6.4.4. Datacenter Scale. The previous analyses considered modest-sized deployments with 1,000 RAID groups for a total population of 16,000 HDDs. However, it is not uncommon for a single organization to hold exabytes of data. To express the likelihood of data loss at this scale, let's consider 100,000 RAID groups for a total of 1.6EB of raw capacity with 1TB disks. Using our RAID-6 equation, an organization possessing this amount of online storage would experience ten TDFs or data-loss events in 2.5 years and approximately 26 events in 5 years. In contrast, the MTTDL equation would predict less than 0.1 TDFs.

7. JAVASCRIPT IMPLEMENTATION

To make it easy for systems designers and practitioners alike to experiment with the closed-form RAID-6 equation, we implemented it in Javascript and made it available as a webpage at `http://raideqn.netapp.com`. It is, in particular, useful for users who are not reliability experts; it gives them the ability to readily explore the design space for their RAID solution and understand its performance and reliability trade-offs without having to worry about Weibull distribution and Gamma function approximations, plotting, and other details.

The calculator has inputs for system parameters and the parameters of the four input distributions for our equation. System parameters include the number of data disks per RAID group, system mission time, and the system size, expressed as the total number of RAID groups. There are inputs for the η, β, and, where applicable, γ parameters for each of the three Weibull distributions, T_{Rest}, T_{Scrub}, and T_{Op}. The T_{Ld} input parameters are expressed as bit-error rate (BER) and the workload, expressed as MB/s accessed per disk. The calculator also provides three sets of predefined values for the four input distributions (listed in Table III) corresponding to our three analyzed HDD models. The user can also plug in custom values.

As the parameters are entered, the calculator immediately updates the input distributions. The calculator also graphs the shape of the respective PDF and CDF to help users visualize the impact of varying the three Weibull distribution parameters. The calculator produces a plot comparing MTTDL and our closed-form equation results, akin to Figure 10.

8. CONCLUSIONS

This work presents a closed-form equation that accurately predicts the reliability of RAID-6 systems and accounts for the various advanced features built into modern systems, including the availability of spare drives and rate-limited continuous media scrubbing. It derives the input-variable distribution from an extensive collection of field data collected over many years from 1,000s of commercially-deployed systems.

To the best of our knowledge, this work is the first formulation of a closed-form RAID-6 reliability equation that is simple enough to be plugged into a spreadsheet and used by system practitioners. Our hope is it will replace the original MTTDL reliability formula whose time has passed. We also believe that, thanks to the Javascript implementation that we made available for everyone, it will be used by RAID architects to explore design trade-offs as well as system dependability experts who wish to accurately capture the reliability of a storage system used in practice.

Finally, we believe this study provides the most-comprehensive analysis of disk failure and repair rates for three distinct disk models based on data from 10,000s of storage systems. Previous studies have provided comprehensive failure data from similarly sized populations of 100,000s of HDDs, but did not include repair rates. We provide distributions for both RAID rebuild and media scrubbing processes observed in the field from a successful commercially-deployed storage system.

ACKNOWLEDGMENTS

The authors would like to thank Steven Kleiman for his support of this project.

REFERENCES

Ascher, H. 1983. Statistical methods in reliability: Discussion. *Technometrics 25*, 4.

Ascher, H. 1999. A set of numbers is not a dataset. *IEEE Trans. Reliab 48*, 2.

Ascher, H. 2010. Personal communication.

Bairavasundaram, L., Goodson, G., Pasupathy, S., and Schindler, J. 2007. An analysis of latent sector errors in disk drives. *ACM SIGMETRICS, Perform. Eval. Rev. 35*, 1, 289–300.

Bairavasundaram, L., Goodson, G., Schroeder, B., Arpaci-Dusseau, A., and Arpaci-Dusseau, R. 2008. An analysis of data corruption in the storage stack. In *Proceedings of the 6th USENIX Conference on File and Storage Technologies*.

Bartlett J., Bartlett, W., Carr, R., Garcia, D., Gray, J., Horst, R., Jardine, R., Lenoski, D., and McGuire, D. 1990. Fault tolerance in tandem computers. NetApp Tech. rep. 90.5.

Bazovsky, I. 1961. *Reliability Theory and Practice*. Prentice Hall.

Blaum, M., Brady, J., Bruck, J., and Menon, J. 1994. EVENODD: An optimal scheme for tolerating double disk failures in RAID architectures. In *Proceedings of the 21st International Symposium on Computer Architecture*.

