RZ 3847 (#Z1211-014) 06/30/2013
Computer Science 10 pages

Research Report

Effect of Latent Errors on the Reliability of Data Storage Systems

Vinodh Venkatesan and Ilias Iliadis

IBM Research – Zurich
8803 Rüschlikon
Switzerland

Email: {ven, ili}@zurich.ibm.com

Research
Africa • Almaden • Austin • Australia • Brazil • China • Haifa • India • Ireland • Tokyo • Watson • Zurich

Effect of Latent Errors on the Reliability of Data Storage Systems

Vinodh Venkatesan and Ilias Iliadis
IBM Research – Zurich
8803 Rüschlikon, Switzerland
{ven, ili}@zurich.ibm.com

***Abstract*—The reliability of data storage systems is adversely affected by the presence of latent sector errors. As the number of occurrences of such errors increases with the storage capacity, latent sector errors have become more prevalent in today's high capacity storage devices. Such errors are typically not detected until an attempt is made to read the affected sectors. When a latent sector error is detected, the redundant data corresponding to the affected sector is used to recover its data. However, if no such redundant data is available, then the data of the affected sector is irrecoverably lost from the storage system. Therefore, the reliability of data storage systems is affected by both the complete failure of storage nodes and the latent sector errors within them. In this article, closed-form expressions for the mean time to data loss (MTTDL) of erasure coded storage systems in the presence of latent errors are derived. The effect of latent errors on systems with various types of redundancy, data placement, and sector error probabilities is studied. For small latent sector error probabilities, it is shown that the MTTDL is reduced by a factor that is independent of the number of parities in the data redundancy scheme as well as the number of nodes in the system. However, for large latent sector error probabilities, the MTTDL is similar to that of a system using a data redundancy scheme with one parity less. The reduction of the MTTDL in the latter case is more pronounced than in the former one.**

I. Introduction

The size of data storage systems grows in terms of the number of storage nodes in the system as well as the storage capacities of individual storage nodes. This growth has two effects on the reliability of the storage system. Firstly, a higher number of storage nodes implies that more nodes are expected to fail each day. Although modern data storage systems use various forms of data redundancy to protect data from node failures, higher frequency of node failures poses a greater risk of data loss in large systems. Secondly, higher capacity nodes are more prone to latent sector errors. This is because, given a certain bit error rate, the number of occurrences of such errors increases with the node capacity. Such latent sector errors increase the risk of irrecoverable loss of data, especially when the data is in a critical state wherein it has lost all of its redundancy as a result of successive storage node failures.

Storage systems employ erasure codes to protect data from node failures. Examples of erasure codes range from simple replication and RAID [1] to more advanced Reed-Solomon [2] and regenerating codes [3]. When nodes fail, the data redundancy is maintained through node rebuild processes that use the data from the surviving nodes to reconstruct the lost data in new replacement nodes. However, as these rebuild process take time to complete, there is a probability of further node failures occurring during rebuild that may eventually result in irrecoverable data loss. The probability of irrecoverable data loss during rebuild is further increased by the presence of latent sector errors in the surviving storage nodes. Such errors are typically not detected until an attempt is made to read the affected sectors. This attempt may be a result of a user read request, a background scrubbing process, or a node rebuild process [4], [5]. Usually, when a latent sector error is detected, the redundant data corresponding to the affected sector is used to rebuild it. However, if no redundant data corresponding to the affected sector is available, then the data of the affected sector is irrecoverably lost from the storage system. This may happen when all the redundant data is lost either by node failures or other latent errors in the system.

The average amount of time taken by the system to lose some data irrecoverably, also known as the mean time to data loss, or MTTDL, is a measure of reliability commonly used to compare different coding schemes and study the effect of various design parameters [6]. In the absence of latent sector errors, the effect of various system designs and parameters, such as, mean time to failure of a node, node rebuild bandwidth, node capacity, data placement scheme, and erasure-code used, on the MTTDL of the system have been extensively studied in literature [1], [7], [8], [9], [10], [11], [12], [13], [14], [15]. A comparison between erasure codes and replication in terms of availability in peer-to-peer systems has been presented in [16]. It has been well-established that erasure codes can provide much higher reliability than replication for the same level of storage efficiency. The trade-off, however, is in the performance as erasure codes may require Galois field arithmetic for encoding and decoding. Therefore, many recent works have laid emphasis on the development of new codes as well as new encoding and decoding techniques to improve the performance of erasure coded systems (see [17] and references therein). Some works have also addressed the reliability assessment of erasure codes through simulation [18].

The occurrence of latent sector errors in disk drives and its effect on the MTTDL of RAID systems was also investigated in the literature [4], [5], [19], [20], [21]. In this paper, closed-form expressions for MTTDL in the presence of latent errors are derived for a wide variety of erasure-coded systems, including replication-based systems and RAID systems, as well as systems with different redundancy placement schemes

Table I
PARAMETERS OF A STORAGE SYSTEM

c	amount of data stored on each storage node (bytes)
n	number of storage nodes
$c\mu$	average read-write rebuild bandwidth of a storage node (bytes/s)
$1/\lambda$	mean time to failure of a storage node (s)
$1/\mu$	mean time to read/write c amount of data from/to a node (s)
s	size of a sector (bytes)
p_S	probability of a sector having a latent error
B	average burst size for sector errors

for realistic node failure and rebuild time distributions. For small latent sector error probabilities, it is shown that the MTTDL is reduced by a factor that is independent of the number of parities in the data redundancy scheme as well as the number of nodes in the system. However, for large latent sector error probabilities, the MTTDL is similar to that of a system using a data redundancy scheme with one parity less. The reduction of the MTTDL in the latter case is more pronounced than in the former one. This is the first work to study the effect of latent sector errors on MTTDL for different *codeword placement schemes.*

The remainder of this article is organized as follows: Section II describes the storage system model. Section III describes the methodology of reliability analysis used. Section IV evaluates the reliability of storage systems in the presence of latent errors for various data placement schemes and erasure codes. Section V discusses the effect of latent errors on system reliability with numerical results. Finally, the paper is concluded in Section VI.

II. SYSTEM MODEL

The storage system is modeled as a collection of n identical *storage nodes* each of which stores c amount of data. In addition to the space required for the c amount of data that is stored, each node is assumed to have sufficient spare space that may be used for a distributed rebuild process (see Section II-F) when other nodes fail. The main parameters used in the storage system model are listed in Table I.

A. Storage Node

A storage node comprises of one or more disks, memory, processor, network interface, and power supply. Any of these components can fail and lead to a temporary node unavailability or a permanent node failure.

1) Node Unavailability vs. Node Failure: The difference between temporary node unavailability and failure (or permanent unavailability) is important to the reliability model. Nodes that become temporarily unavailable may only result in temporary data unavailability, whereas node failures may cause irrecoverable data loss. The focus of this paper will be on the study of irrecoverable data loss.

2) Independence of Node Failures: It is known that strong correlations exist among node unavailabilities [22]. These correlations may be due to short power outages in the datacenter, or part of a rolling reboot or upgrade activity at the datacenter management layer [22]. However, only less than 10% of the node unavailabilities last longer than 15 minutes and are treated as node failures which trigger a rebuild process. There is no indication that correlations exist among such node failures. It has been argued that disk (as opposed to node) replacement rates in large storage systems show correlations [23]. However, as disks have been observed to be more reliable than other components of a node [24], the failure of a node is mainly determined by the failure of these other components. As there is no evidence that correlations exist among node failures (or permanent unavailabilities), we assume node failures to be independent in our model.

B. Redundancy

To protect data from node failures and sector errors, the user data is divided into blocks (or symbols) of a fixed size and each set of l blocks is encoded into a set of $m > l$ blocks, called a codeword, before storing them on m distinct nodes. In this paper, we consider (l, m)-MDS codes (or maximum distance separable codes), in which the encoding is done such that any subset of l symbols of a codeword can be used to decode the l symbols of user data corresponding to that codeword. In other words, each codeword can sustain the loss of up to any $m - l$ symbols. Replication-based systems with a replication factor r, are a subset of such erasure coded systems where the parameters l and m are equal to 1 and r, respectively. For notational convenience, we define $\tilde{r} = m - l + 1$.

