Expected Annual Fraction of Entity Loss as a Metric for Data Storage Durability

Ilias Iliadis

IBM Research Europe – Zurich

8803 Rüschlikon, Switzerland

email: ili@zurich.ibm.com

Abstract—Data protection schemes are designed into storage systems to ensure data durability and accessibility in the presence of device failures. The effectiveness of these schemes has been evaluated based on the Mean Time to Data Loss (MTTDL) and the Expected Annual Fraction of Data Loss (EAFDL) metrics. The EAFDL metric assesses data loss at a low data processing unit level whereas durability refers to losses at a higher entity, say file or object, level. To evaluate the durability of storage systems we introduce the following reliability metric: the Expected Annual Fraction of Entity Loss (EAFEL), that is, the fraction of entities that is expected to be lost by the system annually. The general methodology that was applied to assess the MTTDL and EAFDL metrics is extended to obtain the EAFEL metric analytically for erasure-coding redundancy schemes and for the clustered, declustered, and symmetric data placement schemes. The theoretical model developed considers the effects of device failures, latent errors, and lazy rebuilds. For realistic values of sector error rates, the results obtained demonstrate that MTTDL and EAFEL degrade, but the EAFEL degradation is more pronounced when entities are large. It is also shown that the declustered data placement scheme offers superior reliability.

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

I. INTRODUCTION

Data protection schemes are designed into storage systems to ensure data durability and accessibility in the presence of device and component failures [1-8]. The replication schemes and the Redundant Arrays of Inexpensive Disks (RAID) schemes, such as RAID-5 and RAID-6, which have been employed extensively in the past thirty years [1-4] are special cases of erasure codes. The reliability of storage systems is adversely affected by the latent or unrecoverable sector errors that are discovered when there is an attempt to access these sectors. Permanent losses of data due to latent errors are quite pronounced in higher-capacity hard-disk drives (HDDs) and storage nodes [9-11]. To cope with the repair problem due to the increased network traffic needed to repair data lost due to device failures a lazy rebuild scheme was proposed in [12] to reduce the amount of data transmitted during rebuilds. The reliability level achieved by this scheme was assessed in [13].

The Mean Time to Data Loss (MTTDL) metric has been widely used to assess the reliability of storage systems [14][15]. For a comprehensive reliability evaluation, it is imperative to assess not only the frequency of data loss events, which is obtained by the MTTDL metric, but also the amount of data loss, which is captured by the Expected Annual Fraction of Data Loss (EAFDL) metric [16-21]. The EAFDL metric is meant to complement, not to replace, the traditional MTTDL metric, as these two metrics provide a useful profile of the magnitude and frequency of data losses.

The MTTDL and EAFDL metrics have been evaluated in parallel in a common general theoretical framework using the *direct-path-approximation* methodology presented in [3, 22, 23], which, for highly reliable storage devices, enables the derivation of analytic expressions that accurately assess the reliability metrics of interest [16][22][24]. This methodology does not involve Markovian analysis and holds for general device failure time distributions, which can be exponential or non-exponential, such as Weibull and gamma.

The EAFDL metric was introduced in [16], because Amazon S3 [25], Facebook [26], LinkedIn [27] and Yahoo! [28] consider the amount of lost data measured in time. However, system durability reflects losses at an entity, say file or object, level whereas the EAFDL metric assesses data losses at a lower data processing unit level. For this reason, in this article we introduce the following reliability metric: the Expected Annual Fraction of Entity Loss (EAFEL), that is, the fraction of entities that is expected to be lost by the system annually.

The key contributions of this article are the following. The general non-Markovian methodology that was applied in prior work to assess the MTTDL and EAFDL metrics is extended to obtain the new EAFEL metric analytically for erasure-coding redundancy schemes and for the clustered, declustered, and symmetric data placement schemes. The validity of this methodology for accurately assessing the reliability of storage systems has been confirmed by simulations in several contexts [3][16][22][29]. 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 the entity size on the reliability of storage systems. The theoretical model developed considers the effects of device failures, latent errors, lazy rebuilds, and can also be used to assess system reliability when scrubbing is employed [14]. 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 EAFEL 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

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) ($m \leq k \leq n$)
J	number of codewords per entity ($J \geq 1$)
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 ($s_{\mathrm{eff}} = l/m$)
U	amount of user data stored in the system ($U = s_{\mathrm{eff}} \, n \, c$)
$\tilde{r}$	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 ($C = c/s$)
N_E	number of entities ($N_E = (n\,C)/(m\,J)$)
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
Y	number of lost entities during rebuild
G	number of lost entities, given that data loss has occurred during rebuild

the effectiveness of erasure coding schemes for improving system reliability as well as the adverse effect of increased entity sizes. 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 [13] and extend it to cover the case of entity 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 contained in entities is divided into blocks (or symbols) of a fixed size s (e.g., sector size of 512 bytes) and complemented with parity symbols to form codewords, as shown in Figure 1. 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

$$s_{\mathrm{eff}} = l/m \quad \text{and} \quad U = s_{\mathrm{eff}} \, n \, c \ = \ l \, n \, c/m \,. \quad (1)$$

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

$$C = c/s \,, \quad (2)$$

and, therefore, the number of symbols stored in the system is $n\,C$ and the number of codewords stored in the system is $n\,C/m$. Assuming fixed-size entities with each one containing J codewords, an entity contains $J\,l$ user-data symbols and

