Effect of Lazy Rebuild on Reliability of Erasure-Coded Storage Systems

Ilias Iliadis
IBM Research Europe – Zurich
8803 Rüschlikon, Switzerland
email: ili@zurich.ibm.com

Abstract—**Erasure-coding redundancy schemes are employed to ensure improved reliability of storage systems against device failures. The effect of the lazy rebuild scheme on the Mean Time to Data Loss** (MTTDL) **and the Expected Annual Fraction of Data Loss** (EAFDL) **reliability metrics is evaluated. A theoretical model that considers the effect of latent errors and device failures is developed. Analytical reliability expressions for the symmetric, clustered, and declustered data placement schemes are derived. It is demonstrated that the employment of lazy rebuild results in a reliability degradation of orders of magnitude. Independently of whether lazy rebuild is used, for realistic values of sector error rates, the results obtained demonstrate that** MTTDL **degrades, whereas** EAFDL **remains practically unaffected. It is also shown that the declustered data placement scheme offers superior reliability.**

Keywords–Storage; Deferred recovery or repair; Unrecoverable or latent sector errors; Reliability analysis; MTTDL; EAFDL; RAID; MDS codes; stochastic modeling.

I. INTRODUCTION

Efficient erasure coding schemes that provide high data reliability are employed in today's large-scale data storage systems to recover data lost due to device and component failures. Special cases of erasure codes are the replication schemes and the Redundant Arrays of Inexpensive Disks (RAID) schemes, such as RAID-5 and RAID-6, which have been deployed extensively in the past thirty years [1-4]. Modern storage systems though use advanced, powerful erasure coding schemes that offer high storage efficiency and improve data reliability [5-8]. The reliability of storage systems is also improved by employing a declustered data placement scheme, but is adversely affected by latent or unrecoverable sector errors that are discovered when there is an attempt to access these sectors [9]. Permanent losses of data due to latent errors are quite pronounced in higher-capacity HDDs and storage nodes [10-12].

Despite the reduction in storage overhead and the improvement of reliability achieved, erasure coding is hindered from becoming more pervasive in large-scale distributed storage systems by the *repair problem*. This issue arises from the increased network traffic needed to repair data lost due to device failures and generated by the downloads and disk IOPS performed during the data recovery process [6][7][13]. To cope with the repair problem and reduce the amount of data transmitted during rebuilds, a lazy rebuild scheme was proposed in [14]. A careful scheduling of rebuild operations substantially reduces recovery bandwidth, while keeping the impact on read performance and data durability low. Lazy recovery reduces repair bandwidth at the expense of increasing the amount of degraded stripes, which in turn affects system reliability. The lazy recovery scheme bears some similarity to the practice of delaying recovery of failed nodes by a fixed amount of time (typically 15 minutes) to avoid unnecessary repairs of short transient failures [5]. The main difference, however, is that a lazy repair is initiated based on the state of the system and does not depend on the time that has elapsed after a node failure. This results in transferring less data than the delayed recovery scheme.

The key contributions of this article are the following. We consider the reliability of erasure-coded storage systems when a lazy rebuild scheme is employed and derive closed-form expressions for the MTTDL and EAFDL reliability metrics for the symmetric, clustered, and declustered data placement schemes. We adopt the non-Markovian methodology developed in prior work [15-17] to evaluate MTTDL and EAFDL of storage systems. The validity of this methodology for accurately assessing the reliability of storage systems has been confirmed by simulations in several contexts [3][15][18][19]. It has been demonstrated that theoretical predictions of the reliability of systems comprising highly reliable storage devices are in good agreement with simulation results. Consequently, the emphasis of the present work is on theoretically assessing the effect of lazy rebuilds on the reliability of storage systems. We extend the reliability model presented in [9] to take into account lazy rebuilds. The model developed is relevant and realistic because it properly captures the characteristics of erasure coding and of the rebuild process associated with the declustered placement scheme currently used by Google [5], Microsoft Azure [7], Facebook [13], and DELL/EMC [20]. The theoretical reliability results obtained here can be used to determine the parameter values that ensure a desired level of reliability. They can also be used to assess system reliability when scrubbing is employed by applying the methodology described in [21]. We subsequently use these results to demonstrate the effect of latent errors and system parameters on system reliability.

The remainder of the article is organized as follows. Section II describes the storage system model and the corresponding parameters considered. Section III presents the general framework and methodology for deriving the MTTDL and EAFDL metrics analytically for the case of erasure-coded systems that employ a lazy rebuild scheme. Closed-form expressions for relevant reliability metrics are derived for the symmetric, clustered, and declustered data placement schemes. Section IV presents numerical results demonstrating the effectiveness of erasure coding schemes for improving system reliability as well as the adverse effect of lazy rebuilds.

TABLE I. NOTATION OF SYSTEM PARAMETERS

Parameter	Definition
n	number of storage devices
c	amount of data stored on each device
l	number of user-data symbols per codeword $(l \geq 1)$
m	total number of symbols per codeword $(m > l)$
(m, l)	MDS-code structure
d	lazy rebuild threshold $(0 \leq d < m - l)$
s	symbol size
k	spread factor of the data placement scheme, or group size (number of devices in a group) $(m \leq k \leq n)$
b	average reserved rebuild bandwidth per device
B_{max}	upper limitation of the average network rebuild bandwidth
X	time required to read (or write) an amount c of data at an average rate b from (or to) a device
$F_X(.)$	cumulative distribution function of X
$F_\lambda(.)$	cumulative distribution function of device lifetimes
P_{bit}	probability of an unrecoverable bit error
s_{eff}	storage efficiency of redundancy scheme $(s_{\text{eff}} = l/m)$
U	amount of user data stored in the system $(U = s_{\text{eff}}\, n\, c)$
$\tilde{r}$	MDS-code distance: minimum number of codeword symbols lost that lead to permanent data loss $(\tilde{r} = m - l + 1$ and $2 \leq \tilde{r} \leq m)$
C	number of symbols stored in a device $(C = c/s)$
μ^{-1}	mean time to read (or write) an amount c of data at an average rate b from (or to) a device $(\mu^{-1} = E(X) = c/b)$
λ^{-1}	mean time to failure of a storage device $(\lambda^{-1} = \int_0^\infty [1 - F_\lambda(t)]dt)$
P_s	probability of an unrecoverable sector (symbol) error
P_{DL}	probability of data loss during rebuild
P_{UF}	probability of data loss due to unrecoverable failures during rebuild
P_{DF}	probability of data loss due to a disk failure during rebuild
Q	amount of lost user data during rebuild
H	amount of lost user data, given that data loss has occurred, during rebuild
S	number of lost symbols during rebuild

Finally, we conclude in Section V.

II. STORAGE SYSTEM MODEL

Here we briefly review the operational characteristics of erasure-coded storage systems. To assess their reliability, we adopt the model used in [9] and extend it to cover the case of lazy rebuilds. The storage system comprises n storage devices (nodes or disks), where each device stores an amount c of data such that the total storage capacity of the system is $n\,c$. This does not account for the spare space used by the rebuild process.

User data is divided into blocks of fixed size s and complemented with parity symbols to form codewords. Maximum Distance Separable (MDS) erasure codes (m, l) that map l user-data symbols to codewords of m symbols are employed. They have the property that any subset containing l of the m codeword symbols can be used to reconstruct (recover) a codeword. The corresponding storage efficiency s_{eff} and amount U of user data stored in the system is

$$s_{\text{eff}} = l/m \quad \text{and} \quad U = s_{\text{eff}}\, n\, c \;=\; l\, n\, c/m \;. \tag{1}$$

Also, the number C of symbols stored in a device is

$$C = c/s \;. \tag{2}$$

Our notation is summarized in Table I. The derived parameters are listed in the lower part of the table. To minimize the risk of permanent data loss, the m symbols of each codeword are spread and stored on m distinct devices. This way, the system can tolerate any $\tilde{r} - 1$ device failures, but $\tilde{r}$ device failures may lead to data loss, with

$$\tilde{r} = m - l + 1 \;, \quad 1 \leq l < m \quad \text{and} \quad 2 \leq \tilde{r} \leq m \;. \tag{3}$$

Examples of MDS erasure codes are the replication, RAID-5, RAID-6, and Reed–Solomon schemes.

Data is stored according to symmetric placement schemes, including the *clustered* and *declustered* placement schemes, as shown in Figure 1 of [9][17]. The system comprises n/k disjoint groups of k devices. Within each group, all $\binom{k}{m}$ possible ways of placing m symbols across k devices are used equally to store all the codewords in that group. Refer to [9] for additional details.