C. Codeword Placement

In a large storage system, the number of nodes, n, is typically much larger than the codeword length, m. Therefore, there exist many ways in which a codeword of m blocks can be stored across n nodes. In this paper, we consider the class of symmetric placement schemes as described in [14]. Two schemes of particular interest within this class are the clustered and declustered placement schemes. These two schemes are first explained below before describing the broader class of symmetric placement schemes.

1) Clustered Placement: If n is divisible by m, one way to place codewords would be to divide the n nodes into disjoint sets, of m nodes each, and store the codewords across the nodes in each set. This type of data placement is known as *clustered* placement, and each of these disjoint sets of nodes as *clusters*. In such a placement scheme, it can be seen that no cluster stores the redundancies corresponding to the data on another cluster. The storage system can essentially be modeled as consisting of n/m independent clusters. Reliability behavior of a RAID cluster in the presence of latent errors is studied in [21].

2) Declustered Placement: A placement scheme that can potentially offer far higher reliability than the clustered placement scheme, especially as the number of nodes in the system grows, is the *declustered* placement scheme. There exists $\binom{n}{m}$ different ways of placing m symbols of a codeword across n nodes. In this scheme, all these $\binom{n}{m}$ possible ways are equally used to store data. It can be seen that, in such a placement scheme, when a node fails, the redundancy corresponding to the data on the failed node is equally spread across the

remaining surviving nodes. This allows one to use the rebuild read-write bandwidth available at all surviving nodes to do a *distributed* rebuild in parallel, which can be extremely fast when the number of nodes is large. In contrast, in clustered placement scheme, when a node fails, the redundancy corresponding to the data on the failed node is only spread across the remaining nodes of a cluster. Therefore, a fast parallel rebuild process that scales with the number of nodes is not possible for clustered placement.

3) Spread Factor: A broader set of symmetric placement schemes can be defined using the concept of *spread factor* [14]. The analysis presented for clustered and declustered placement schemes can be easily extended to this set of symmetric placement schemes.

D. Node Failure

The times to node failures are modeled as independent and identically distributed random variables. Denote the cumulative distribution function of the times to node failure by F_λ, with mean, $1/\lambda$. It has been shown that the system MTTDL is invariant within a large class of failure time distributions that includes the exponential distribution and, most importantly, real-world distributions like Weibull and gamma [15].

E. Latent Errors

Latent sectors errors in disk drives have been studied extensively in literature [19], [20]. The number of latent sector errors are observed to be higher for disks with higher capacity. In a given disk, the probability of a sector having an error increases with its age [20]. However, in a system with a large number of disks, the age distribution of the disks becomes stationary as disks fail and are replaced with new disks. Consequently, the steady state probability, p_S, of a sector having a latent error remains constant. Typical values of p_S lie in the range of 10^{-9} to 10^{-8} [20]. The analysis presented in this paper holds for

$$p_S \ll \lambda/\mu, \tag{1}$$

which holds true for practical values of λ/μ (in the order of 10^-5 to 10^-3) and practical values of p_S. This assumption ensures that the probability of two different nodes having sector errors in the same position is negligible. Latent sector errors in each disk are found to exhibit high spatial and temporal locality [19], [20]. In other words, sector errors are found to occurs in bursts with respect to both their locations within a disk as well as their times of occurrence. To counter these errors, various scrubbing and intradisk redundancy schemes have been developed [19], [25]. In effect, these schemes reduce p_S, and therefore improve reliability. Clearly, the extent to which p_S is reduced depends on the factors mentioned above and also on the parameter settings of these schemes [19], [25, Propositions 7.9 and 7.10]. Thus the effect of these factors is accounted and inherently captured by p_S. In Section A, it will be shown that the probability of data loss depends on the probability, $P_S(D)$, of encountering at least one sector error while rebuilding D amount of data. If sector errors were independent, $P_S(D)$, is given by $(D/s)p_S$, where s denotes the size of a sector in bytes. However, as sector errors occur in bursts, it can be shown that (see Section 6 in [21])

$$P_S(D) \approx (D/(sB))p_S, \tag{2}$$

where B is the average burst length. The above result can be intuitively explained as follows. If the errors were independent, then the errors will be spread uniformly across the entire disk. However, if the errors are bursty, the same number of errors will be clumped together into a fewer number of error bursts with larger gaps separating these bursts. Therefore, the probability that a sector error exists in a space of size D is reduced when the errors are bursty. Since 90-98% of cases consist of bursts of just one sector error, and less than 2.5% consist of bursts of more than two errors [19], the typical value of B is close to one. Clearly, for high values of p_S or large amounts of data D, for which the right hand side of (2) is greater than one, the approximation does not hold. Therefore, we approximate the variation of $P_S(D)$ as a linear function in D that saturates at one,

$$P_S(D) \approx \min\left(1, (D/(sB))p_S\right). \tag{3}$$

F. Node Rebuild

When storage nodes fail, codewords lose some of their symbols and this leads to a reduction in data redundancy. The system attempts to maintain the redundancy of the system by reconstructing the lost codeword symbols using the surviving symbols of the affected codewords from other nodes.

1) Codeword Reconstruction: For a system using an (l, m)-MDS code for redundancy, a simple way to reconstruct a codeword that has lost up to $m - l$ symbols is to read any of its l symbols, decode the original l user data blocks, re-encode these l user data blocks using the (l, m)-MDS code, and recover the lost codeword symbols. The time taken by the reconstruction process depends on the amount of data to be read and written and the number of nodes involved.

2) Sector Errors During Rebuild: Latent sector errors in surviving nodes may be detected during rebuilds. When the codewords corresponding to the detected bad sectors have at least l other surviving symbols, these bad sectors are restored by the codeword reconstruction process described above. As the number of such detected errors is relatively small, the time taken to restore these sectors is small and therefore has negligible effect on the system reliability. However, if the codewords corresponding to the detected bad sectors have less than l other surviving symbols, these sectors cannot be restored and the data corresponding to the codewords is irrecoverably lost from the system. So only the bad sectors in critical data, that is, data whose codewords have only l surviving symbols, is considered for the reliability analysis of this paper. Sector errors in critical data are referred to as critical sector errors, or CSE.

3) Intelligent Rebuild: In an intelligent rebuild process, the system attempts to first recover the codewords of the user data that have the least number of codeword symbols left.

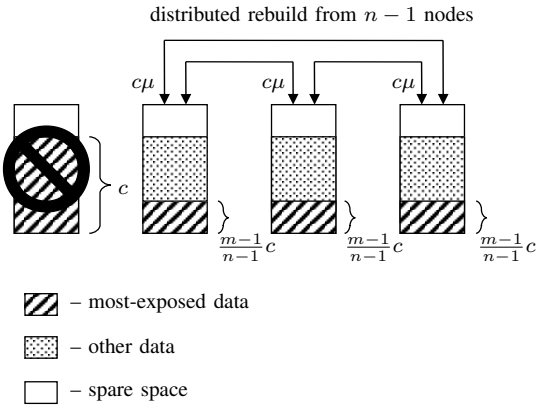


Figure 1. Distributed rebuild in declustered placement.

In contrast to intelligent rebuild, one may consider a *blind* rebuild, where lost codeword symbols are being recovered in an order that is not specifically aimed at recovering the codewords with the least number of surviving symbols first. Clearly, such a blind rebuild is more vulnerable to data loss. So, in the remainder of the paper, we only consider intelligent rebuild.

4) Distributed Rebuild: For placement schemes with spread factor $k > m$, the surviving codeword symbols that the system needs to read to recover the lost codeword symbols may be spread across $\tilde{k} \leq k-1$ surviving nodes. For such schemes, as illustrated in Figure 1, a distributed rebuild process may be used, which involves reading the required codeword symbols of the data to be rebuilt from all the $\tilde{k}$ nodes, computing the lost codeword symbols, and writing them to the spare space of these $\tilde{k}$ nodes in such a way that no symbol is written to a node in which another symbol corresponding to the same codeword is already present. Once all lost codeword symbols are recovered, they are transferred to a new replacement node. Due to the parallel nature of distributed rebuild, the rebuild times can be extremely short for large storage systems. In this paper, distributed rebuild is assumed to be used for all placement schemes with spread factors $k > m$. However, for $k = m$, that is, for clustered placement, such a rebuild cannot be used, as it is not possible to write the reconstructed symbols to the spare space of the $\tilde{k} \leq k-1 = m-1$ nodes in such a way that no symbol is written to a node in which another symbol corresponding to the same codeword is already present. Therefore, for clustered placement, it is assumed that the required codeword symbols are read from a set of l nodes of the surviving cluster, the lost symbols are reconstructed on the fly, and the reconstructed symbols are directly written to a new replacement node.