$J\,(m - l)$ parity symbols for a total of $J\,m$ symbols. Consequently, the entity size (including parity symbols) is $J\,m\,s$ and the number N_E of entities in the system is

$$N_E = \frac{n}{m} \cdot \frac{C}{J} \stackrel{(2)}{=} \frac{n}{m} \cdot \frac{c}{J\,s} \,. \quad (3)$$

Our notation is summarized in Table I. The derived parameters are listed in the lower part of the table. Two different ways (A and B) for storing the $J\,l$ user-data symbols of an entity in J codewords are shown in Figure 1. Note that $S_{j,i}$ denotes the ith symbol of the jth codeword. Thus, $S_{1,2}$, which is the second symbol of codeword C-1, in the case of A is the second symbol of the first chunk, whereas in the case of B is the first symbol of the second shard. In the case of A, successive symbols of a chunk are stored on different devices whereas in the case of B, successive symbols of a shard are stored on the same device. To minimize the risk of permanent data loss, the m symbols of each of the J codewords are spread and stored successively in a set of m devices that constitute a placement group. This way, the system can tolerate any $\tilde{r} - 1$ device failures, but $\tilde{r}$ device failures may lead to data loss, with

$$\tilde{r} = m - l + 1 \,, \quad 1 \leq l < m \quad \text{and} \quad 2 \leq \tilde{r} \leq m \,. \quad (4)$$

Examples of MDS erasure codes are the replication, RAID-5, and RAID-6 schemes. In terms of encoding operations, MDS erasure codes are either bitwise exclusive-OR (XOR) or non-XOR. The computation complexity of the non-XOR-based codes, such as Reed–Solomon, is much higher than that of the XOR-based ones. Additional categories of erasure codes are the Minimum Bandwidth Regenerating (MBR) codes that aim to reduce the amount of data transferred over the storage network during rebuilds, and the Minimum Storage Regenerating (MSR) codes that aim to reduce storage overheads [30].

The system comprises n/k groups of k devices. These groups are disjoint in that codewords are stored within, but not across groups. Within each group, the number of placement groups is at most $\binom{k}{m}$, which is the number of possible ways of placing m symbols across k devices. Successive entities, such as Entity-1 and Entity-2 shown in Figure 1, may be stored in different placement groups.

Data is stored according to symmetric placement schemes, including the *clustered* and *declustered* placement schemes that represent the two extremes of spreading codewords across devices. For declustered placement, the symbols are spread equally across all remaining devices, whereas for clustered placement, the symbols are spread across the smallest possible number of devices, as shown in Figure 1. For clustered placement, the n devices are divided into disjoint sets of m devices, referred to as *clusters*, and each codeword is stored across the devices of a particular cluster as shown in Figure 1. In such a placement scheme, there are n/m placement groups. The entire storage system is modeled as consisting of n/m independent clusters. In any given cluster, data loss occurs when $\tilde{r}$ devices fail successively before rebuild operations can be completed successfully. For declustered placement, all $\binom{n}{m}$ possible ways of placing m symbols across n devices are used equally to store all the codewords in the system, as shown in Figure 1.
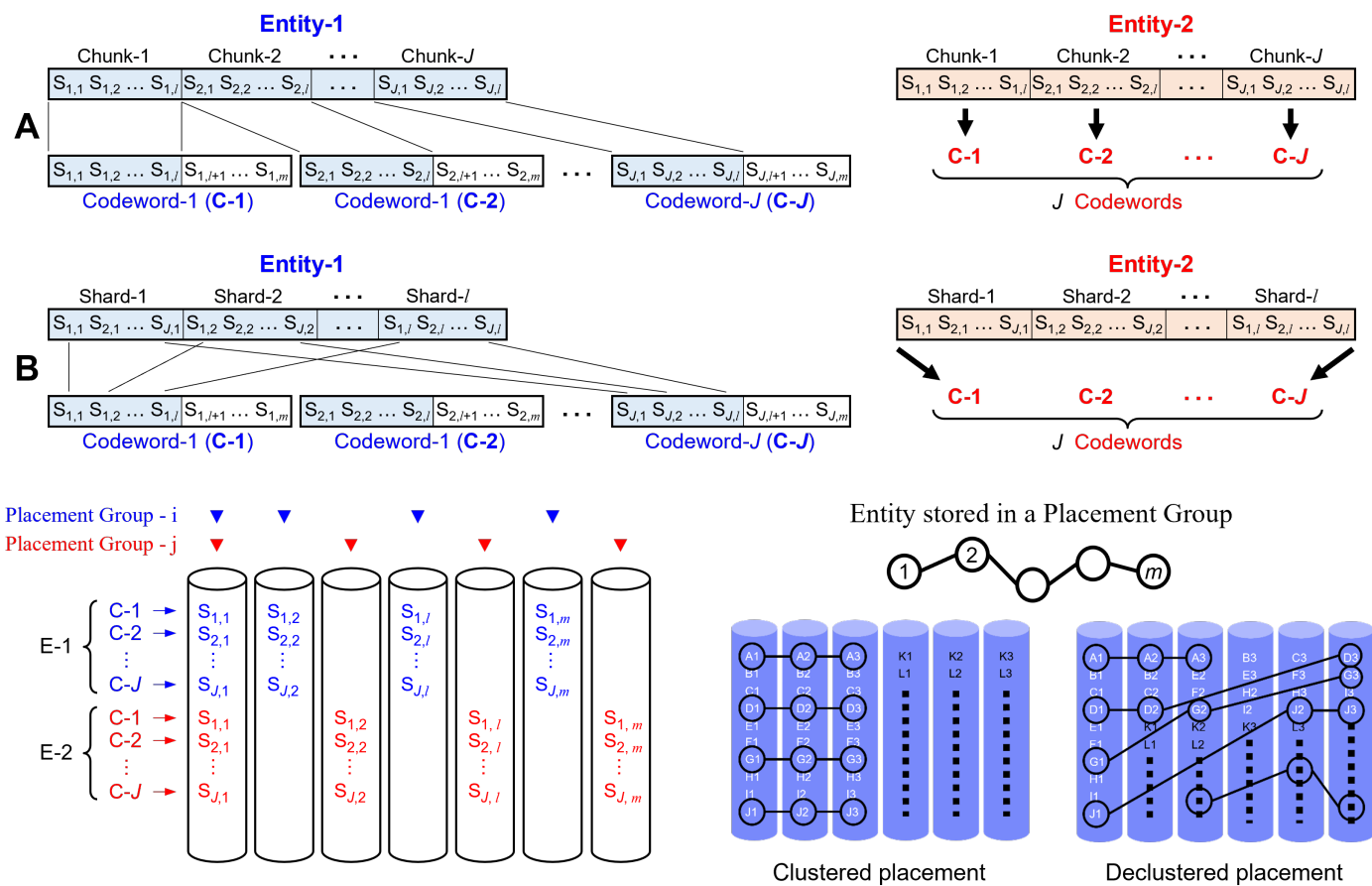
Figure 1. Formation of codewords from entity data contents and their placement on storage devices. Two methods (A and B) for identical formation of codewords. Clustered and declustered placement of codewords of length $m = 3$ on $n = 6$ devices. X1, X2, X3 represent codewords of an entity (X = A, B, C, . . . , L).