A. Codeword Reconstruction and Rebuild Process

When storage devices fail, codewords lose some of their symbols, and this reduces data redundancy. The system attempts to maintain its redundancy by reconstructing the lost codeword symbols using the surviving symbols of the affected codewords. As the times to detect device failures are much shorter than rebuild times, we assume that failures are detected instantaneously. When a *lazy rebuild* scheme is used, the rebuild process is not triggered immediately, but is delayed until additional device failures occur that result in d additional symbol losses within some of the codewords. Consequently, the rebuild process is initiated when codewords have lost $1 + d$ symbols. To avoid permanent data losses, the number of symbols lost within codewords should be less than the MDS-code distance $\tilde{r}$, that is, this number should not exceed $\tilde{r} - 1$, which implies that $d + 1 \leq \tilde{r} - 1 = m - l$. Thus, we have

$$l < m \leq k \leq n \quad (n/k \in \mathbb{N}) \quad \text{and} \quad 0 \leq d \leq m - l - 1 \;. \tag{4}$$

1) Exposure Levels: The system is at exposure level u $(0 \leq u \leq \tilde{r})$ when there are codewords that have lost u symbols owing to device failures, but there are no codewords that have lost more symbols. These codewords are referred to as the *most-exposed* codewords. Transitions to higher exposure levels are caused by device failures, whereas transitions to lower ones are caused by successful rebuilds. We denote by C_u the number of most-exposed codewords upon entering exposure level u, $(u \geq 1)$. Upon the first device failure it holds that

$$C_1 = C \;, \tag{5}$$

where C is determined by (2). In Section III, we will derive the reliability metrics of interest using the *direct path approximation*, which considers only transitions from lower to higher exposure levels [3][15][18][19][22]. This implies that each exposure level is entered only once.

2) Prioritized Lazy Rebuild: When a symmetric or declustered placement scheme is used, as shown in Figure 2 of [9][17], spare space is reserved on each device for temporarily storing the reconstructed codeword symbols before they are transferred to a new replacement device. The rebuild process to restore the data lost by failed devices is assumed to be both *prioritized* and *distributed*. A prioritized (or intelligent) rebuild process always attempts first to rebuild the *most-exposed* codewords, namely, the codewords that have lost the largest number of symbols [3][5][7][14][17][18]. According to the lazy rebuild scheme, no recovery actions are performed at exposure levels u not exceeding the threshold d. However, when the system enters a higher exposure level u, the rebuild process is triggered and attempts to bring the system back to exposure level $u - 1$ by reading l symbols and recovering one of the u symbols that each of the C_u most-exposed codewords has lost. To improve reliability, the vulnerability window is reduced by recovering only one symbol as opposed to the scheme considered in [14] that recovers multiple symbols. In a distributed rebuild process, the codewords are reconstructed by reading symbols from an appropriate set of surviving devices and storing the recovered symbols in the reserved spare space of these devices. During this process, it is desirable to reconstruct the lost codeword symbols on devices in which another symbol of the same codeword is not already present.

In the case of clustered placement, the codeword symbols are spread across all $k \ (= m)$ devices in each group (cluster). Therefore, reconstructing the lost symbols on the surviving devices of a group would result in more than one symbol of the same codeword on the same device. To avoid this, the lost symbols are reconstructed directly in spare devices as described and shown in Figure 3 of [17].

3) Rebuild Process: A certain portion of the device bandwidth is reserved for read/write data recovery during the rebuild process, and the remaining bandwidth is used to serve user requests. Let b denote the actual average reserved rebuild bandwidth per device. Lost symbols are rebuilt in parallel using the rebuild bandwidth b available on each surviving device. The amount of data corresponding to the number C_u of symbols to be rebuilt at exposure level u is written at an average rate $b_u \ (\leq b)$ to selected device(s). For the time X required to read (or write) an amount c of data from (or to) a device it holds that

$$\mu^{-1} \ \triangleq \ E(X) = c/b \, . \tag{6}$$

4) Failure and Rebuild Time Distributions: The lifetimes of the n devices are assumed to be independent and identically distributed, with a cumulative distribution function $F_\lambda(.)$ and a mean of $1/\lambda$. The results in this article hold for *highly reliable* storage devices, which satisfy the condition [17][19]

$$\mu \int_0^\infty F_\lambda(t)[1 - F_X(t)]dt \ll 1, \quad \text{with } \frac{\lambda}{\mu} \ll 1 \, . \tag{7}$$

5) Amount of Data to Rebuild and Rebuild Times at Each Exposure Level: We denote by $\tilde{n}_u$ the number of devices at exposure level u whose failure causes an exposure level transition to level $u+1$, and V_u the fraction of the C_u most-exposed codewords that have a symbol stored on any given such device. Note that $\tilde{n}_u$ depends on the codeword placement scheme. Let R_u denote the rebuild time of the most-exposed codewords at exposure level u and α_u be the fraction of the rebuild time R_u still left when another device fails, causing the exposure level transition $u \to u+1$. For $u \leq d$, no rebuild is performed and therefore $\alpha_u = 1$. For $u > d$, α_u is approximately uniformly distributed in $(0,1)$ [23, Lemma 2]. Therefore,

$$\alpha_u \ \approxeq \begin{cases} 1 \, , & \text{for } \ u = 1, \ldots, d \\ U(0,1) \, , & \text{for } \ u = d+1, \ldots, \tilde{r}-1 \, . \end{cases} \tag{8}$$

We proceed by considering that the rebuild time R_{u+1} is determined completely by R_u and α_u in the same manner as in [16][17][22]. For the rebuild schemes considered, the fraction of the C_u most-exposed codewords that were not yet considered by the rebuild process upon the next device failure is roughly equal to the fraction α_u of the rebuild time R_u still left. Therefore, upon the next device failure, an approximate number $\alpha_u \, C_u$ of the C_u codewords were not yet considered by the rebuild process. Clearly, the fraction V_u of these codewords that have symbols stored on the newly failed device depends only on the codeword placement scheme. Consequently, the number C_{u+1} of the most-exposed codewords upon entering exposure level $u+1$ is

$$C_{u+1} \approx \ V_u \, \alpha_u \, C_u \ , \ \text{ for } \ u = 1, \ldots, \tilde{r}-1 \, . \tag{9}$$

Repeatedly applying (9) and using (5) and the convention that for any sequence δ_i, $\prod_{i=1}^{0} \delta_i \triangleq 1$, yields

$$C_u \ \approx \ C \prod_{i=1}^{u-1} V_i \, \alpha_i \ , \ \text{ for } \ u = 1, \ldots, \tilde{r} \, . \tag{10}$$

6) Unrecoverable Errors: The reliability of storage systems is affected by the occurrence of unrecoverable or latent errors. Let P_{bit} denote . According to the specifications, the unrecoverable bit-error probability P_{bit} is equal to 10^{-15} for SCSI drives and 10^{-14} for SATA drives [21]. Assuming that bit errors occur independently over successive bits, the unrecoverable sector (symbol) error probability P_s is

$$P_s = 1 - (1 - P_{\text{bit}})^s \, , \tag{11}$$

with s expressed in bits. Assuming a sector size of 512 bytes, the equivalent unrecoverable sector error probability is $P_s \approx P_{\text{bit}} \times 4096$, which is 4.096×10^{-12} in the case of SCSI and 4.096×10^{-11} in the case of SATA drives. In practice, however, and also owing to the accumulation of latent errors over time, these probability values are higher. Indeed, empirical field results suggest that the actual values can be orders of magnitude higher, reaching $P_s \approx 5 \times 10^{-9}$ [24].

III. DERIVATION OF MTTDL AND EAFDL

The reliability metrics are derived using the direct-path-approximation methodology presented in [9][15][16][17] and extending it to assess the effect of lazy rebuilds.

At any point in time, the system is in one of two modes: non-rebuild or rebuild mode. Note that part of the non-rebuild mode is the normal mode of operation where all devices are operational and all data in the system has the original amount of redundancy. In the context of lazy rebuild, when the first device fails, the system does not enter the rebuild mode. Subsequently, we refer to the device failure that causes the transition from non-rebuild to rebuild mode as an *initial device* failure, which should not be confused with the first device failure. Consequently, an *initial device* failure triggers a rebuild process that attempts to restore the lost data, which eventually leads the system either to a Data Loss (DL) with probability P_{DL} or back to the original normal mode by restoring initial redundancy, with probability $1 - P_{\text{DL}}$.