5) Node Rebuild Bandwidth: During the rebuild process, an average read-write bandwidth of $c\mu$ bytes/s is assumed to be reserved at each node exclusively for the rebuild. This implies that the average time required to read (or write) c amount of data from (or to) a node is equal to $1/\mu$. The average rebuild bandwidth is usually only a fraction of the total bandwidth available at each node, with the remainder being used to serve user requests. Denote the cumulative distribution function of the time required to read (or write) c amount of data from (or to) a node by G_μ, and its corresponding probability density function by g_μ.

G. Failure and Rebuild Time Distributions

It is known that real-world storage nodes are *generally reliable*, that is, the mean time to repair a node (which is typically of the order of tens of hours) is much smaller than the mean time to failure of a node (which is typically at least of the order of thousands of hours). As $1/\lambda$ denotes the mean time to failure of a node and $1/\mu$ denotes the mean time to read (or write) c amount of data from (or to) a storage node, it follows that generally reliable nodes satisfy the following condition:

$$1/\mu \ll 1/\lambda, \quad \text{or} \quad \lambda/\mu \ll 1. \tag{4}$$

In the subsequent analysis, this condition implies that terms involving powers of λ/μ greater than one are negligible compared to λ/μ and can be ignored.

Let the cumulative distribution functions F_λ and G_μ satisfy the following condition:

$$\mu \int_0^\infty F_\lambda(t)(1-G_\mu(t))dt \ll 1, \quad \text{with } \frac{\lambda}{\mu} \ll 1. \tag{5}$$

The results of this paper are derived for the class of failure and rebuild distributions that satisfy the above condition. In particular, the mean time to data loss of a system is shown to be insensitive to the failure distributions within this class. This result is of great importance because it turns out that this condition holds for a wide variety of failure and rebuild distributions, including, most importantly, distributions that are seen in real-world storage systems [15].

III. Reliability Analysis

The reliability analysis in this article uses a methodology similar to [13], [15]. It involves a series of approximations, each of which are justified for generally reliable nodes satisfying (4) and for failure and rebuild time distributions satisfying (5). Note that this methodology does not necessarily assume exponential failure and rebuild distributions and therefore does not involve any Markov chain analysis. The theoretical estimates of mean times to data loss predicted using this methodology have also been shown to match with simulations, which avoid all the approximations made in the methodology, over a wide range of system parameters [13], [14], [15]. This establishes a confidence in the results obtained and conclusions drawn in this article.

A. Mean Time to Data Loss (MTTDL)

In an erasure coded system, a data loss is said to have occurred when sufficient number of blocks of at least one codeword have been lost, rendering the codeword(s) undecodeable. The average time taken for the system to end up in data loss, also referred to as the mean time to data loss, or MTTDL,

is a commonly used measure that is useful for assessing trade-offs, for comparing schemes, and for estimating the effect of the various parameters on the system reliability [6].

At any point of time, the system can be thought to be in one of two modes: *fully-operational* mode or *rebuild* mode. During the fully-operational mode, all data in the system has the original amount of redundancy and there is no active rebuild process. During the rebuild mode, some data in the system has less than the original amount of redundancy and there is an active rebuild process that is trying to restore the lost redundancy. A transition from fully-operational mode to rebuild mode occurs when a node fails; we refer to this node failure that causes a transition from the fully-operational mode to the rebuild mode as a *first-node* failure. Following a first-node failure, a complex sequence of rebuilds and subsequent node failures may occur, which eventually lead the system either to irrecoverable data loss, with probability P_{DL}, or back to the original fully-operational mode by restoring all replicas, with probability $1-P_{DL}$. The following proposition has been proved in [15].

Proposition 1: Consider a system with generally reliable nodes whose failure and rebuild distributions, F_λ and G_μ, satisfy (5). Its MTTDL is given by MTTDL $\approx 1/(n\lambda P_{DL})$. The relative error in the approximation tends to zero as λ/μ tends to zero.

B. Probability of Data Loss in Rebuild Mode (P_{DL})

This section show how P_{DL} is estimated so that MTTDL can be obtained using Proposition 1.

1) Exposure Levels: Consider an erasure coded storage system with an (l,m)-MDS code. We model the system as evolving from one exposure level to another as nodes fail and rebuilds complete. At time $t \geq 0$, let $D_j(t)$ be the amount of user data that have lost j symbols of their corresponding codewords, for $0 \leq j \leq \tilde{r}$ (note that $\tilde{r} = m-l+1$). At time t, the system is said to be in exposure level e, $0 \leq e \leq \tilde{r}$, if $e = \max_{D_j(t)>0} j$.

2) Direct Path Approximation: Denote the probability of the direct path to data loss by $P_{DL,\text{direct}}$, that is,

$$P_{DL,\text{direct}} := \Pr\{\text{exposure level path } 1 \to 2 \to \cdots \to \tilde{r}\}. \quad (6)$$

Now the following approximation holds for generally reliable nodes satisfying (5) [13].

$$P_{DL} \approx P_{DL,\text{direct}}. \quad (7)$$

The relative error in the approximation tends to zero as λ/μ tends to zero.

C. Probability of the Direct Path to Data Loss ($P_{DL,direct}$)

Consider the direct path to data loss, that is, the path $1 \to 2 \to \cdots \to \tilde{r}$ through the exposure levels. At each exposure level, the *intelligent* rebuild process attempts to rebuild the most-exposed data, that is, the data with the least number of codeword symbols left (see Section II-F). Let the rebuild times of the most-exposed data at each exposure level in this path be denoted by R_e, $e = 1, \cdots, \tilde{r}-1$. Let t_e, $e = 2, \cdots, \tilde{r}$, be the times of transitions from exposure level $e-1$ to e following a first-node failure. Let $\tilde{n}_e$ be the number of nodes in exposure level e whose failure before the rebuild of most-exposed data causes an exposure level transition to level $e+1$. Denote the time period from t_e until the next failure of node i by $E_{t_e}^{(i)}$. The time, F_e, until the first failure among the $\tilde{n}_{e-1}$ nodes that causes the system to enter exposure level e from $e-1$, is

$$F_e := \min_{i \in \{1,\cdots,\tilde{n}_{e-1}\}} E_{t_{e-1}}^{(i)}, \quad e = 2, \cdots, \tilde{r}. \quad (8)$$

At exposure level e, let α_e be the fraction of the rebuild time R_e still left when a node failure occurs causing an exposure level transition, that is, let

$$\alpha_e := (R_e - F_{e+1})/R_e, \quad e = 1, \cdots, \tilde{r}-2. \quad (9)$$

It can be shown that α_e is uniformly distributed in $(0,1)$ [26, Lemma 2]. Now, denote by $1/\mu_e$ the following conditional means of R_e:

$$1/\mu_e := E[R_e | R_{e-1}, \alpha_{e-1}], \quad e = 2, \cdots, r-1. \quad (10)$$

The actual values of $1/\mu_e$ depend on the codeword placement and this will be further discussed in later sections of this paper. Now, the distribution of R_e given R_{e-1} and α_{e-1} could be modeled in several ways. We consider the model B presented in [15], namely,

$$R_e | R_{e-1}, \alpha_{e-1} = 1/\mu_e \qquad w.p.\,1 \text{ for } e = 2, \cdots, \tilde{r}-1. \quad (11)$$

This model assumes that the rebuild time R_e is determined completely by R_{e-1} and α_{e-1} and no new randomness is introduced in the rebuild time of exposure level e. For further discussion on this model see [15]. Now, in the critical exposure level $\tilde{r}-1$, let $S_{\tilde{r}-1}$ denote the speed of rebuild and $E[D_{\tilde{r}-1}]$ denote the expected amount of data to be rebuilt. Define

$$\text{Region A:} \qquad p_S \leq sB/\left(lE[D_{\tilde{r}-1}]\right), \quad (12)$$

$$\text{Region B:} \qquad sB/\left(lE[D_{\tilde{r}-1}]\right) < p_S \overset{(1)}{\ll} \lambda/\mu. \quad (13)$$

These two regions represent the two main ways in which the presence of sector errors affects the reliability of the system. This is because, in the critical exposure level, to rebuild $E[D_{\tilde{r}-1}]$amount of data, $lE[D_{\tilde{r}-1}]$ amount of data has be read from the surviving nodes. The probability of a critical sector error in this $lE[D_{\tilde{r}-1}]$ amount of data is $p_S lE[D_{\tilde{r}-1}]/(sB)$. In region B, the probability of a critical sector error is essentially one. This means that as soon as the system enters the critical state (i.e. exposure level $\tilde{r}-1$), there is a sector error in the critical data with probability almost one, and the system experiences irrecoverable data loss. In region A, the probability of a critical sector error is less than one and its influence on the system reliability depends on the relative magnitudes of the probability of a critical sector error and the probability of a node failure. As $E[D_{\tilde{r}-1}]$ depends on the underlying codeword placement scheme, the regions A and B also depend on the placement scheme. Now, the probability of the direct path to data loss is given by the following proposition.