A. Rebuild Process for Entity and Codeword Reconstruction

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

$$l < m \leq k \leq n \quad (n/k \in \mathbb{N}) \quad \text{and} \quad 0 \leq d \leq m - l - 1 . \quad (5)$$

When a codeword of an entity loses $\tilde{r}$ or more symbols, this codeword, and consequently the entity is permanently lost. Note that the reconstruction of successive codewords leads to the successive reconstruction of entities.

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

$$C_1 = C , \quad (6)$$

where C is determined by (2). The reliability metrics of interest are derived in Section III using the *direct path approximation*, which considers only transitions from lower to higher exposure levels [3][16][22][23][29]. 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 [18][20][21], 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][12][18][29]. 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. 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, the lost codeword symbols of successive entities are reconstructed on other appropriately chosen placement groups, which ensures that the lost codeword symbols are reconstructed on devices in which another symbol of the same codeword is not already present. In the case of clustered placement, the lost symbols are reconstructed directly in spare devices as described and shown in Figure 3 of [18]. As an entity contains J codewords, it is assumed that the number E_u of entities associated with the C_u codewords is approximately C_u/J.

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

$$\mu^{-1} \triangleq E(X) = c/b . \tag{7}$$

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 [18][22]

$$\mu \int_0^\infty F_\lambda(t)[1 - F_X(t)]dt \ll 1, \quad \text{with } \frac{\lambda}{\mu} \ll 1 . \tag{8}$$

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 $u \rightarrow u+1$. For $u \leq d$, no rebuild is performed and therefore $\alpha_u = 1$. For $u > d$, α_u is approximately uniformly distributed in $(0,1)$ [31, Lemma 2]. Therefore,

$$\alpha_u \approx \begin{cases} 1, & \text{for } u = 1, \ldots, d \\ U(0,1), & \text{for } u = d+1, \ldots, \tilde{r}-1 . \end{cases} \tag{9}$$

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 [13][17][18][23]. For the rebuild schemes considered, it holds that [13, Eq. (10)]:

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

with the convention that for any integer j and for any sequence δ_i, $\prod_{i=j}^0 \delta_i \triangleq 1$. Unconditioning (10) on $\alpha_1, \ldots, \alpha_{u-1}$ yields

$$E(C_u) = C \left(\prod_{j=1}^{u-1} V_j \right) E \left(\prod_{j=1}^{u-1} \alpha_j \ \middle| \ \text{level } u \text{ was entered} \right),$$
$$\text{for } u = 1, \ldots, \tilde{r} . \tag{11}$$

It also holds that [13, Eq. (49)]:

$$R_{d+1} \approx \left(\prod_{j=1}^d V_j \right) \frac{b}{b_{d+1}} \, X . \tag{12}$$

6) Unrecoverable Errors: The reliability of storage systems is affected by the occurrence of unrecoverable or latent errors. According to the specifications of enterprise quality HDDs, the unrecoverable bit-error probability P_{bit} is equal to 10^{-15} [14]. Assuming that bit errors occur independently over successive bits, the unrecoverable sector error probability P_s is

$$P_s = 1 - (1 - P_{\text{bit}})^s , \tag{13}$$

with s expressed in bits. For a sector size of 512 bytes, we have $P_s \approx P_{\text{bit}} \times 4096 = 4.096 \times 10^{-12}$. In practice, however, and also owing to the accumulation of latent errors over time, these probability values are higher. Recent investigations show that P_{bit} can be orders of magnitude higher, reaching $P_{\text{bit}} \approx 10^{-12}$ [32], which results in $P_s \approx 4.096 \times 10^{-9}$. Moreover, latent errors are found to exhibit spatial locality and they occur in bursts of multiple contiguous sector errors. The mean burst length $\bar{B}$ reported in [21] is 1.0291. In the next section, we show that the correlation of latent errors affects both MTTDL and EAFEL reliability metrics.