Let T be a typical interval of a non-rebuild period, that is, the time interval from the time the system is brought to its original state until a subsequent initial device failure occurs that causes the system to enter exposure level $d+1$. It then holds that $T = \sum_{u=0}^{d} T_u$, where T_0 denotes the time interval from the time the system is brought to its original state until the first device failure and T_u denotes the time that the system spends at exposure level u. For a system comprising n devices with a mean time to failure of a device equal to $1/\lambda$, it holds that $E(T_0) = 1/(n\,\lambda)$. Given that the number of devices at exposure level u whose failure causes an exposure level transition to level $u+1$ is $\tilde{n}_u$, it holds that $E(T_u) = 1/(\tilde{n}_u\, \lambda)$. From the above, it follows that

$$E(T) = \sum_{u=0}^{d} E(T_u) = \left(\sum_{u=0}^{d} \frac{1}{\tilde{n}_u} \right) \Big/ \lambda \, , \quad \text{where } \tilde{n}_0 \triangleq n \, , \tag{12}$$

where $\tilde{n}_u$ is determined by (35) or (38).

The MTTDL metric is then obtained by [15, Eq. (5)]:

$$\text{MTTDL} \approx \frac{E(T)}{P_{\text{DL}}} \, . \tag{13}$$

The EAFDL is obtained as the ratio of the expected amount $E(Q)$ of lost user data, normalized to the amount U of user data, to the expected duration of T [15, Eq. (9)]:

$$\text{EAFDL} \approx \frac{E(Q)}{E(T) \cdot U} \overset{(1)}{=} \frac{m \, E(Q)}{n \, l \, c \, E(T)} \, . \tag{14}$$

where $E(T)$ is determined by (12) and expressed in years.

The expected amount $E(H)$ of lost user data, given that data loss has occurred, is determined by [15, Eq. (8)]:

$$E(H) = \frac{E(Q)}{P_{\text{DL}}} \, . \tag{15}$$

A. Reliability Analysis

The reliability evaluation of the lazy rebuild scheme is based on the reliability analysis presented in [9]. The MTTDL and EAFDL reliability metrics were determined by first deriving the probability of data loss P_{DL} and the expected amount $E(Q)$ of lost user data. Central to these derivations are the variables α_u that represent the fractions of the rebuild times R_u still left when device failures cause exposure level transitions. These variables were assumed to be independent and approximately uniformly distributed in $(0, 1)$. However, in the case of lazy rebuild, these variables are distributed according to (8). We now proceed to derive the various measures of interest.

At any exposure level u $(u = d+1, \ldots, \tilde{r}-1)$, data loss may occur during rebuild owing to one or more unrecoverable failures, which is denoted by the transition $u \to$ UF. Moreover, at exposure level $\tilde{r}-1$, data loss occurs owing to a subsequent device failure, which leads to the transition to exposure level $\tilde{r}$. Consequently, the direct paths that lead to data loss are the following:

- $\overrightarrow{UF_u}$: the direct path of successive transitions $1 \to 2 \to \cdots \to u \to$ UF, for $u = d+1, \ldots, \tilde{r}-1$, and
- $\overrightarrow{DF}$: the direct path of successive transitions $1 \to 2 \to \cdots \to \tilde{r}-1 \to \tilde{r}$,

with corresponding probabilities P_{UF_u} and P_{DF}, respectively.

1) Data Loss: It holds that

$$P_{\text{UF}_u} = P_u \; P_{u \to \text{UF}} \; , \quad \text{for} \quad u = d+1, \ldots, \tilde{r}-1 \; , \tag{16}$$

where P_u is the probability of entering exposure level u, which is derived in Appendix A as follows:

$$P_u \approx \frac{(\lambda \, c \prod_{j=1}^{d} V_j)^{u-d-1}}{(u-d-1)!} \, \frac{E(X^{u-d-1})}{[E(X)]^{u-d-1}} \prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} \, V_i^{u-1-i} \, , \tag{17}$$

and $P_{u \to \text{UF}}$ is the probability of encountering an unrecoverable failure during the rebuild process at this exposure level.

In [25], it was shown that P_{DL} is accurately approximated by the probability of all direct paths to data loss. Therefore,

$$P_{\text{DL}} \approx P_{\text{DF}} + \sum_{u=d+1}^{\tilde{r}-1} P_{\text{UF}_u} \, . \tag{18}$$

Approximate expressions for the probabilities of data loss P_{UF_u} and P_{DF} are subsequently obtained by the following proposition.

Proposition 1: For $u = d+1, \ldots, \tilde{r}-1$, it holds that

$$\begin{aligned} & P_{\text{UF}_u} \\ & \approx - \left(\lambda \, c \prod_{j=1}^{d} V_j \right)^{u-d-1} \frac{E(X^{u-d-1})}{[E(X)]^{u-d-1}} \left(\prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} \, V_i^{u-1-i} \right) \\ & \qquad \cdot \log(\hat{q}_u)^{-(u-d-1)} \left(\hat{q}_u - \sum_{i=0}^{u-d-1} \frac{\log(\hat{q}_u)^i}{i!} \right) , \end{aligned} \tag{19}$$

where

$$\hat{q}_u \triangleq q_u^{C \prod_{j=1}^{u-1} V_j} \, , \tag{20}$$

$$q_u = 1 - \sum_{j=\tilde{r}-u}^{m-u} \binom{m-u}{j} P_s^j (1-P_s)^{m-u-j} \, , \tag{21}$$

$$P_{\text{DF}} \approx \frac{(\lambda \, c \prod_{j=1}^{d} V_j)^{\tilde{r}-d-1}}{(\tilde{r}-d-1)!} \, \frac{E(X^{\tilde{r}-d-1})}{[E(X)]^{\tilde{r}-d-1}} \prod_{i=d+1}^{\tilde{r}-1} \frac{\tilde{n}_i}{b_i} \, V_i^{\tilde{r}-1-i} . \tag{22}$$

Proof: Equation (19) is obtained in Appendix A. Equation (22) is obtained from the fact that $P_{\text{DF}} = P_{\tilde{r}}$ and, subsequently, from (17) by setting $u = \tilde{r}$. ■

The MTTDL metric is obtained by substituting (18) into (13) as follows:

$$\text{MTTDL} \approx \frac{E(T)}{P_{\text{DF}} + \sum_{u=d+1}^{\tilde{r}-1} P_{\text{UF}_u}} \, , \tag{23}$$

where $E(T)$, P_{UF_u} and P_{DF} are determined by (12), (19), and (22), respectively.

2) Amount of Data Loss: We proceed to derive the amount of data loss during rebuild. Let Q, H, and S be the amount of lost user data, the conditional amount of lost user data, given that data loss has occurred, and the number of lost symbols, respectively. Let also Q_{DF} and Q_{UF_u} denote the amount of lost user data associated with the direct paths $\overrightarrow{DF}$ and $\overrightarrow{UF_u}$, respectively. Similarly, we consider the variables H_{DF}, H_{UF_u}, S_{DF}, and S_{UF_u}. Then, the amount Q of lost user data is obtained by

$$Q \approx \begin{cases} H_{\text{DF}} \, , & \text{if } \overrightarrow{DF} \\ H_{\text{UF}_u} \, , & \text{if } \overrightarrow{UF_u} \, , \ \text{ for } \ u = d+1, \ldots, \tilde{r}-1 \\ 0 \, , & \text{otherwise} \, . \end{cases} \tag{24}$$

$$\text{Thus,}\quad E(Q) \approx P_{\text{DF}}\, E(H_{\text{DF}}) + \sum_{u=d+1}^{\tilde{r}-1} P_{\text{UF}_u}\, E(H_{\text{UF}_u}) \tag{25}$$

$$= E(Q_{\text{DF}}) + \sum_{u=d+1}^{\tilde{r}-1} E(Q_{\text{UF}_u}) \;, \tag{26}$$

$$\text{where}\quad E(Q_{\text{DF}}) = P_{\text{DF}}\, E(H_{\text{DF}}) \;, \tag{27}$$

$$\text{and}\quad E(Q_{\text{UF}_u}) = P_{\text{UF}_u}\, E(H_{\text{UF}_u}) \;,\;\; u = d+1, \ldots, \tilde{r}-1 \;. \tag{28}$$