Proposition 2: Consider an (l, m)-MDS erasure coded storage system with generally reliable nodes whose failure and rebuild distributions, F_λ and G_μ, satisfy (5). Let the probability of a sector being in error be p_S ($p_S \ll \lambda/\mu$). Consider the direct path $1 \to 2 \to \cdots \to \tilde{r}$ through the exposure levels in which the rebuild times R_e satisfy (11). The probability of this direct path is given by

$$P_{DL,\text{direct}} \approx \begin{cases} \left(1 + \frac{p_S l S_{\tilde{r}-1}}{sB\lambda\tilde{n}_{\tilde{r}-1}}\right) P^{\text{noSE}}_{DL,\text{direct}}(\tilde{r}) & \text{in region A,} \\ P^{\text{noSE}}_{DL,\text{direct}}(\tilde{r}-1) & \text{in region B,} \end{cases} \quad (14)$$

where $P^{\text{noSE}}_{DL,\text{direct}}(x)$ denotes the probability of the direct path to data loss for a system with zero sector error probability and maximum exposure level $\tilde{r} = x$. Regions A and B are as defined in (12) and (13). The relative error in the approximation in (14) tends to zero as λ/μ tends to zero.

Proof: See Appendix A. ∎

The expressions for $P^{\text{noSE}}_{DL,\text{direct}}(\tilde{r})$ are given by [26, Prop. 3].

IV. Effect of Latent Errors on Reliability

Let $\text{MTTDL}^{\text{noSE}}(\tilde{r})$ denote the MTTDL in the absence of sector errors. Then, from (7) and Propositions 1 and 2,

$$\text{MTTDL}(\tilde{r}) = \begin{cases} \frac{\text{MTTDL}^{\text{noSE}}(\tilde{r})}{\left(1+\frac{p_S l S_{\tilde{r}-1}}{sB\lambda\tilde{n}_{\tilde{r}-1}}\right)} & \text{in region A,} \\ \text{MTTDL}^{\text{noSE}}(\tilde{r}-1) & \text{in region B.} \end{cases} \quad (15)$$

As expected, the presence of latent sector errors reduces the MTTDL. In region A, the MTTDL is reduced by a factor that increases with p_S. However, in region B, the MTTDL is equal to the MTTDL of a redundancy scheme with one parity less. The extents of the two regions in terms of p_S depend on the expected amount of data to be rebuilt in the critical exposure level, which in turn depends on the underlying codeword placement scheme.

A. Clustered Codeword Placement

Let $D^{\text{clus.}}_1, \cdots, D^{\text{clus.}}_{\tilde{r}-1}$ denote the amounts of data to be rebuilt in exposure levels $1, \cdots, \tilde{r}-1$, respectively. Following the first node failure, the amount of data to be rebuilt is $D^{\text{clus.}}_1 = c$. The fraction of data not rebuilt in exposure level one when a transition to exposure level two occurs is α_1 and it is uniformly distributed in $(0, 1)$. Therefore, the expected amount of data to be rebuilt in exposure level two is $E[D^{\text{clus.}}_2] = D^{\text{clus.}}_1/2 = c/2$. Continuing by the same logic, we obtain $E[D^{\text{clus.}}_{\tilde{r}-1}] = c/2^{\tilde{r}-2}$. Therefore, by the definitions (12) and (13),

$$\text{Region A (clus.):} \quad p_S \le sB2^{\tilde{r}-2}/(lc), \quad (16)$$

$$\text{Region B (clus.):} \quad sB2^{\tilde{r}-2}/(lc) < p_S \overset{(1)}{\ll} \lambda/\mu. \quad (17)$$

For a system using clustered codeword placement, the rebuild process involves reading data from l nodes of the affected cluster at an average bandwidth of $c\mu$ from each node, computing the lost codeword symbols, and writing them to a spare node at an average bandwidth of $c\mu$. Therefore, the average speed of rebuild in exposure level $\tilde{r}-1$ is $S^{\text{clus.}}_{\tilde{r}-1} = c\mu$. The failure of any of the surviving l nodes of the critical cluster before rebuild completion causes data loss. Therefore, $\tilde{n}^{\text{clus.}}_{\tilde{r}-1} = l$. Substituting these values of $S^{\text{clus.}}_{\tilde{r}-1}$ and $\tilde{n}^{\text{clus.}}_{\tilde{r}-1}$ in (15),

$$\text{MTTDL}^{\text{clus.}}(\tilde{r}) = \begin{cases} \frac{\text{MTTDL}^{\text{clus., noSE}}(\tilde{r})}{\left(1+\frac{p_S c\mu}{sB\lambda}\right)} & \text{(region A)} \\ \text{MTTDL}^{\text{clus., noSE}}(\tilde{r}-1) & \text{(region B)} \end{cases} \quad (18)$$

The expression for $\text{MTTDL}^{\text{clus., noSE}}(\tilde{r})$ is given by [26, Prop. 4]. Furthermore, condition (37) can be seen to hold true for clustered placement as follows:

$$\tilde{n}^{\text{clus.}}_{\tilde{r}-1}\lambda D^{\text{clus.}}_{\tilde{r}-1}/S^{\text{clus.}}_{\tilde{r}-1} = l\lambda D^{\text{clus.}}_{\tilde{r}-1}/(c\mu) \le l\lambda c/(c\mu) \ll 1. \quad (19)$$

Here, the first inequality follows from noting that the amount of most critical data to be rebuilt in the direct path never exceeds the total amount of data stored on one node (that is, c), and the final inequality follows from the assumption (4).

B. Declustered Codeword Placement

Let $D^{\text{declus.}}_1, \cdots, D^{\text{declus.}}_{\tilde{r}-1}$ denote the amounts of data to be rebuilt in exposure levels $1, \cdots, \tilde{r}-1$, respectively. Following the first node failure, the amount of data to be rebuilt is $D^{\text{declus.}}_1 = c$. The fraction of data not rebuilt in exposure level one when a transition to exposure level two occurs is α_1 and it is uniformly distributed in $(0, 1)$. In contrast to clustered placement scheme, not all of the unrebuilt part of this $D^{\text{declus.}}_1$ amount of data loses its second codeword symbol. Due to the nature of the declustered placement scheme, the two failed nodes store codewords of only a fraction $(m-1)/(n-1)$ of this data. The intelligent rebuild process only rebuilds this fraction of the data in the second exposure level. Therefore, the expected amount of data to be rebuilt in exposure level two is $E[D^{\text{declus.}}_2] = \frac{1}{2}\frac{m-1}{n-1}D^{\text{declus.}}_1 = \frac{1}{2}\frac{m-1}{n-1}c$. Continuing by the same logic, we obtain $E[D^{\text{declus.}}_{\tilde{r}-1}] = \frac{1}{2^{\tilde{r}-2}}c\prod_{e=1}^{\tilde{r}-2}\left(\frac{m-e}{n-e}\right)$. Therefore, by the definitions (12) and (13), the regions A and B for clustered placement are given by

$$\text{Region A (declus.):} \ p_S \le \frac{sB2^{\tilde{r}-2}}{lc}\prod_{e=1}^{\tilde{r}-2}\left(\frac{n-e}{m-e}\right), \quad (20)$$

$$\text{Region B (declus.):} \ \frac{sB2^{\tilde{r}-2}}{lc}\prod_{e=1}^{\tilde{r}-2}\left(\frac{n-e}{m-e}\right) < p_S \overset{(1)}{\ll} \frac{\lambda}{\mu}. \quad (21)$$