III. DERIVATION OF EAFEL

The EAFEL reliability metric is derived using the direct-path-approximation methodology presented in [16-21]. 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 $E(T_0) = 1/(n\,\lambda)$ [22]. 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

$$E(T) = \sum_{u=0}^{d} E(T_u) = \left(\sum_{u=0}^{d} \frac{1}{\tilde{n}_u} \right) \Big/ \lambda \;, \quad \text{where } \tilde{n}_0 \triangleq n \;, \tag{14}$$

where $\tilde{n}_u$ is determined by (58) or (61).

The MTTDL metric is then obtained by [16, Eq. (5)]:

$$\text{MTTDL} \approx \frac{E(T)}{P_{\text{DL}}} \;. \tag{15}$$

The EAFEL is obtained as the ratio of the expected number $E(Y)$ of lost entities, normalized to the number N_E of entities, to the expected duration of T:

$$\text{EAFEL} \approx \frac{E(Y)}{E(T) \cdot N_E} \stackrel{(3)}{=} \frac{m\, J\, s\, E(Y)}{n\, c\, E(T)} \;. \tag{16}$$

where $E(T)$ is determined by (14) and expressed in years.

The expected number $E(G)$ of lost entities, given that data loss has occurred, is determined by

$$E(G) = \frac{E(Y)}{P_{\text{DL}}} \;. \tag{17}$$

It follows from (15), (16), and (17) that

$$\text{EAFEL} = \frac{E(G)}{\text{MTTDL} \cdot N_E} \;, \tag{18}$$

with the MTTDL expressed in years.

A. Reliability Analysis

The EAFEL is evaluated in parallel with MTTDL using the theoretical framework presented in [13]. At any exposure level u ($u = d+1, \dots, \tilde{r}-1$), data loss may occur during rebuild owing to one or more unrecoverable failures, which is denoted by the transition $u \to \text{UF}$. 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 $1 \to 2 \to \cdots \to u \to \text{UF}$, for $u = d+1, \dots, \tilde{r}-1$, and

$\overrightarrow{DF}$: the direct path of successive transitions $1 \to 2 \to \cdots \to \tilde{r}-1 \to \tilde{r}$,

with their respective probabilities P_{UF_u} and P_{DF} determined by virtue of Proposition 1 of [13] and Eq. (32) of [21] as follows:

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

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

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

$$f_{\text{cor}} \triangleq \begin{cases} 1, & \text{for independent latent errors} \\ \frac{1}{B}, & \text{for correlated latent errors, and} \end{cases} \tag{22}$$

$$P_{\text{DF}} \approx \frac{(\lambda c \prod_{j=1}^{d} V_j)^{\tilde{r}-d-1}}{(\tilde{r}-d-1)!} \frac{E(X^{\tilde{r}-d-1})}{[E(X)]^{\tilde{r}-d-1}} \prod_{i=d+1}^{\tilde{r}-1} \frac{\tilde{n}_i}{b_i} V_i^{\tilde{r}-1-i} \;. \tag{23}$$

Remark 1: Independently of threshold d, it holds that [21]

$$q_u \approx \begin{cases} 1 - \binom{m-u}{\tilde{r}-u} P_s^{\tilde{r}-u}, & \text{for } P_s \ll \frac{1}{m-u} < \binom{m-u}{\tilde{r}-u}^{-\frac{1}{\tilde{r}-u}} \\ 0, & \text{for } P_s \gg \binom{m-u}{\tilde{r}-u}^{-\frac{1}{\tilde{r}-u}} > \frac{1}{m-u} \;, \end{cases} \tag{24}$$

$$\hat{q}_u \approx \begin{cases} 1 - Z_u P_s^{\tilde{r}-u}, & \text{for } P_s \ll P_{s,u}^* \\ 0, & \text{for } P_s \gg P_{s,u}^* \;, \end{cases} \tag{25}$$

where $\quad Z_u \triangleq f_{\text{cor}} C \binom{m-u}{\tilde{r}-u} \prod_{j=1}^{u-1} V_j \;, \quad P_{s,u}^* \triangleq Z_u^{-\frac{1}{\tilde{r}-u}} \;. \tag{26}$

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

$$P_{\text{UF}_u} \approx \begin{cases} A_u P_s^{\tilde{r}-u}, & \text{for } P_s \ll P_{\tilde{s},u}^{(\tilde{r})} \\ P_u, & \text{for } P_s \gg P_{\tilde{s},u}^{(\tilde{r})} \;, \end{cases} \tag{27}$$

where

$$A_u \triangleq \frac{Z_u}{u-d} P_u \;, \tag{28}$$

$$P_u \approx \frac{(\lambda c \prod_{j=1}^{d} V_j)^{u-d-1}}{(u-d-1)!} \frac{E(X^{u-d-1})}{[E(X)]^{u-d-1}} \prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} V_i^{u-1-i} \;, \tag{29}$$

$$P_{\tilde{s},u}^{(\tilde{r})} \triangleq \left(\frac{u-d}{Z_u} \right)^{\frac{1}{\tilde{r}-u}} \;, \tag{30}$$

P_u is the probability of entering exposure level u, and f_{cor} and Z_u are determined by (22) and (26), respectively.