Note that the expected amount $E(Q)$ of lost user data is equal to the product of the storage efficiency and the expected amount of lost data, where the latter is equal to the product of the expected number of lost symbols $E(S)$ and the symbol size s. Consequently, it follows from (1) that

$$E(Q) = \frac{l}{m}\, E(S)\, s \overset{(2)}{=} \frac{l}{m}\, \frac{E(S)}{C}\, c \;. \tag{29}$$

$$\text{Similarly,}\quad E(Q_{\text{DF}}) = \frac{l}{m}\, E(S_{\text{DF}})\, s \overset{(2)}{=} \frac{l}{m}\, \frac{E(S_{\text{DF}})}{C}\, c \;, \tag{30}$$

$$\text{and}\quad E(Q_{\text{UF}_u}) = \frac{l}{m}\, E(S_{\text{UF}_u})\, s \overset{(2)}{=} \frac{l}{m}\, \frac{E(S_{\text{UF}_u})}{C}\, c \;. \tag{31}$$

Proposition 2: For $u = d+1, \ldots, \tilde{r}-1$, it holds that

$$E(Q_{\text{UF}_u}) \approx c\, \frac{l\,\tilde{r}}{m}\, \frac{\left(\lambda\, c \prod_{j=1}^{d} V_j\right)^{u-d-1}}{(u-d)!}\, \frac{E(X^{u-d-1})}{[E(X)]^{u-d-1}} \left(\prod_{j=1}^{d} V_j\right) \cdot \left(\prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i}\, V_i^{u-i}\right) \binom{m-u}{\tilde{r}-u} P_s^{\tilde{r}-u}, \quad P_s \ll \frac{1}{m-\tilde{r}} \;, \tag{32}$$

$$E(Q_{\text{DF}}) \approx c\, \frac{l\,\tilde{r}}{m} \left(\lambda\, c \prod_{j=1}^{d} V_j\right)^{\tilde{r}-d-1} \frac{1}{(\tilde{r}-d)!}\, \frac{E(X^{\tilde{r}-d-1})}{[E(X)]^{\tilde{r}-d-1}} \cdot \left(\prod_{j=1}^{d} V_j\right) \left(\prod_{i=d+1}^{\tilde{r}-1} \frac{\tilde{n}_i}{b_i}\, V_i^{\tilde{r}-i}\right) . \tag{33}$$

Proof: Equation (32) is obtained in Appendix B. Equation (33) is obtained from (32) by setting $u = \tilde{r}$. ■

The EAFDL metric is obtained by substituting (26) into (14) as follows:

$$\text{EAFDL} \approx \frac{m\,[E(Q_{\text{DF}}) + \sum_{u=d+1}^{\tilde{r}-1} E(Q_{\text{UF}_u})]}{n\; l\; c\; E(T)} \;, \tag{34}$$

where $E(Q_{\text{UF}_u})$ and $E(Q_{\text{DF}})$ are determined by (32) and (33), respectively, and $E(T)$ by (12) and expressed in years.

The conditional amount $E(H)$ of lost user data, given that data loss has occurred, is obtained from (15), P_{DL} is determined by (18), (19), and (22), and $E(Q)$ is determined by (26), (32), and (33).

B. *Symmetric and Declustered Placement*

We consider the case $m < k \le n$. The special case $k = m$ corresponding to the clustered placement has to be considered separately for the reasons discussed in Section II-A2. At each exposure level u, for $u = 1, \cdots, \tilde{r}-1$, it holds that [16][17]

$$\tilde{n}_u^{\text{sym}} = k - u \;, \tag{35}$$

$$b_u^{\text{sym}} = \frac{\min((k-u)\, b, B_{\max})}{l+1} \;, \tag{36}$$

$$V_u^{\text{sym}} = \frac{m-u}{k-u} \;. \tag{37}$$

The corresponding parameters $\tilde{n}_u^{\text{declus}}$, b_u^{declus}, and V_u^{declus} for the declustered placement are derived from (35), (36), and (37) by setting $k = n$.

C. *Clustered Placement*

At each exposure level u, for $u = 1, \cdots, \tilde{r}-1$, it holds that [16][17]

$$\tilde{n}_u^{\text{clus}} = m - u \;,\quad b_u^{\text{clus}} = \min(\, b\,, B_{\max}/l\,) \;,\quad V_u^{\text{clus}} = 1 \;. \tag{38}$$

Remark 1: From (19), (22), (32), and (33), and considering expressions (35) through (38), it follows that P_{UF_u} and $E(Q_{\text{UF}_u})$ are mainly determined by the term $(\lambda\, c/b)^{u-d-1}$, and P_{DF} and $E(Q_{\text{DF}})$ by the term $(\lambda\, c/b)^{\tilde{r}-d-1}$. According to (7), $\lambda\, c/b \ll 1$, such that, for fixed values of $\tilde{r}$ and u, increasing d causes these parameters to increase. Therefore, by virtue of (23) and (34), increasing d causes MTTDL to decrease and EAFDL to increase. Consequently, for fixed values of m and l, deferring rebuilds degrades reliability.

D. *Equivalent Systems*

We call *equivalent systems* those that employ a given codeword length m and have the same number $m - l - d$ of exposure levels at which the rebuild process is active. In this case, it holds that $l + d = z$, and from (3) and (4), it follows that

$$0 \le d < z < m \quad \text{and} \quad d+1 \le u \le m - z + d + 1 \;. \tag{39}$$

Next, we compare the MTTDL and EAFDL of equivalent systems. For $P_s = 0$, substituting (12), (19), and (22) into (23) yields

$$\frac{\text{MTTDL}(d+1)}{\text{MTTDL}(d)} \approx \frac{E(T|d+1)}{E(T|d)} \cdot \frac{1}{\prod_{u=d+1}^{m-z+d} V_u} \cdot \frac{\prod_{i=d+1}^{m-z+d} \frac{\tilde{n}_i(d)}{b_i(d)}}{\prod_{i=d+2}^{m-z+d+1} \frac{\tilde{n}_i(d+1)}{b_i(d+1)}} \;. \tag{40}$$

From (12), it follows that $E(T|d+1) > E(T|d)$. Also, V_u represent fractions, which implies that $V_u \le 1$. Consequently, the product of the first two terms of (40) is greater than 1.

For a symmetric placement scheme that is not bandwidth constrained, it follows from (36) that $\tilde{n}_u(d)/b_u(d) = (l+1)/b = (z-d+1)/b$. Substituting this into (40) yields

$$\frac{\text{MTTDL}(d+1)}{\text{MTTDL}(d)} \approx \frac{E(T|d+1)}{E(T|d)} \cdot \prod_{u=d+1}^{m-z+d} V_u^{-1} \cdot \left(\frac{z-d+1}{z-d}\right)^{m-z} > 1 \;. \tag{41}$$

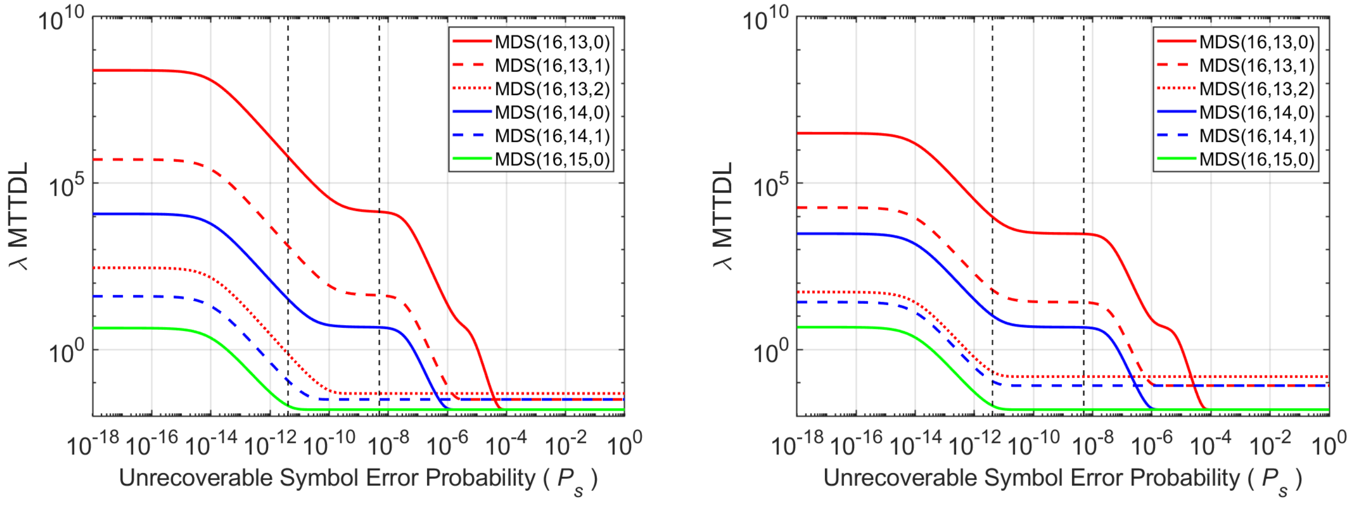
For a clustered placement scheme that is not bandwidth constrained, it follows from (38) that $\tilde{n}_u(d)/b_u(d) = (m-u)/b$. Substituting this into (40), and using (38), yields

$$\frac{\text{MTTDL}(d+1)}{\text{MTTDL}(d)} \approx \frac{E(T|d+1)}{E(T|d)} \cdot \frac{m-d-1}{z-d-1} > 1 \;. \tag{42}$$