For a system using declustered codeword placement, the distributed rebuild process in exposure level $\tilde{r}-1$ involves reading the required codeword symbols of the data to be rebuilt from all the $n-\tilde{r}+1$ surviving nodes of the system, computing the lost codeword symbols, and writing them to the spare space of these nodes. This process requires reading lc amount of data, as well as writing c amount of data, from and to all $n-\tilde{r}+1$ surviving nodes in parallel. As each of the $n-\tilde{r}+1$ nodes has an average read-write rebuild bandwidth of $c\mu$, and as l times more data is read from each node than what is written during the distributed rebuild process, the average rate of rebuild in exposure level e is $S^{\text{declus.}}_{\tilde{r}-1} = (n-\tilde{r}+1)c\mu/(l+1)$. The failure of any of the $n-\tilde{r}+1$ surviving nodes during rebuild causes data loss. Therefore, $\tilde{n}^{\text{declus.}}_{\tilde{r}-1} = n-\tilde{r}+1$.

Table II
RANGE OF VALUES OF DIFFERENT PARAMETERS

Parameter	Meaning	Range
c	amount of data stored on each node	10 TB
n	number of storage nodes	10 to 1000
$1/\lambda$	mean time to failure of a storage node	30000 h
$1/\mu$	mean time to read/write c amount of data from/to a node	30 h
s	size of a sector	512 B
p_S	probability of a sector having an error	10^{-15} to 10^{-6}
B	average burst size for sector errors	1.05

Substituting these values of $S_{\tilde{r}-1}^{\text{declus.}}$ and $\tilde{n}_{\tilde{r}-1}^{\text{declus.}}$ in (15),

$$\text{MTTDL}^{\text{declus.}}(\tilde{r})=\begin{cases}\frac{\text{MTTDL}^{\text{declus., noSE}}(\tilde{r})}{\left(1+\frac{p_S c\mu l}{sB\lambda(l+1)}\right)} & \text{(region A)}\\ \text{MTTDL}^{\text{declus., noSE}}(\tilde{r}-1) & \text{(region B)}\end{cases} \quad (22)$$

The expression for $\text{MTTDL}^{\text{declus., noSE}}(\tilde{r})$ is given by [26, Prop. 5]. Moreover, condition (37) can be seen to hold true for declustered placement, similar to (19).

C. Other Symmetric Placement Schemes

The regions A and B and the corresponding MTTDL expressions can be derived for all other symmetric placement schemes in a similar manner as above.

Remark 1: When the probability of encountering a latent error in the critical state is less than one (i.e. in region A), the MTTDL is reduced by a factor $\left(1+\frac{p_S c\mu l}{sB\lambda(l+1)}\right)$ due to the presence of latent sector errors. This factor is the same for all symmetric data placement schemes except clustered placement. In addition, it is independent of the number of nodes in the system as well as the number of parities in the erasure code. For clustered placement, the MTTDL is scaled down by a factor $\left(1+\frac{p_S c\mu}{sB\lambda}\right)$ in region A due to the presence of latent errors. This factor is independent of both parameters, l and m, of the erasure code.

Remark 2: When the probability of encountering a latent error in the critical state is almost one (i.e. in region B), the MTTDL of the system is equal to the MTTDL of a system with one less parity and no sector errrors.

V. NUMERICAL RESULTS

In this section, we present the effect of latent errors on reliability by plotting the MTTDL of clustered and declustered placement schemes for a range of values of the system parameters. According to [19], [20], [22], the range of practical relevance of parameters used in the numerical results are listed in Table II.

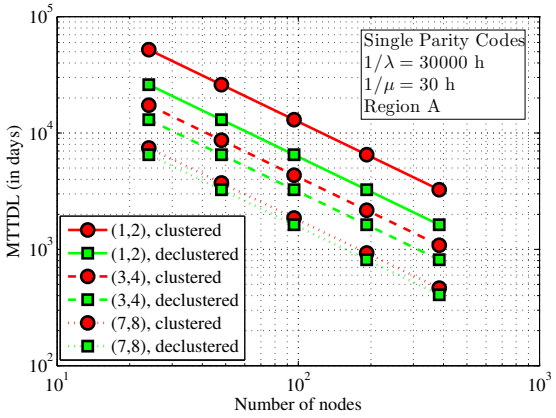
Figures 2, 3, and 4 show the difference in the MTTDL behavior with respect to the number of nodes in the system between regions A and B. Figures 2(a), 3(a), and 4(a) show the MTTDL behavior when the effect of sector errors is negligible in region A. However, when the probability of sector errors increases and reaches a level at which, when the system enters a critical state, it always experiences data loss due to a critical sector error, then the MTTDL behavior of the system resembles that of a system with one parity less. This can be observed by comparing Figure 3(b) with Figure 2(a), and by comparing Figure 4(b) with Figure 3(a).

Figures 5(a) and 5(b) show the MTTDL behavior with respect to p_S for a fixed number of nodes. For $1/\lambda = 30000$ h and $1/\mu = 30$ h, the practical values of p_S in the range of 10^-9 to 10^-8 (without any scrubbing or intradisk redundancy) lie mostly in region B. For double parity codes, this means that declustered and clustered placement have the same reliability for practical values of p_S. For triple parity codes, the difference in reliability between declustered and clustered placement is reduced. By reducing p_S by two to three orders of magnitude and using declustered placement, significant improvements in reliability can be achieved. Note also that these curves depend on the ratio λ/μ. All curves move to the right when λ/μ increases. Therefore, when node failures become more frequent, it is more likely that declustered placement performs much better in terms of reliability than clustered placement for systems with two or more parities.

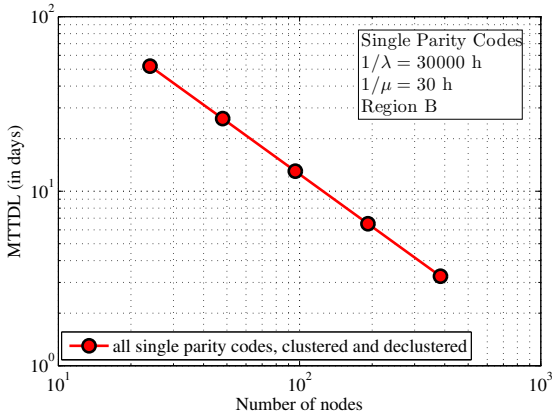
VI. CONCLUSIONS

The effect of latent sector errors on the reliability of a variety of erasure coded data storage systems was investigated. The key findings of this article can be summarized as follows:

- The effect of latent errors depends on the relative magnitudes of the probability of encountering a critical sector error versus the probability of encountering a critical node failure.
- When the probability of encountering a sector error in the critical state is almost one, the reliability of the system is similar to the reliability of a system with no sector errors and one parity less, e.g., reliability of RAID-6 with sector errors resembles that of RAID-5 without sector errors. This effect has been observed in literature for RAID systems with clustered data placement [21]. In this paper, this effect is shown to be true for all symmetric placement schemes and all MDS erasure codes.
- When the probability of encountering a sector error in the critical state is less than one, the MTTDL is scaled down by a factor that depends on the probability, p_S, of a sector having a latent error. This factor is the same for all symmetric data placement schemes except clustered placement. In addition, it is independent of the number of nodes in the system as well the number of parities in the erasure code.
- If the effective probability, p_S, of a sector having an error is reduced by means of scrubbing or intradisk redundancy, larger gains in reliability can be achieved for declustered placement when compared to other symmetric placement schemes.
- It is observed that the practical values of p_S and λ/μ lie in an important range of values where the reliability can be significantly improved by using a combination of declustered data placement and either intradisk redundancy or scrubbing.