Proof: From [13, Eqs. (57), (61), and (65)], it follows that

$$P_{\text{UF}_u}(R_{d+1}) \approx -(\lambda b_{d+1} R_{d+1})^{u-d-1}$$
$$\cdot \left(\prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} V_i^{u-1-i} \right) \sum_{i=1}^{\infty} \frac{\log(\hat{q}_u)^i}{(u-d-1+i)!} \;. \tag{31}$$

From (25) and for $P_s \ll P_{s,u}^*$, it follows that $\hat{q}_u \approx 1$. Furthermore, $\log(\hat{q}_u) = -(1-\hat{q}_u) + O((1-\hat{q}_u)^2) \approx -(1-\hat{q}_u)$. Consequently, by virtue of (25), it holds that $\log(\hat{q}_u) \approx -Z_u P_s^{\tilde{r}-u}$. For small values of P_s, all the terms of the summation in (31) are negligible compared with the first one. Therefore, from the above, and unconditioning (31) on R_{d+1} using (7) and (12), yields (27). Also, from (25) and for $P_s \gg P_{s,u}^*$, it follows that $\hat{q}_u \approx 0$, which by virtue of Eq. (59) of [13] implies that $P_{u \to \text{UF}}(R_{d+1}, \vec{\alpha}_{u-1}) \approx 1$. Consequently, it follows from Eq. (46) of [13] that $P_{\text{UF}_u} \approx P_u$, which is obtained by Eq. (17) of [13] and determined by (29). Also, $P_{\tilde{s},u}^{(\tilde{r})}$ is obtained from the approximation (27) $P_{\text{UF}_u} \approx A_u (P_{\tilde{s},u}^{(\tilde{r})})^{\tilde{r}-u} = P_u$ and (28). ∎

In [24], it was shown that the probability of data loss P_{DL} is accurately approximated by the probability of all direct paths to data loss. Therefore,

$$P_{\text{DL}} \approx P_{\text{DF}} + \sum_{u=d+1}^{\tilde{r}-1} P_{\text{UF}_u} \approx P_{\text{DF}} + P_{\text{UF}} \;, \tag{32}$$

where P_{UF} is the probability of data loss due to unrecoverable failures determined by

$$P_{\text{UF}} \approx \sum_{u=1}^{\tilde{r}-1} P_{\text{UF}_u} \;, \tag{33}$$

The MTTDL metric is obtained by substituting (32) into (15) as follows:

$$\text{MTTDL} \approx \frac{E(T)}{P_{\text{DF}} + \sum_{u=d+1}^{\tilde{r}-1} P_{\text{UF}_u}} , \qquad (34)$$

where $E(T)$, P_{UF_u} and P_{DF} are determined by (14), (19), and (23), respectively. Note that P_{DL} and, consequently, MTTDL do not depend on the entity size.

We proceed to derive the number of lost entities during rebuild. Let Y and G be the number of lost entities and the conditional number of lost entities, given that data loss has occurred, respectively. Let also Y_{DF} and Y_{UF_u} denote the number of lost entities associated with the direct paths $\overrightarrow{DF}$ and $\overrightarrow{UF_u}$, respectively. Similarly, we consider the variables G_{DF} and G_{UF_u}. Then, the number Y of lost entities is obtained by

$$Y \approx \begin{cases} G_{\text{DF}} , & \text{if } \overrightarrow{DF} \\ G_{\text{UF}_u} , & \text{if } \overrightarrow{UF_u} , \ \text{for } u = d+1, \dots, \tilde{r}-1 \\ 0 , & \text{otherwise} . \end{cases} \quad (35)$$

Thus, $\quad E(Y) \approx P_{\text{DF}}\, E(G_{\text{DF}}) + \sum_{u=d+1}^{\tilde{r}-1} P_{\text{UF}_u}\, E(G_{\text{UF}_u})$ (36)

$$= E(Y_{\text{DF}}) + \sum_{u=d+1}^{\tilde{r}-1} E(Y_{\text{UF}_u}) , \qquad (37)$$

$$\approx E(Y_{\text{DF}}) + E(Y_{\text{UF}}) , \qquad (38)$$

where $\quad E(Y_{\text{DF}}) = P_{\text{DF}}\, E(G_{\text{DF}}) ,$ (39)

$$E(Y_{\text{UF}_u}) = P_{\text{UF}_u}\, E(G_{\text{UF}_u}) , \quad u = d+1, \dots, \tilde{r}-1 , \quad (40)$$

$$E(Y_{\text{UF}}) = P_{\text{UF}}\, E(G_{\text{UF}}) \approx \sum_{u=1}^{\tilde{r}-1} E(Y_{\text{UF}_u}) , \qquad (41)$$

where Y_{UF} denotes the number of lost entities due to unrecoverable failures and G_{UF} the conditional number of lost entities, given that data loss due to unrecoverable failures has occurred.

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

$$E(Y_{\text{UF}_u}) \approx \frac{C}{J} \frac{P_u}{u-d} \left(\prod_{j=1}^{u-1} V_j \right) \tilde{q}_u , \qquad (42)$$

where $\qquad \tilde{q}_u \triangleq 1 - q_u^{f_{\text{cor}} J} ,$ (43)

$$E(Y_{\text{DF}}) \approx \frac{C}{J} \frac{P_{\text{DF}}}{\tilde{r}-d} \prod_{j=1}^{\tilde{r}-1} V_j , \qquad (44)$$

and P_u is determined by (29).