Similarly, from (32) it follows that

$$\frac{\frac{E(Q_{\text{UF}_{u+1}}|d+1)}{l_{d+1}\, E(T|d+1)}}{\frac{E(Q_{\text{UF}_u}|d)}{l_d\, E(T|d)}} \approx \frac{E(T|d)}{E(T|d+1)} \cdot \frac{\tilde{r}(d+1)}{\tilde{r}(d)} \cdot \prod_{i=d+1}^{u} V_i \cdot \frac{z-d-1}{m-u} \cdot A \,, \tag{43}$$

$$\text{where}\quad A = \begin{cases} \left(\frac{z-d}{z-d+1}\right)^{u-d-1} , & \text{for symmetric placement} \\ \frac{m-u}{m-d-1} \,, & \text{for clustered placement} \;. \end{cases} \tag{44}$$



(a) $k = 64$ (declustered data placement scheme)

(b) $k = 16$ (clustered data placement scheme)

Figure 1. Normalized MTTDL vs. P_s for various MDS(m, l, d) codes ; $n = 64$, $\lambda/\mu = 0.0002$, $c = 12$ TB, and $s = 512$ B.

TABLE II. TYPICAL VALUES OF DIFFERENT PARAMETERS

Parameter	Definition	Values
n	number of storage devices	64
c	amount of data stored on each device	12 TB
l	user-data symbols per codeword	13, 14, 15
m	symbols per codeword	16
s	symbol (sector) size	512 B
b	rebuild bandwidth per device	50 MB/s
λ^{-1}	mean time to failure of a storage device	**300,000** h to 1,000,000 h
U	amount of user data stored in the system	624 to 720 TB
μ^{-1}	time to read an amount c of data at a rate b from a storage device	66.7 h

It can be shown that $\frac{E(T|d)}{E(T|d+1)} \cdot \frac{\tilde{r}(d+1)}{\tilde{r}(d)} < 1$. Consequently, from (34), (39), and (43), and recognizing that $A < 1$ and $E(Q_{\text{DF}}) = E(Q_{\text{UF}_{\tilde{r}}})$, it follows that

$$\frac{\text{EAFDL}(d+1)}{\text{EAFDL}(d)} < 1 \; . \tag{45}$$

Remark 2: Within the class of equivalent systems, according to (42) and (45), deferring rebuilds improves reliability, despite the fact that rebuilds are performed at the same number of exposure levels. This is because increasing d amounts to decreasing l, and therefore at a reduced number of symbols read at each exposure level. This in turn results in reduced vulnerability window and therefore improved reliability.

IV. NUMERICAL RESULTS

Here, we assess the reliability of the clustered and declustered schemes for a system comprised of $n = 64$ devices (disks), where each device stores an amount $c = 12$ TB, and $m = 16$, $l = 13$, 14, and 15, and the symbol size s is equal to a sector size of 512 bytes.

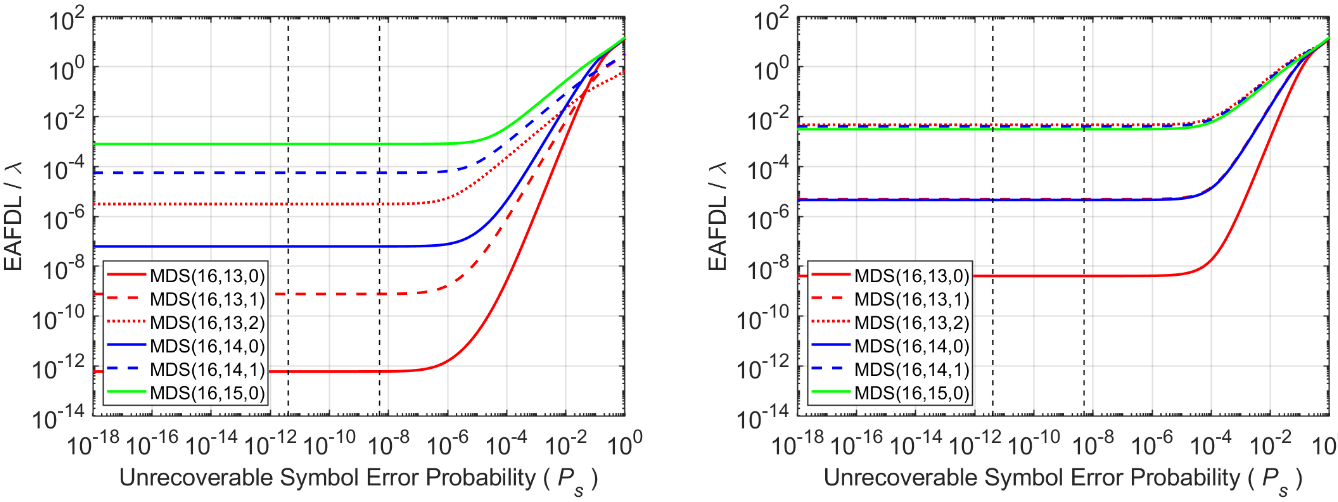
Typical parameter values are listed in Table II. The Annualized Failure Rate (AFR) is in the range of 0.9% to 3%, which corresponds to a mean time to failure in the range of $300,000$ h to $1,000,000$ h. The parameter λ^{-1} is chosen to be equal to $300,000$ h. It is assumed that the reserved rebuild bandwidth b is equal to 50 MB/s, which yields a rebuild time of a device $\mu^{-1} = c/b = 66.7$ h, and that the network rebuild bandwidth is sufficiently large ($B_{\text{max}} \geq n\,b = 3.2$ GB/s). We assume that the rebuild time distribution is deterministic, such that $E(X^k) = [E(X)]^k$. The obtained results are accurate because (7) is satisfied, given that $\lambda/\mu = 2.2 \times 10^{-4} \ll 1$.

First, we assess the reliability for the declustered placement scheme ($k = n = 64$) for various MDS-coded configurations with $m = 16$ and varying values of l and d. These configurations are denoted by MDS(m,l,d) and the corresponding results are shown in Figures 1, 2, and 3 by solid lines for $d = 0$ (no lazy rebuild employed), dashed lines for $d = 1$ and dotted lines for $d = 2$. Six configurations are considered: MDS(16,13,0), MDS(16,13,1), MDS(16,13,2), MDS(16,14,0), MDS(16,14,1), and MDS(16,15,0), for each of the declustered and clustered data placement schemes. In particular, for the clustered placement scheme, the MDS(16,15,0) and MDS(16,14,0) configurations correspond to the RAID-5 and RAID-6 systems.

The normalized $\lambda\,$MTTDL measure is obtained from (13) as a function of P_s and shown in Figure 1(a) for the declustered data placement scheme. We observe that MTTDL decreases monotonically with P_s and exhibits $m - l - d$ plateaus. In the interval $[4.096 \times 10^{-12}, 5 \times 10^{-9}]$ of practical importance for P_s, which is indicated between the two vertical dashed lines, MTTDL is degraded by orders of magnitude. Increasing the number of parities (reducing l) improves reliability by orders of magnitude. By contrast, and according to Remark 1, employing lazy rebuild degrades reliability by orders of magnitude. Moreover, for equivalent systems, such as MDS(16,15,0), MDS(16,14,1) and MDS(16,13,2), and according to Remark 2, MTTDL increases as d increases.

The normalized $\lambda\,$MTTDL measure for the clustered data placement scheme is shown in Figure 1(b). We observe that the declustered placement scheme achieves a significantly higher MTTDL than the clustered one.