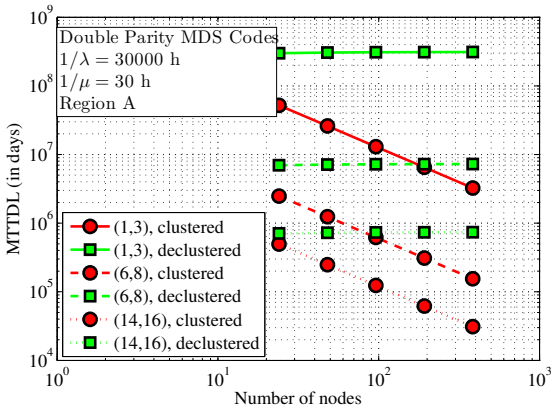


(a) Region A

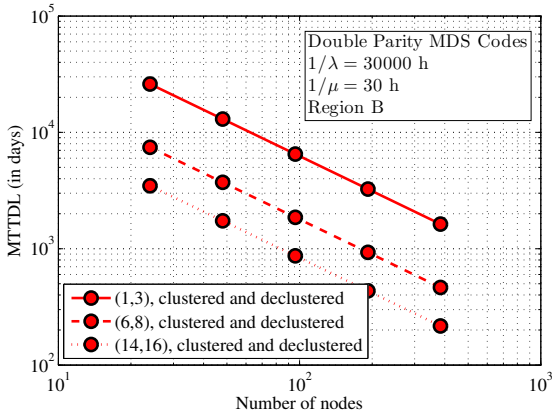


(b) Region B

Figure 2. MTTDL vs. number of nodes for systems using single parity codes with $p_S = 0$ for region A and $p_S = 10^{-7}$ for region B.



(a) Region A



(b) Region B

Figure 3. MTTDL vs. number of nodes for systems using double parity codes with $p_S = 0$ for region A and $p_S = 10^{-7}$ for region B.

REFERENCES

[1] D. A. Patterson, G. Gibson, and R. H. Katz, "A case for redundant arrays of inexpensive disks (RAID)," in *Proc. 1988 ACM SIGMOD Int'l Conference on Management of Data*, 1988, pp. 109–116.

[2] J. Kubiatowicz, D. Bindel, Y. Chen, S. Czerwinski, P. Eaton, D. Geels, R. Gummadi, S. Rhea, H. Weatherspoon, W. Weimer, C. Wells, and B. Zhao, "OceanStore: an architecture for global-scale persistent storage," *SIGPLAN Notices*, vol. 35, no. 11, pp. 190–201, November 2000.

[3] A. G. Dimakis, P. B. Godfrey, M. J. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," in *INFOCOM 2007. 26th IEEE Int'l Conference on Computer Communications*, 2007, pp. 2000–2008.

[4] T. Schwarz, Q. Xin, E. L. Miller, D. D. E. Long, A. Hospodor, and S. Ng, "Disk scrubbing in large archival storage systems," in *Proc. 12th Int'l Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS '04)*, October 2004, pp. 409–418.

[5] I. Iliadis, "Reliability modeling of RAID storage systems with latent errors," in *Proc. IEEE Int'l Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS '09)*, 2009.

[6] A. Thomasian and M. Blaum, "Higher reliability redundant disk arrays: Organization, operation, and coding," *ACM Trans. Storage*, vol. 5, no. 3, pp. 1–59, 2009.

[7] P. M. Chen, E. K. Lee, G. A. Gibson, R. H. Katz, and D. A. Patterson, "RAID: high-performance, reliable secondary storage," *ACM Computing Surveys*, vol. 26, no. 2, pp. 145–185, June 1994.

[8] D. Leong, A. G. Dimakis, and T. Ho, "Distributed storage allocation for high reliability," in *Proc. IEEE Int'l Conference on Communications*, 2010, pp. 1–6.

[9] M. Leslie, J. Davies, and T. Huffman, "A comparison of replication strategies for reliable decentralised storage," *Journal of Networks*, vol. 1, no. 6, pp. 36–44, December 2006.

[10] A. Thomasian and M. Blaum, "Mirrored disk organization reliability analysis," *IEEE Trans. on Computers*, vol. 55, pp. 1640–1644, December 2006.

[11] X. Li, M. Lillibridge, and M. Uysal, "Reliability analysis of deduplicated and erasure-coded storage," *ACM SIGMETRICS Performance Evaluation Review*, vol. 38, no. 3, pp. 4–9, January 2011.

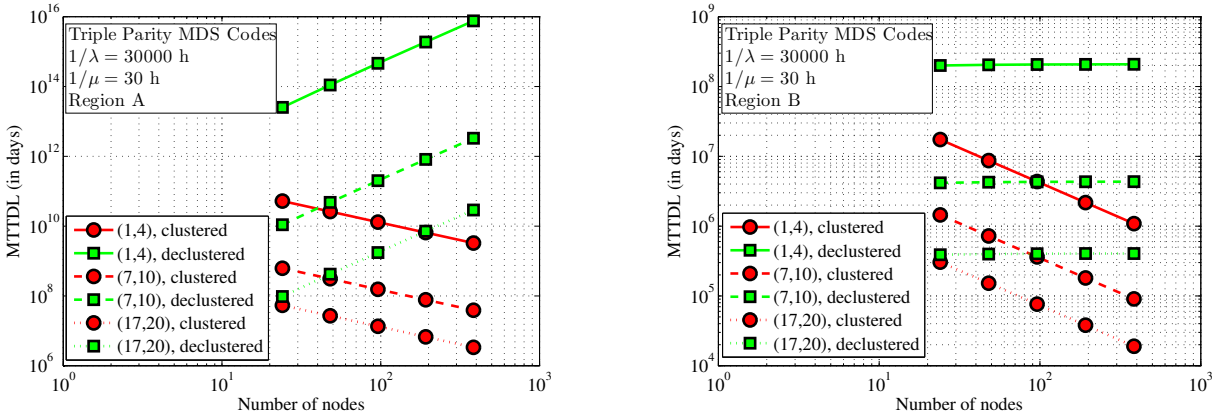
[12] Q. Xin, E. L. Miller, and T. J. E. Schwarz, "Evaluation of distributed recovery in large-scale storage systems," in *Proc. 13th IEEE Int'l Symposium on High Performance Distributed Computing (HPDC'04)*, 2004, pp. 172–181.

[13] V. Venkatesan, I. Iliadis, C. Fragouli, and R. Urbanke, "Reliability of clustered vs. declustered replica placement in data storage systems," in *Proc. 19th Annual IEEE/ACM Int'l Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS'11)*, 2011, pp. 307–317.

[14] V. Venkatesan, I. Iliadis, and R. Haas, "Reliability of data storage systems under network rebuild bandwidth constraints," in *Proc. 2012 IEEE 20th Annual Int'l Symposium on Modelling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS '12)*, 2012, pp. 189 –197.

[15] V. Venkatesan and I. Iliadis, "A general reliability model for data storage systems," in *Proc. 2012 9th Int'l Conference on Quantitative Evaluation of Systems (QEST '12)*, 2012, pp. 209–219.

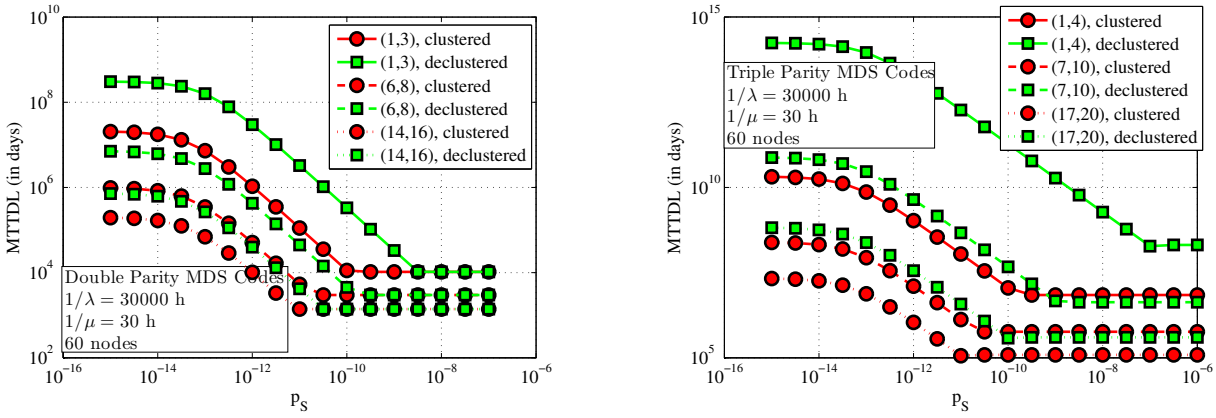
[16] H. Weatherspoon and J. Kubiatowicz, "Erasure coding vs. replication:



(a) Region A

(b) Region B

Figure 4. MTTDL vs. number of nodes for systems using triple parity codes with $p_S = 0$ for region A and $p_S = 10^{-7}$ for region B.