Proof: Equation (42) is obtained in Appendix A. Equation (44) is obtained from (42) by setting $u = \tilde{r}$ and recognizing that $q_{\tilde{r}} = 0$, $\tilde{q}_{\tilde{r}} = 1$, and $P_{\tilde{r}} = P_{\text{DF}}$. ∎

Remark 2: It follows from (24) and (43) that

$$\tilde{q}_u \approx \begin{cases} f_{\text{cor}} J \binom{m-u}{\tilde{r}-u} P_s^{\tilde{r}-u}, & \text{for } P_s \ll [\, f_{\text{cor}} J \binom{m-u}{\tilde{r}-u}]^{-\frac{1}{\tilde{r}-u}} \\ 1 , & \text{for } P_s \gg [\, f_{\text{cor}} J \binom{m-u}{\tilde{r}-u}]^{-\frac{1}{\tilde{r}-u}} . \end{cases} \quad (45)$$

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

$$E(Y_{\text{UF}_u}) \approx \begin{cases} A_u\, P_s^{\tilde{r}-u} , & \text{for } P_s \ll [\, f_{\text{cor}} J \binom{m-u}{\tilde{r}-u}]^{-\frac{1}{\tilde{r}-u}} \\ \dfrac{C}{J} \dfrac{P_u}{u-d} \displaystyle\prod_{j=1}^{u-1} V_j , & \text{for } P_s \gg [\, f_{\text{cor}} J \binom{m-u}{\tilde{r}-u}]^{-\frac{1}{\tilde{r}-u}} , \end{cases} \quad (46)$$

where A_u and P_u are determined by (28) and (29), respectively.

Proof: Immediate by substituting (45) into (42) and invoking (26), (28), and (29). ∎

Remark 3: For small values of P_s, from (27) and (46), and by virtue of (33) and (41), it follows that

$$E(Y_{\text{UF}_u}) \approx P_{\text{UF}_u} , \ \forall\, u ; \quad E(Y_{\text{UF}}) \approx P_{\text{UF}} , \ \text{for } P_s \to 0 . \quad (47)$$

Consequently, (40) and (41) imply that

$$E(G_{\text{UF}_u}) \approx 1 , \ \forall\, u ; \quad E(G_{\text{UF}}) \approx 1 , \ \text{for } P_s \to 0 . \quad (48)$$

For small values of P_s, and according to (27) and (46), all the terms of the summations in (33) and (41) are negligible compared with the last ones, which implies that $P_{\text{UF}} \approx P_{\text{UF}_{\tilde{r}-1}}$ and $E(Y_{\text{UF}}) \approx E(Y_{\text{UF}_{\tilde{r}-1}})$. Thus, when an unrecoverable failure occurs, this failure most likely occurs when rebuilding a codeword after it has lost $\tilde{r}-1$ of its symbols owing to $\tilde{r}-1$ device failures and an unrecoverable error is encountered. In this case, and according to (48), there is a single entity lost.

Remark 4: From (39) and (44) it follows that

$$E(G_{\text{DF}}) \approx \frac{C}{J} \frac{1}{\tilde{r}-d} \prod_{j=1}^{\tilde{r}-1} V_j . \qquad (49)$$

For small values of P_s, it holds that $P_{\text{UF}} \to 0$ and $E(Y_{\text{UF}}) \to 0$, which by virtue of (32) and (38) implies that $P_{\text{DL}} \to P_{\text{DF}}$ and $E(Y) \to E(Y_{\text{DF}})$. Consequently, from (17), (39), and (49), it follows that

$$E(G) \approx E(G_{\text{DF}}) \approx \frac{C}{J} \frac{1}{\tilde{r}-d} \prod_{j=1}^{\tilde{r}-1} V_j , \ \text{for } P_s \to 0 . \quad (50)$$

The EAFEL metric is obtained by substituting (37) into (16) as follows:

$$\text{EAFEL} \approx \frac{m\, J\, s\, [E(Y_{\text{DF}}) + \sum_{u=d+1}^{\tilde{r}-1} E(Y_{\text{UF}_u})]}{n\, c\, E(T)} , \qquad (51)$$

where $E(Y_{\text{UF}_u})$ and $E(Y_{\text{DF}})$ are determined by (42) and (44), respectively, and $E(T)$ by (14) and expressed in years.

The conditional number $E(G)$ of lost entities, given that data loss has occurred, is obtained from (17), where $E(Y)$ is determined by (37), (42), and (44), and P_{DL} is determined by (32).

Note that when entering exposure level $\tilde{r}$, the $C_{\tilde{r}}$ most-exposed codewords cause $C_{\tilde{r}}/J$ entities to be lost, that is,

$$E(G_{\text{DF}} \,|\, C_{\tilde{r}}) = C_{\tilde{r}}/J . \qquad (52)$$

Unconditioning (52) on $C_{\tilde{r}}$ yields

$$E(G_{\text{DF}}) = E(C_{\tilde{r}})/J . \qquad (53)$$

Combining (49) and (53) yields

$$E(C_{\tilde{r}}) \approx \left(\prod_{i=1}^{\tilde{r}-1} V_i \right) \frac{C}{\tilde{r}-d} . \qquad (54)$$

From (11) and (54), it follows that

$$E\left(\prod_{j=1}^{\tilde{r}-1}\alpha_j \,\middle|\, \text{level } \tilde{r} \text{ was entered}\right) \approx \frac{1}{\tilde{r}-d} \;. \quad (55)$$