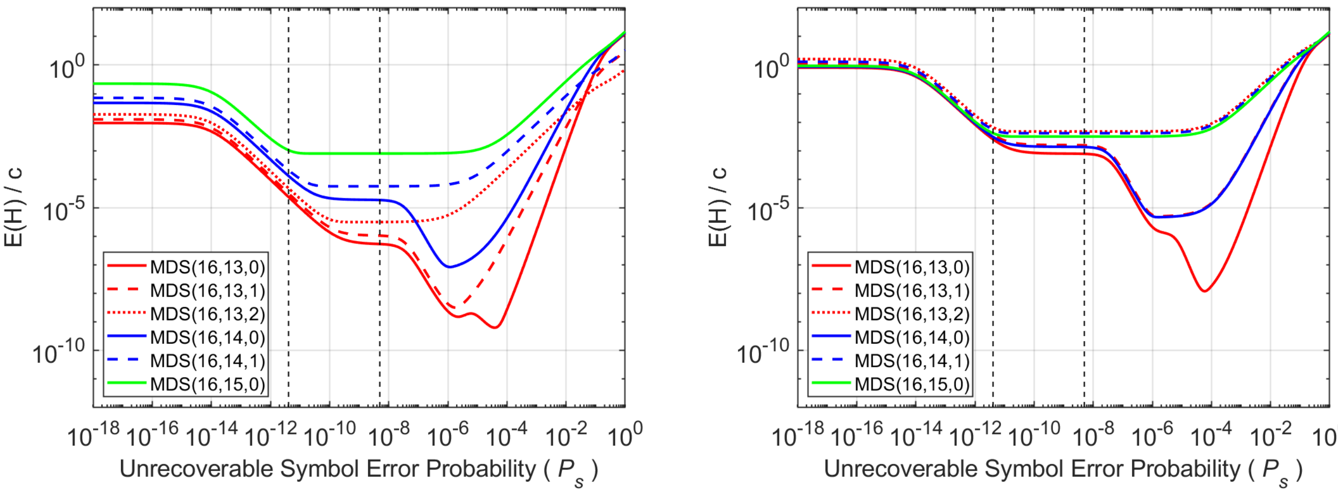
The normalized EAFDL$/\lambda$ measure is obtained from (14) and shown in Figure 2. We observe that EAFDL increases monotonically, but it is practically unaffected in the interval of interest because it degrades only when P_s is much larger than the typical sector error probabilities. For the EAFDL



(a) $k = 64$ (declustered data placement scheme)

(b) $k = 16$ (clustered data placement scheme)

Figure 2. Normalized EAFDL vs. P_s for various MDS(m, l, d) codes ; $n = 64$, $\lambda/\mu = 0.0002$, $c = 12$ TB, and $s = 512$ B.



(a) $k = 64$ (declustered data placement scheme)

(b) $k = 16$ (clustered data placement scheme)

Figure 3. Normalized $E(H)$ vs. P_s for various MDS(m, l, d) codes ; $n = 64$, $\lambda/\mu = 0.0002$, $c = 12$ TB, and $s = 512$ B.

metric too, increasing the number of parities (reducing l) results in a reliability improvement by orders of magnitude. By contrast, employing lazy rebuild degrades reliability by orders of magnitude. Moreover, for equivalent systems, such as MDS(16,15,0), MDS(16,14,1) and MDS(16,13,2), and according to Remark 2, EAFDL decreases as d increases, but for the clustered placement scheme it is not significantly affected. We observe that for both MTTDL and EAFDL reliability metrics, the reliability level achieved by the declustered data placement scheme is higher than that of the clustered one.

The normalized expected amount $E(H)/c$ of lost user data, given that a data loss has occurred, relative to the amount of data stored in a device is obtained from (15) and shown in Figure 3. In contrast to the P_{DL}, EAFDL, and $E(Q)$ metrics that increase monotonically with P_s, we observe that $E(H)$ does not do so. The reason for that is the following. For $P_s \gg 10^{-14}$, data loss is more likely to be due to sector errors than to device failures. Given that sector errors result in a negligible amount of data loss compared with the substantial data losses caused by device failures, when P_s increases over the value of 10^{-14}, the conditional amount of lost data decreases. Clearly, this is reversed for high values of P_s, and the conditional amount of lost data increases.

Also, in the interval $[4.096 \times 10^{-12}, 5 \times 10^{-9}]$ of practical importance for P_s, and by contrast to MTTDL and EAFDL, employing lazy rebuild does not affect $E(H)$ significantly. Moreover, for equivalent systems, such as MDS(16,15,0), MDS(16,14,1) and MDS(16,13,2), and for higher values of d, $E(H)$ is lower for the declustered data placement scheme, but it is not significantly affected for the clustered one.

V. CONCLUSIONS

The effect of the lazy rebuild scheme on the reliability of erasure-coded data storage systems was investigated. A methodology was developed for deriving the Mean Time to Data Loss (MTTDL) and the Expected Annual Fraction of

Data Loss (EAFDL) reliability metrics analytically. Closed-form expressions capturing the effect of unrecoverable latent errors were obtained for the symmetric, clustered and declustered data placement schemes. We demonstrated that system reliability is significantly degraded by the employment of the lazy rebuild scheme. We also demonstrated that the declustered placement scheme offers superior reliability in terms of both metrics. We established that, for realistic unrecoverable sector error rates, MTTDL is adversely affected by the presence of latent errors, whereas EAFDL is not. The analytical reliability expressions derived here can identify lazy rebuild schemes that reduce the volumes of repair traffic and at the same time ensure a desired level of reliability.

Applying these results to assess the effect of network rebuild bandwidth constraints is a subject of further investigation. The reliability evaluation of erasure-coded systems when device failures, as well as unrecoverable latent errors are correlated is also part of future work.

APPENDIX A

Proof of Proposition 1.

We consider the direct path $\overrightarrow{UF_u} = 1 \to 2 \to \cdots \to u \to \text{UF}$ and proceed to evaluate $P_{\text{UF}_u}(R_{d+1}, \vec{\alpha}_{u-1})$, the probability of entering exposure level u through vector $\vec{\alpha}_{u-1} \triangleq (\alpha_1, \ldots, \alpha_{u-1})$ and given a rebuild time R_{d+1}, and then encountering an unrecoverable failure during the rebuild process at this exposure level. It follows from (16) that

$$P_{\text{UF}_u}(R_{d+1}, \vec{\alpha}_{u-1}) = P_u(R_{d+1}, \vec{\alpha}_{u-2}) \cdot P_{u \to \text{UF}}(R_{d+1}, \vec{\alpha}_{u-1}) \, . \tag{46}$$

To evaluate the above product, we first establish the following lemma.

LEMMA 1: For $u = d+1, \ldots, \tilde{r}-1$, it holds that

$$P_u(R_{d+1}, \vec{\alpha}_{u-2}) \approx (\lambda b_{d+1} R_{d+1})^{u-d-1} \prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} \, (V_i \, \alpha_i)^{u-1-i} \tag{47}$$

with the convention that for any integer j and for any sequence δ_i , $\prod_{i=j}^{0} \delta_i \triangleq 1$.

Proof: As the rebuild times are proportional to the amount of data to be rebuilt and are inversely proportional to the rebuild rates, it holds that

$$\frac{R_{d+1}}{X} = \frac{C_{d+1}}{C} \, \frac{b}{b_{d+1}} \, , \tag{48}$$

Using (8) and (10), (48) yields

$$R_{d+1} \approx \left(\prod_{j=1}^{d} V_j \right) \frac{b}{b_{d+1}} \, X \, , \tag{49}$$

Also,

$$\frac{R_{u+1}}{R_u} = \frac{C_{u+1}}{C_u} \, \frac{b_u}{b_{u+1}} \, , \quad \text{for} \quad u = d+1, \ldots, \tilde{r}-2 \, . \tag{50}$$

Combining (9) and (50) yields

$$R_{u+1} \approx V_u \, \alpha_u \, \frac{b_u}{b_{u+1}} \, R_u \, , \quad \text{for} \quad u = d+1, \ldots, \tilde{r}-2 \, . \tag{51}$$

Repeatedly applying (51) yields

$$R_u \approx \frac{b_{d+1}}{b_u} \, R_{d+1} \prod_{j=d+1}^{u-1} V_j \, \alpha_j \, , \quad u = d+1, \ldots, \tilde{r}-1 \, . \tag{52}$$