(a) Double parity codes

(b) Triple parity codes

Figure 5. MTTDL vs. probability of sector error p_S for systems with $n = 60$ nodes.

A quantitative comparison," in *Proc. 1st Int'l Workshop on Peer-to-Peer Systems (IPTPS)*, Mar. 2002, pp. 328–338.

[17] J. S. Plank and C. Huang, "Tutorial: Erasure coding for storage applications," Slides presented at 11th Usenix Conference on File and Storage Technologies (FAST'13), San Jose, February 2013.

[18] K. M. Greenan, E. L. Miller, and J. Wylie, "Reliability of flat XOR-based erasure codes on heterogeneous devices," in *Proc. 38th Annual IEEE/IFIP Int'l Conference on Dependable Systems and Networks (DSN'08)*, June 2008, pp. 147–156.

[19] B. Schroeder, S. Damouras, and P. Gill, "Understanding latent sector errors and how to protect against them," *ACM Trans. on Storage*, vol. 6, no. 3, pp. 9:1–9:23, September 2010.

[20] L. N. Bairavasundaram, G. R. Goodson, S. Pasupathy, and J. Schindler, "An analysis of latent sector errors in disk drives," in *Proc. 2007 ACM SIGMETRICS Int'l Conference on Measurement and Modeling of Computer Systems*, ser. SIGMETRICS '07, 2007, pp. 289–300.

[21] A. Dholakia, E. Eleftheriou, X.-Y. Hu, I. Iliadis, J. Menon, and K. Rao, "A new intra-disk redundancy scheme for high-reliability RAID storage systems in the presence of unrecoverable errors," *ACM Trans. on Storage*, vol. 4, no. 1, pp. 1–42, May 2008.

[22] D. Ford, F. Labelle, F. I. Popovici, M. Stokely, V.-A. Truong, L. Barroso, C. Grimes, and S. Quinlan, "Availability in globally distributed storage systems," in *Proc. 9th USENIX Symposium on Operating Systems Design and Implementation (OSDI'10)*, 2010, pp. 61–74.

[23] B. Schroeder and G. A. Gibson, "Understanding disk failure rates: What does an MTTF of 1,000,000 hours mean to you?" *ACM Trans. on Storage*, vol. 3, no. 3, pp. 1–31, October 2007.

[24] W. Jiang, C. Hu, Y. Zhou, and A. Kanevsky, "Are disks the dominant contributor for storage failures?: A comprehensive study of storage subsystem failure characteristics," *ACM Trans. on Storage*, vol. 4, no. 3, pp. 1–25, November 2008.

[25] I. Iliadis, R. Haas, X.-Y. Hu, and E. Eleftheriou, "Disk scrubbing versus intradisk redundancy for RAID storage systems," *ACM Trans. on Storage*, vol. 7, no. 2, pp. 5:1–5:42, July 2011.

[26] V. Venkatesan and I. Iliadis, "Effect of codeword placement on the reliability of data storage systems," in *Proc. 10th Int'l Conference on Quantitative Evaluation of Systems (QEST'13)*, 2013.

APPENDIX A
PROOF OF PROPOSITION 2

Consider a sample direct path with $R_e = \tau_e$, $e = 1, \cdots, \tilde{r} - 1$, and $\alpha_e = a_e$, $e = 1, \cdots, \tilde{r} - 2$.[1] Denote the vector $(\tau_1, \cdots, \tau_{\tilde{r}-1})$ by $\vec{\tau}$ and $(a_1, \cdots, a_{\tilde{r}-2})$ by $\vec{a}$ for notational convenience. Then, the probability of this direct path, denoted by $P_{DL,\text{direct}}(\vec{\tau}, \vec{a})$, is

$$\begin{aligned} P_{DL,\text{direct}}(\vec{\tau}, \vec{a}) &= \Pr\{R_1 = \tau_1\} \times \Pr\{F_2 < R_1 | R_1 = \tau_1\} \\ &\times \Pr\{\alpha_1 = a_1 | R_1 = \tau_1, F_2 < R_1\} \\ &\times \Pr\{R_2 = \tau_2 | R_1 = \tau_1, F_2 < R_1, \alpha_1 = a_1\} \\ &\times \Pr\{F_3 < R_2 | R_1 = \tau_1, F_2 < R_1, \alpha_1 = a_1, R_2 = \tau_2\} \\ &\cdots \times \Pr\left\{F_{\tilde{r}} < R_{\tilde{r}-1} \text{ or CSE} \middle| R_e = \tau_e, F_{e'+1} < R_{e'}, \right. \\ &\left. \alpha_{e'} = a_{e'}, \forall\, e \in \{1, \cdots, \tilde{r} - 1\}, \forall\, e' \in \{1, \cdots, \tilde{r} - 2\}\right\}. \quad (23) \end{aligned}$$

[1] This is a shorthand notation referring to a direct path to data loss with $\tau_e < R_e \leq \tau_e + \delta\tau_e$, $e = 1, \cdots, \tilde{r} - 1$, and $a_e < \alpha_e \leq \delta a_e$, $e = 1, \cdots, \tilde{r} - 2$, where $\delta\tau_e$ and δa_e are positive infinitesimal quantities.

If we denote the mean of R_1 by $1/\mu_1$, based on the rebuild model described in Section II-F, it follows that R_1 is distributed according to some distribution G_{μ_1} that satisfies (5), that is, $R_1 \sim G_{\mu_1}$. Therefore, the first term in (23) reduces to

$$\Pr\{R_1 = \tau_1\} = g_{\mu_1}(\tau_1)\delta\tau_1, \tag{24}$$

where $\delta\tau_1$ denotes an infinitesimal increment in τ_1. Now, denote by p_{CSE} the probability of critical sector error given that the system has reached exposure level $\tilde{r}-1$ through this sample direct path:

$$p_{\text{CSE}} = \Pr\big\{\text{CSE}\big|R_e = \tau_e, F_{e'+1} < R_{e'}, \alpha_{e'} = a_{e'}, \forall\, e \in \{1,\cdots,\tilde{r}-1\}, \forall\, e' \in \{1,\cdots,\tilde{r}-2\}\big\} \tag{25}$$

Using p_{CSE} the last term in (23) can be split into two:

$$\begin{aligned}&\Pr\big\{F_{\tilde{r}} < R_{\tilde{r}-1} \text{ or CSE}\big|R_e = \tau_e, F_{e'+1} < R_{e'}, \alpha_{e'} = a_{e'}, \\ &\forall\, e \in \{1,\cdots,\tilde{r}-1\}, \forall\, e' \in \{1,\cdots,\tilde{r}-2\}\big\} \\ &= p_{\text{CSE}} + (1-p_{\text{CSE}})\Pr\big\{F_{\tilde{r}} < R_{\tilde{r}-1}\big|R_e = \tau_e, F_{e'+1} < R_{e'}, \\ &\alpha_{e'} = a_{e'}, \forall\, e \in \{1,\cdots,\tilde{r}-1\}, \forall\, e' \in \{1,\cdots,\tilde{r}-2\}\big\}.\end{aligned} \tag{26}$$

All terms in (23) other than (24) and (26) fall into three types:

$$\text{A: } \Pr\big\{F_e < R_{e-1}\big|R_{e'} = \tau_{e'}, F_{e''+1} < R_{e''}, \alpha_{e''} = a_{e''}, \forall\, e' \in \{1,\cdots,e-1\}, \forall\, e'' \in \{1,\cdots,e-2\}\big\}, \tag{27}$$

$$\text{B: } \Pr\big\{\alpha_e = a_e\big|R_{e'} = \tau_{e'}, F_{e'+1} < R_{e'}, \alpha_{e''} = a_{e''}, \forall\, e' \in \{1,\cdots,e\}, \forall\, e'' \in \{1,\cdots,e-1\}\big\}, \tag{28}$$

$$\text{C: } \Pr\big\{R_e = \tau_e\big|R_{e'} = \tau_{e'}, F_{e'+1} < R_{e'}, \alpha_{e'} = a_{e'}, \forall\, e' \in \{1,\cdots,e-1\}\big\}. \tag{29}$$