Remark 5: It turns out that when a data loss has occurred, the variables $\alpha_{d+1}, \ldots, \alpha_{\tilde{r}-1}$ are not distributed identically. More specifically, for a rebuild time R_u, the uniform distribution of α_u in the interval $(0,1)$, given by (9), holds under the assumption that there is a failure during this rebuild period, that is, an exposure level transition $u \to u+1$. However, conditioning on the exposure level transitions $u \to u+1 \to \cdots \to u' \to u'+1$ $(u' > u)$, α_u is no longer uniformly distributed in $(0,1)$. This is due to the fact that, conditioning on the fact that additional failures occur during the rebuild times $R_{u+1}, \ldots, R_{u'}$, it is more likely that the R_{u+1} period is long rather than short. In this case, only α'_u is uniformly distributed in $(0,1)$. Assuming that the system has entered exposure level u, we deduce from (55) that

$$E\left(\prod_{j=1}^{u-1}\alpha_j \,\middle|\, \text{level } u \text{ was entered}\right) \approx \frac{1}{u-d} \;, \quad u = d+1, \ldots, \tilde{r}.$$
$$(56)$$

Substituting (56) into (11) and using (9) yields

$$E(C_u) \approx \left(\prod_{i=1}^{u-1}V_i\right)\frac{C}{\max(u-d,1)} \;, \quad u = 1, \ldots, \tilde{r} \;. \quad (57)$$

Remark 6: From (42) and considering (57), it holds that $E(Y_{\text{UF}_u}) \approx [E(C_u)/J\,\tilde{q}_u]\,P_u$, that is, the expected number of lost entities at exposure level u is the product of the expected number of lost entities at exposure level u, given that the system has entered exposure level u, and the probability of entering exposure level u. Similarly, from (44) and considering (54), it holds that $E(Y_{\text{DF}}) \approx [E(C_{\tilde{r}})/J]\,P_{\text{DF}}$, given that at exposure level $\tilde{r}$, all of the $E(C_{\tilde{r}})$ exposed codewords and corresponding $E(C_{\tilde{r}})/J$ entities are lost.

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 [17][18]

$$\tilde{n}_u^{\text{sym}} = k - u \;, \quad (58)$$

$$b_u^{\text{sym}} = \frac{\min((k-u)\,b, B_{\max})}{l+1} \;, \quad (59)$$

$$V_u^{\text{sym}} = \frac{m-u}{k-u} \;. \quad (60)$$

The corresponding parameters $\tilde{n}_u^{\text{declus}}$, b_u^{declus}, and V_u^{declus} for the declustered placement are derived from (58), (59), and (60) by setting $k = n$.

C. Clustered Placement

At each exposure level u, for $u = 1, \cdots, \tilde{r}-1$, it holds that [17][18]

$$\tilde{n}_u^{\text{clus}} = m - u \;, \quad b_u^{\text{clus}} = \min(\,b, B_{\max}/l\,) \;, \quad V_u^{\text{clus}} = 1 \;. \quad (61)$$

Remark 7: From (42), (44), (19), and (23), and considering expressions (58) through (61), it follows that $E(Y_{\text{UF}_u})$ and

P_{UF_u} are mainly determined by the term $(\lambda\,c/b)^{u-d-1}$, and $E(Y_{\text{DF}})$ and P_{DF} by the term $(\lambda\,c/b)^{\tilde{r}-d-1}$. According to (8), $\lambda\,c/b \ll 1$, such that, for fixed values of $\tilde{r}$ and u, increasing d causes these parameters to increase. Therefore, by virtue of (34) and (51), increasing d causes EAFEL to increase and MTTDL to decrease. 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 (4) and (5), it follows that

$$0 \le d < z < m \quad \text{and} \quad d+1 \le u \le m-z+d+1 \;. \quad (62)$$

Next, we compare the EAFEL of equivalent systems. From (42) and (45), it follows that

$$\frac{\frac{E(Y_{\text{UF}_{u+1}}|d+1)}{E(T|d+1)}}{\frac{E(Y_{\text{UF}_u}|d)}{E(T|d)}} \approx \frac{E(T|d)}{E(T|d+1)} \cdot \prod_{i=d+1}^{u}V_i \cdot \frac{z-d-1}{m-u} \cdot A \;, \quad (63)$$

$$\text{where} \quad A = \begin{cases} \left(\frac{z-d}{z-d+1}\right)^{u-d-1}, & \text{for symmetric placement} \\ \frac{m-u}{m-d-1}, & \text{for clustered placement} \end{cases} \quad (64)$$

From (14), it follows that $\frac{E(T|d)}{E(T|d+1)} < 1$. Consequently, from (51), (62), and (63), and recognizing that $A < 1$ and $E(Y_{\text{DF}}) = E(Y_{\text{UF}_{\tilde{r}}})$, it follows that

$$\frac{\text{EAFEL}(d+1)}{\text{EAFEL}(d)} < 1 \;. \quad (65)$$

It also holds that [13, Eqs. (46) and (47)]:

$$\frac{\text{MTTDL}(d+1)}{\text{MTTDL}(d)} > 1 \;. \quad (66)$$

Remark 8: Within the class of equivalent systems, according to (65) and (66), 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) and protected by an erasure coding scheme with $m = 16$, which is the codeword length used by Microsoft® Azure [7], and $l = 13$, 14, and 15. Each device stores an amount of $c = 20$ TB, which is the capacity of the latest generation of Seagate drives, and the symbol size s is equal to a sector size of 512 bytes [33].