We denote by $\tilde{n}_u$ the number of devices at exposure level u whose failure causes an exposure level transition to level $u+1$. Subsequently, the transition probability $P_{u \to u+1}$ from exposure level u to $u+1$ depends on the duration of the corresponding rebuild time R_u and the aggregate failure rate of these $\tilde{n}_u$ highly reliable devices, and is given by [19]

$$P_{u \to u+1} \approx \tilde{n}_u \, \lambda \, R_u \, , \quad \text{for } u = d+1, \ldots, \tilde{r}-1 \, . \tag{53}$$

Substituting (52) into (53) yields

$$P_{u \to u+1}(R_{d+1}, \vec{\alpha}_{u-1}) \approx \tilde{n}_u \, \lambda \, \frac{b_{d+1}}{b_u} \, R_{d+1} \prod_{j=d+1}^{u-1} V_j \, \alpha_j \, . \tag{54}$$

The probability P_u of entering exposure level u can be approximated by the probability of the direct path $d+1 \to d+2 \to \cdots \to u$ of successive transitions from exposure level $d+1$ to u, that is,

$$P_u \approx \prod_{i=d+1}^{u-1} P_{i \to i+1}, \, , \quad \text{for } u = d+2, \ldots, \tilde{r} \, . \tag{55}$$

Substituting (54) into (55), and using the fact that $P_{d+1} = 1$, yields (47). ■

Given that the elements of $\vec{\alpha}_{u-2}$ are independent random variables approximately distributed according to (8), such that $E(\alpha_i^k) \approx 1/(k+1)$ for $i \geq d+1$, we have

$$\begin{aligned} E\left(\prod_{i=d+1}^{u-1} \alpha_i^{u-1-i} \right) &= \prod_{i=d+1}^{u-1} E(\alpha_i^{u-1-i}) \\ &\approx \prod_{i=d+1}^{u-1} \frac{1}{u-i} = \frac{1}{(u-d-1)!} \, . \end{aligned} \tag{56}$$

Unconditioning (47) on $\vec{\alpha}_{u-2}$ and using (56) yields

$$P_u(R_{d+1}) \approx \frac{(\lambda b_{d+1} R_{d+1})^{u-d-1}}{(u-d-1)!} \prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} \, V_i^{u-1-i} \, . \tag{57}$$

Unconditioning (57) on R_{d+1} and using (6) and (49) yields (17).

We now proceed to calculate $P_{u \to \text{UF}}(R_{d+1}, \vec{\alpha}_{u-1})$. Upon entering exposure level u, the rebuild process attempts to restore the C_u most-exposed codewords, each of which has $m-u$ remaining symbols. The probability q_u that a codeword can be restored is determined by (21), which is Equation (16) of [9]. Note that, if a codeword is corrupted, then at least one of its l user-data symbols is lost. Owing to the independence of symbol errors, codewords are independently corrupted. Consequently, the conditional probability $P_{\text{UF}|C_u}$ of encountering an unrecoverable failure during the rebuild process of the C_u codewords is

$$P_{\text{UF}|C_u} = 1 - q_u^{C_u} \, , \quad \text{for } u = d+1, \ldots, \tilde{r} \, . \tag{58}$$

Substituting (10) into (58) and using (20) yields

$$P_{u \to \text{UF}}(R_{d+1}, \vec{\alpha}_{u-1}) \approx 1 - q_u^{C \, \prod_{j=1}^{u-1} V_j \, \alpha_j} = 1 - \hat{q}_u^{\prod_{j=1}^{u-1} \alpha_j} \, . \tag{59}$$

Substituting (59) into (46) yields

$$P_{\mathrm{UF}_u}(R_{d+1}, \vec{\alpha}_{u-1}) \;\approx\; P_u(R_{d+1}, \vec{\alpha}_{u-2}) \left[1 - \hat{q}_u^{\,\prod_{j=1}^{u-1} \alpha_j} \right] . \qquad (60)$$

Unconditioning (60) on $\vec{\alpha}_{u-1}$, and using (8) and (47), yields

$$P_{\mathrm{UF}_u}(R_{d+1}) \;\approx\; P_u(R_{d+1}) - (\lambda b_{d+1} R_{d+1})^{u-d-1} \cdot \left(\prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} \, V_i^{u-1-i} \right) E_{\vec{\alpha}_{u-1}} \left[\left(\prod_{i=d+1}^{u-1} \alpha_i^{u-1-i} \right) \hat{q}_u^{\,\prod_{j=d+1}^{u-1} \alpha_j} \right] . \qquad (61)$$

LEMMA 2: For α_i distributed according to (8) it holds that

$$E \left[\left(\prod_{i=d+1}^{u-1} \alpha_i^{u-1-i} \right) q^{\,\prod_{i=d+1}^{u-1} \alpha_i} \right] = \frac{1}{(u-d-1)!} + \log(q)^{-(u-d-1)} \left(q - \sum_{i=0}^{u-d-1} \frac{\log(q)^i}{i!} \right) . \qquad (62)$$

Proof: It holds that

$$q^{\,\prod_{i=d+1}^{u-1} \alpha_i} = e^{\log(q) \, \prod_{i=d+1}^{u-1} \alpha_i} = \sum_{j=0}^{\infty} \frac{\log(q)^j \, (\prod_{i=d+1}^{u-1} \alpha_i)^j}{j!} \, , \qquad (63)$$

which implies that

$$\begin{aligned} \left(\prod_{i=d+1}^{u-1} \alpha_i^{u-1-i} \right) q^{\,\prod_{i=d+1}^{u-1} \alpha_i} &= \left(\prod_{i=d+1}^{u-1} \alpha_i^{u-1-i} \right) \left(\sum_{j=0}^{\infty} \frac{\log(q)^j \, (\prod_{i=d+1}^{u-1} \alpha_i)^j}{j!} \right) \\ &= \sum_{j=0}^{\infty} \frac{\log(q)^j \, \prod_{i=d+1}^{u-1} \alpha_i^{u-1-i+j}}{j!} \, . \qquad (64) \end{aligned}$$

Consequently,

$$\begin{aligned} & E \left[\left(\prod_{i=d+1}^{u-1} \alpha_i^{u-1-i} \right) q^{\,\prod_{i=d+1}^{u-1} \alpha_i} \right] - \frac{1}{(u-d-1)!} \\ &= \sum_{j=0}^{\infty} \frac{\log(q)^j \, \prod_{i=d+1}^{u-1} E(\alpha_i^{u-1-i+j})}{j!} - \frac{1}{(u-d-1)!} \\ &\approx \sum_{j=0}^{\infty} \frac{\log(q)^j \, \prod_{i=d+1}^{u-1} E(\alpha_i^{u-1-i+j})}{j!} - \frac{1}{(u-d-1)!} \\ &\approx \sum_{j=0}^{\infty} \frac{\log(q)^j \, \prod_{i=d+1}^{u-1} \frac{1}{u-i+j}}{j!} - \frac{1}{(u-d-1)!} \\ &= \sum_{j=0}^{\infty} \frac{\log(q)^j}{(u-d-1+j)!} - \frac{1}{(u-d-1)!} \\ &= \sum_{j=1}^{\infty} \frac{\log(q)^j}{(u-d-1+j)!} \\ &= \log(q)^{-(u-d-1)} \sum_{i=u-d}^{\infty} \frac{\log(q)^i}{i!} \\ &= \log(q)^{-(u-d-1)} \left(\sum_{i=0}^{\infty} \frac{\log(q)^i}{i!} - \sum_{i=0}^{u-d-1} \frac{\log(q)^i}{i!} \right) \\ &= \log(q)^{-(u-d-1)} \left(e^{\log(q)} - \sum_{i=0}^{u-d-1} \frac{\log(q)^i}{i!} \right) . \qquad (65) \end{aligned}$$

■

From (57) and (62), (61) yields

$$P_{\mathrm{UF}_u}(R_{d+1}) \;\approx\; -(\lambda b_{d+1} R_{d+1})^{u-d-1} \left(\prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} \, V_i^{u-1-i} \right) \cdot \log(\hat{q}_u)^{-(u-d-1)} \left(\hat{q}_u - \sum_{i=0}^{u-d-1} \frac{\log(\hat{q}_u)^i}{i!} \right) . \qquad (66)$$

Unconditioning (66) on R_{d+1}, and using (6) and (49), yields (19).

☐

APPENDIX B

Proof of Proposition 2.