From Lemmas 1, 2, and 3 in [26], expressions A, B, and C reduce to $\tilde{n}_{e-1}\lambda\tau_{e-1}$, δa_e, and $\delta(\tau_e - 1/\mu_e)\delta\tau_e$, respectively. Here, $\delta(\tau_e - 1/\mu_e)$ denotes the Dirac delta function with a spike at $1/\mu_e$, and $\delta\alpha_e$ and $\delta\tau_e$ denote an infinitesimal increment of α_e and τ_e, respectively. Substituting (24), (26), and the expressions A, B, and C in (23), the probability of a sample direct path, $P_{DL,\text{direct}}(\vec{\tau},\vec{a})$, becomes

$$\begin{aligned}P_{DL,\text{direct}}(\vec{\tau},\vec{a}) \approx{}& \lambda^{\tilde{r}-2} \times \tilde{n}_1 \cdots \tilde{n}_{\tilde{r}-2} \times \tau_1 \cdots \tau_{\tilde{r}-2} \times g_{\mu_1}(\tau_1) \\ &\times \delta a_1 \cdots \delta a_{\tilde{r}-2} \times \delta\tau_1 \cdots \delta\tau_{\tilde{r}-1} \\ &\times \delta(\tau_2 - 1/\mu_2)\cdots\delta(\tau_{\tilde{r}-1} - 1/\mu_{\tilde{r}-1}) \\ &\times ((1-p_{\text{CSE}})\tilde{n}_{\tilde{r}-1}\lambda\tau_{\tilde{r}-1} + p_{\text{CSE}}).\end{aligned} \tag{30}$$

The probability of the direct path to data loss, $P_{DL,\text{direct}}$, is the sum of the probabilities, $P_{DL,\text{direct}}(\vec{\tau},\vec{a})$, of all possible sample direct paths. As the infinitesimal increments in (30) tend to zero, the sum becomes an integral. Therefore,

$$\begin{aligned}P_{DL,\text{direct}} \approx{}& \lambda^{\tilde{r}-2} \times \tilde{n}_1 \cdots \tilde{n}_{\tilde{r}-2} \\ &\times \int_{\tau_1}\cdots\int_{\tau_{\tilde{r}-1}}\int_{a_1}\cdots\int_{a_{\tilde{r}-2}} \Bigg(\tau_1 \cdots \tau_{\tilde{r}-2} g_{\mu_1}(\tau_1) \\ &\times \delta\left(\tau_2 - \frac{1}{\mu_2}\right)\cdots\delta\left(\tau_{\tilde{r}-1} - \frac{1}{\mu_{\tilde{r}-1}}\right) \\ &\times ((1-p_{\text{CSE}})\tilde{n}_{\tilde{r}-1}\lambda\tau_{\tilde{r}-1} + p_{\text{CSE}}) d\vec{a} d\vec{\tau}\Bigg).\end{aligned} \tag{31}$$

Here, the integrals are from 0 to ∞ for τ_e, $e = 1,\cdots,\tilde{r}-1$, and from 0 to 1 for a_e, $e = 1,\cdots,\tilde{r}-2$. Changing the order of integrals, integrating over $\tau_{\tilde{r}-1}$, and rearranging, we obtain

$$\begin{aligned}P_{DL,\text{direct}} \approx{}& \lambda^{\tilde{r}-1} \times \tilde{n}_1 \cdots \tilde{n}_{\tilde{r}-1} \\ &\times \int_{a_1}\cdots\int_{a_{\tilde{r}-2}}\int_{\tau_1}\cdots\int_{\tau_{\tilde{r}-2}} \Bigg(\tau_1 \cdots \tau_{\tilde{r}-2}\frac{1}{\mu_{\tilde{r}-1}} g_{\mu_1}(\tau_1) \\ &\times \delta\left(\tau_2 - \frac{1}{\mu_2}\right)\cdots\delta\left(\tau_{\tilde{r}-2} - \frac{1}{\mu_{\tilde{r}-2}}\right) \\ &\times \left(1 + p_{\text{CSE}}\frac{\mu_{\tilde{r}-1}}{\tilde{n}_{\tilde{r}-1}\lambda}\left(1 - \frac{\tilde{n}_{\tilde{r}-1}\lambda}{\mu_{\tilde{r}-1}}\right)\right) d\tau_{\tilde{r}-2}\cdots d\tau_1 d\vec{a}\Bigg).\end{aligned} \tag{32}$$

Let the amount of critical data to be rebuilt in exposure level $\tilde{r}-1$ be $D_{\tilde{r}-1}$. This amount is dependent on the sample direct path. For the codeword reconstruction process, the l surviving symbols of this data needs to be read from the surviving nodes. So the total amount of data that is read during the reconstruction process is $lD_{\tilde{r}-1}$. The probability of a critical sector error in this data, p_{CSE}, is given by (3):

$$p_{\text{CSE}} = P_S(lD_{\tilde{r}-1}) \approx \min(1, lD_{\tilde{r}-1}p_S/(sB)). \tag{33}$$

Since the exact value of $D_{\tilde{r}-1}$ is path dependent, we consider the expected value of $D_{\tilde{r}-1}$, namely, $E[D_{\tilde{r}-1}]$, and distinguish between two regions, A and B, as defined in (12) and (13). In region B, the probability of a critical sector error is essentially one. So,

$$p_{\text{CSE}} \approx \begin{cases}\frac{lD_{\tilde{r}-1}}{sB}p_S & \text{in region A,}\\ 1 & \text{in region B.}\end{cases} \tag{34}$$

Region A: Let the average speed of rebuild in exposure level $\tilde{r}-1$ be $S_{\tilde{r}-1}$. The average amount of time taken to rebuild this data is then given by

$$1/\mu_{\tilde{r}-1} = D_{\tilde{r}-1}/S_{\tilde{r}-1}. \tag{35}$$

The speed of rebuild $S_{\tilde{r}-1}$ depends on the codeword placement scheme but is independent of the sample direct path. From (35) and (34), we obtain

$$p_{\text{CSE}}\frac{\mu_{\tilde{r}-1}}{\tilde{n}_{\tilde{r}-1}\lambda}\left(1 - \frac{\tilde{n}_{\tilde{r}-1}\lambda}{\mu_{\tilde{r}-1}}\right) = \frac{p_S l S_{\tilde{r}-1}}{sB\lambda\tilde{n}_{\tilde{r}-1}}\left(1 - \frac{\tilde{n}_{\tilde{r}-1}\lambda D_{\tilde{r}-1}}{S_{\tilde{r}-1}}\right). \tag{36}$$

For systems with generally reliable nodes satisfying (4), it can be shown that

$$\tilde{n}_{\tilde{r}-1}\lambda D_{\tilde{r}-1}/S_{\tilde{r}-1} \ll 1 \tag{37}$$

for all symmetric placement schemes (see Sections IV-A, IV-B, and IV-C). Therefore, (36) reduces to

$$p_{\text{CSE}}\frac{\mu_{\tilde{r}-1}}{\tilde{n}_{\tilde{r}-1}\lambda}\left(1 - \frac{\tilde{n}_{\tilde{r}-1}\lambda}{\mu_{\tilde{r}-1}}\right) \approx \frac{p_S l S_{\tilde{r}-1}}{sB\lambda\tilde{n}_{\tilde{r}-1}}, \tag{38}$$

which is independent of the sample direct path because $S_{\tilde{r}-1}$ is independent of the sample path. Substituting (38) into (32), and recognizing the remaining terms as $P^{\text{noSE}}_{DL,\text{direct}}(\tilde{r})$, we get (14) for region A.

Region B: In region B, the probability, p_{CSE}, of a critical sector error is equal to one. Therefore, (31) reduces to $P^{\text{noSE}}_{DL}(\tilde{r}-1)$ and we obtain (14) for region B.