Corbett, P., English, B., Goel, A., Grcanac, T., Kleiman, S., Leong, J., and Shankar, S. 2004. Row-diagonal parity for double disk failure correction. In *Proceedings of the 3rd USENIX Conference on File and Storage Technologies*.

Dholakia, A., Eleftheriou, E., Hu, X., Iliadis, I., Menon, J., and Rao, K. K. 2006. Analysis of a new intra-disk redundancy scheme for high-reliability RAID storage systems in the presence of unrecoverable errors. *ACM SIGMETRICS Perform. Eval. Rev. 34*, 1.

Elerath, J. 2009a. A simple equation for estimating reliability of an N+1 redundant array of independent disks. In *Proceedings of the 39th International Conference on Dependable Systems and Networks*.

Elerath, J. 2009b. Hard disk drives: The good, the bad and the ugly. *ACM Queue 52*, 6.

Elerath, J. and Pecht, M. 2009. A highly accurate method for assessing reliability of redundant arrays of inexpensive disks (RAID). *IEEE Trans. Comput. 58*, 3.

EMC. 2007. EMC CLARiion RAID 6 Technology: A detailed review. http://www.emc.com/collateral/hardware/white-papers/h2891-clariion-raid-6.pdf. (Last accessed July 2012.)

Gao, Y., Meister, D., and Binkmann, A. 2010. Reliability analysis of declustered-parity RAID 6 with disk scrubbing and considering irrecoverable read errors. In *Proceedings of the IEEE International Conference on Networking, Architecture, and Storage*.

Gibson, G. and Patterson, D. 1993. Designing disk arrays for high data reliability. *J. Parallel Distrib. Comput. 17*.

Greenan, K., Plank, J., and Wylie, J. 2010. Mean time to meaningless: MTTDL, Markov models, and storage system reliability. In *Proceedings of the 2nd USENIX Conference on Hot Topics in Storage and File Systems*.

Kececioglu, D. 1993. *Reliability & Life Testing Handbook, Volumes 1 & 2*. Prentice Hall.

Malhotra, M. and Trivedi, K. 1993. Reliability analysis of redundant arrays of inexpensive disks. *J. Parallel Distrib. Comput. 17*, 1–2.

Nelson, W. 1982. *Applied Life Data Analysis*. Addison-Wesley.

Nelson, W. 1990. *Accelerated Testing*. Wiley & Sons.

Nelson, W. 2003. *Recurrent Events Data Analysis for Product Repairs, Disease Recurrences, and Other Applications*. ASA-SIAM Series on Statistics and Applied Probability, Society for Industrial and Applied Mathematics.

NetApp. 2013. NetApp data ONTAP 8 operating system. http://www.netapp.com/us/products/platform-os/data-ontap-8/index.aspx.

Oracle. 2010. A better RAID strategy for high capacity drives in mainframe storage. http://www.oracle.com/technetwork/articles/systems-hardware-architecture/raid-strategy-hi-capacity-drives-170907.pdf.

Paris, J., Amer, A., Long, D., and Schwarz, T. 2009. Evaluating the impact of irrecoverable read errors on disk array reliability. In *Proceedings of the 15th IEEE Pacific Rim International Symposium on Dependable Computing*.

Patterson, D. A., Gibson, G., and Katz, R. 1988. A case for redundant arrays of inexpensive disks (RAID). In *Proceedings of the ACM SIGMOD International Conference on Management of Data*.

Pinheiro, E., Weber, W., and Barroso, L. 2007. Failure trends in a large disk drive population. In *Proceedings of the 5th USENIX Conference on File and Storage Technologies*.

Rao, K. K., Hafner, J., and Golding, R. 2006. Reliability for networked storage nodes. In *Proceedings of the 36th International Conference on Dependable Systems and Networks*.

Schroeder, B. and Gibson, G. 2007. 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*.

Serve The Home. 2011. The RAID reliability anthology – The primer. http://www.servethehome.com/raid-reliability-failureanthology-part-primer.

Shah, S. and Elerath, J. 2005. Reliability analysis of disk drive failure mechanisms. In *Proceedings of the IEEE Reliability and Maintainability Symposium*.

Thompson, W. 1981. On the foundations of reliability. *Technometrics 23*, 1.

Thomasian, A. and Blaum, M. 2009. Higher reliability in redundant disk arrays: Organization, operation, and coding. *ACM Trans. Storage 5*, 3.

Tobias, P. and Trindade, D. 2011. *Applied Reliability* 3rd Ed. CRC Press.

Received March 2013; revised June 2013; accepted July 2013