The expected number $E(S_{\mathrm{U}} | C_u)$ of symbols lost due to unrecoverable failures during the rebuild of the C_u codewords at exposure level u is determined by Equation (53) of [9]:

$$E(S_{\mathrm{U}} | C_u) \;\approx\; C_u \, \tilde{r} \binom{m-u}{\tilde{r}-u} P_s^{\tilde{r}-u}, \;\; P_s \ll \frac{1}{m - \tilde{r}} \, . \qquad (67)$$

Substituting (10) into (67) yields

$$E(S_{\mathrm{U}} | \vec{\alpha}_{u-1}) \;\approx\; C \left(\prod_{j=1}^{u-1} V_j \, \alpha_j \right) \tilde{r} \binom{m-u}{\tilde{r}-u} P_s^{\tilde{r}-u} \, . \qquad (68)$$

Subsequently, the expected number $E(S_{\mathrm{UF}_u} | R_{d+1}, \vec{\alpha}_{u-1})$ of symbols lost due to unrecoverable failures encountered during rebuild in conjunction with entering exposure level u through vector $\vec{\alpha}_{u-1}$, and given a rebuild time R_{d+1}, is

$$E(S_{\mathrm{UF}_u} | R_{d+1}, \vec{\alpha}_{u-1}) = P_u(R_{d+1}, \vec{\alpha}_{u-1}) \, E(S_{\mathrm{U}} | \vec{\alpha}_{u-1}) \, . \qquad (69)$$

Substituting (47) and (68) into (69) yields

$$E(S_{\mathrm{UF}_u} | R_{d+1}, \vec{\alpha}_{u-1}) \approx (\lambda b_{d+1} R_{d+1})^{u-d-1} \left[\prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} \, (V_i \, \alpha_i)^{u-i} \right] \cdot \; C \left(\prod_{j=1}^{d} V_j \right) \tilde{r} \binom{m-u}{\tilde{r}-u} P_s^{\tilde{r}-u}, \quad P_s \ll \frac{1}{m - \tilde{r}} \, . \qquad (70)$$

Unconditioning (70) on $\vec{\alpha}_{u-1}$, and given that (56) implies that $E(\prod_{i=d+1}^{u-1} \alpha_i^{u-i}) = 1/(u-d)!$, yields

$$E(S_{\mathrm{UF}_u} | R_{d+1}) \approx \; (\lambda b_{d+1} R_{d+1})^{u-d-1} \left(\prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} \, V_i^{u-i} \right) \frac{1}{(u-d)!} \cdot \; C \left(\prod_{j=1}^{d} V_j \right) \tilde{r} \binom{m-u}{\tilde{r}-u} P_s^{\tilde{r}-u}, \quad P_s \ll \frac{1}{m - \tilde{r}} \, . \qquad (71)$$

Unconditioning (71) on R_{d+1}, and using (6) and (49), yields

$$E(S_{\mathrm{UF}_u}) \approx \left(\lambda \, c \prod_{j=1}^{d} V_j \right)^{u-d-1} \frac{E(X^{u-d-1})}{[E(X)]^{u-d-1}} \left(\prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} \, V_i^{u-i} \right) \cdot \frac{1}{(u-d)!} \; C \left(\prod_{j=1}^{d} V_j \right) \tilde{r} \binom{m-u}{\tilde{r}-u} P_s^{\tilde{r}-u}, \quad P_s \ll \frac{1}{m - \tilde{r}} \, . \qquad (72)$$

Substituting (72) into (31) yields (32).

☐

REFERENCES

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

[2] P. M. Chen, E. K. Lee, G. A. Gibson, R. H. Katz, and D. A. Patterson, “RAID: High-Performance, reliable secondary storage,” ACM Comput. Surv., vol. 26, no. 2, Jun. 1994, pp. 145–185.

[3] 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), Jul. 2011, pp. 307–317.

[4] I. Iliadis, D. Sotnikov, P. Ta-Shma, and V. Venkatesan, “Reliability of geo-replicated cloud storage systems,” in Proc. 2014 IEEE 20th Pacific Rim Int’l Symposium on Dependable Computing (PRDC), Nov. 2014, pp. 169–179.

[5] D. Ford et al., “Availability in globally distributed storage systems,” in Proc. 9th USENIX Symposium on Operating Systems Design and Implementation (OSDI), Oct. 2010, pp. 61–74.

[6] A. G. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, “A survey on network coding for distributed storage,” Proc. IEEE, vol. 99, no. 3, Mar. 2011, pp. 476–489.

[7] C. Huang et al., “Erasure coding in Windows Azure Storage,” in Proc. USENIX Annual Technical Conference (ATC), Jun. 2012, pp. 15–26.

[8] S. Muralidhar et al., “f4: Facebook’s Warm BLOB Storage System,” in Proc. 11th USENIX Symposium on Operating Systems Design and Implementation (OSDI), Oct. 2014, pp. 383–397.

[9] I. Iliadis, “Reliability assessment of erasure-coded storage systems with latent errors,” in Proc. 14th Int’l Conference on Communication Theory, Reliability, and Quality of Service (CTRQ), Apr. 2021, pp. 15–24.

[10] 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. Storage, vol. 4, no. 1, 2008, pp. 1–42.

[11] I. Iliadis, “Reliability modeling of RAID storage systems with latent errors,” in Proc. 17th Annual IEEE/ACM Int’l Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Sep. 2009, pp. 111–122.

[12] V. Venkatesan and I. Iliadis, “Effect of latent errors on the reliability of data storage systems,” in Proc. 21th Annual IEEE Int’l Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Aug. 2013, pp. 293–297.

[13] K. V. Rashmi et al., “A ”Hitchhiker’s” guide to fast and efficient data reconstruction in erasure-coded data centers,” in Proc. 2014 ACM conference on SIGCOMM, Aug. 2014, pp. 331–342.

[14] M. Silberstein, L. Ganesh, Y. Wang, L. Alvisi, and M. Dahlin, “Lazy means smart: Reducing repair bandwidth costs in erasure-coded distributed storage,” in Proc. 7th ACM Int’l Systems and Storage Conference (SYSTOR), Jun. 2014, pp. 15:1–15:7.

[15] I. Iliadis and V. Venkatesan, “Expected annual fraction of data loss as a metric for data storage reliability,” in Proc. 22nd Annual IEEE Int’l Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Sep. 2014, pp. 375–384.

[16] ——, “Reliability evaluation of erasure coded systems,” Int’l J. Adv. Telecommun., vol. 10, no. 3&4, Dec. 2017, pp. 118–144.

[17] I. Iliadis, “Reliability evaluation of erasure coded systems under rebuild bandwidth constraints,” Int’l J. Adv. Networks and Services, vol. 11, no. 3&4, Dec. 2018, pp. 113–142.

[18] V. Venkatesan, I. Iliadis, and R. Haas, “Reliability of data storage systems under network rebuild bandwidth constraints,” in Proc. 20th Annual IEEE Int’l Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Aug. 2012, pp. 189–197.

[19] V. Venkatesan and I. Iliadis, “A general reliability model for data storage systems,” in Proc. 9th Int’l Conference on Quantitative Evaluation of Systems (QEST), Sep. 2012, pp. 209–219.

[20] DELL/EMC Whitepaper, ”PowerVault ME4 Series ADAPT Software,” Feb. 2019. [Online]. Available: https://www.dellemc.com/ [retrieved: March, 2022]

[21] I. Iliadis, R. Haas, X.-Y. Hu, and E. Eleftheriou, “Disk scrubbing versus intradisk redundancy for RAID storage systems,” ACM Trans. Storage, vol. 7, no. 2, 2011, pp. 1–42.

[22] V. Venkatesan and I. Iliadis, “Effect of codeword placement on the reliability of erasure coded data storage systems,” in Proc. 10th Int’l Conference on Quantitative Evaluation of Systems (QEST), Sep. 2013, pp. 241–257.

[23] ——, “Effect of codeword placement on the reliability of erasure coded data storage systems,” IBM Research Report, RZ 3827, Aug. 2012.

[24] I. Iliadis and X.-Y. Hu, “Reliability assurance of RAID storage systems for a wide range of latent sector errors,” in Proc. 2008 IEEE Int’l Conference on Networking, Architecture, and Storage (NAS), Jun. 2008, pp. 10–19.

[25] I. Iliadis and V. Venkatesan, “Most probable paths to data loss: An efficient method for reliability evaluation of data storage systems,” Int’l J. Adv. Syst. Measur., vol. 8, no. 3&4, Dec. 2015, pp. 178–200.