Typical parameter values are listed in Table II. The Annualized Failure Rate (AFR) of HDDs for the year 2021 is in the range of 0.11% to 4.79% [34], which corresponds to a mean time to failure in the range of $180,000$ h to $8,000,000$ h. The parameter λ^{-1} is chosen to be equal to $876,000$ h (100 years) that corresponds to an AFR of 1%, which is the average AFR across all drive models [34]. Considering that 35% of the maximum transfer rate of 285 MB/s [33] is allocated for
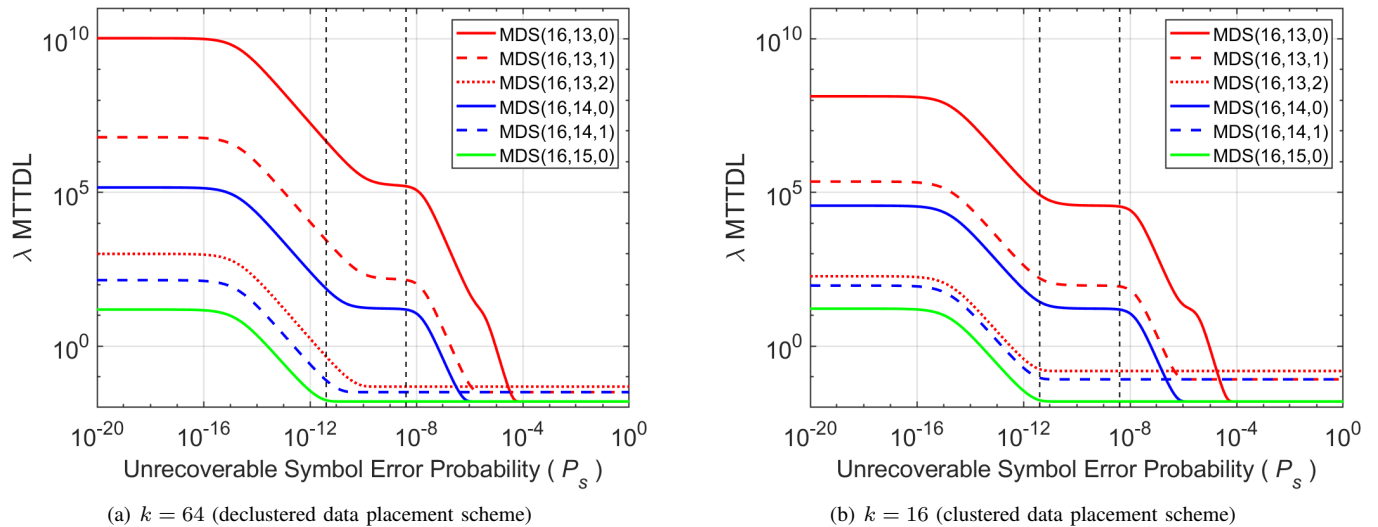
(a) $k = 64$ (declustered data placement scheme)

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

Figure 2.   Normalized MTTDL vs. P_s for various MDS(m, l, d) codes; $n = 64$, $\lambda/\mu = 0.00006$, $c = 20$ 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	20 TB
s	symbol (sector) size	512 B
λ^{-1}	mean time to failure of a storage device	876,000 h
b	rebuild bandwidth per device	100 MB/s
m	symbols per codeword	16
l	user-data symbols per codeword	13, 14, 15
U	amount of user data stored in the system	1.04 to 1.2 PB
μ^{-1}	time to read an amount c of data at a rate b from a storage device	55.5 h

recovery operations, the reserved rebuild bandwidth b is then equal to 100 MB/s, which yields a rebuild time of a device $\mu^{-1} = c/b = 55.5$ h. Also, it is assumed that the maximum network rebuild bandwidth is sufficiently large ($B_{\max} \geq n\,b = 6.4$ GB/s), that the rebuild time distribution is deterministic, such that $E(X^k) = [E(X)]^k$, and that sector errors are correlated with $B \approx 1$, which implies that $f_{\text{cor}} \approx 1$. The obtained results are accurate, because (8) is satisfied, given that $\lambda/\mu = 6.3 \times 10^{-5} \ll 1$.

First, we assess the reliability for the declustered placement scheme ($k = n = 64$) for the MDS-coded configurations considered in [13] 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 2, 3, 4, 5, and 6 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 λ MTTDL measure, which does not depend on the entity size, is obtained from (34) as a function of P_s and shown in Figure 2(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 $[4.096 \times 10^{-12}, 4.096 \times 10^{-9}]$ 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 7, 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 8, MTTDL increases as d increases.

The normalized λ MTTDL measure for the clustered data placement scheme is shown in Figure 2(b). We observe that the declustered placement scheme achieves a significantly higher MTTDL than the clustered one.

In contrast to MTTDL that is not dependent on the entity size, the other reliability metrics depend on it. Its effect is assessed by considering the following three cases: entity sizes of 6.656 KB, 6.656 MB, and 6.656 GB that correspond to entities containing 1, 1000, and 1,000,000 codewords, respectively.

The normalized EAFEL/λ measure corresponding to the declustered data placement scheme is obtained from (51) and shown in Figure 3 for various entity sizes. We observe that EAFEL increases monotonically, but for small entity sizes, 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, as shown in Figure 3(a). However, for medium and large entity sizes, Figures 3(b) and 3(c) reveal that EAFEL degrades in the interval of interest and the larger the entity size, the more pronounced the degradation. By contrast, for clustered data placement and in the interval of interest, Figure 4 reveals that EAFEL degrades only when entities are very large.

For the EAFEL 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 8, EAFEL decreases as d increases, but according to Figures 3 and 4 the EAFEL reduction for the clustered placement is much less than that for the declustered placement. We observe that for both MTTDL

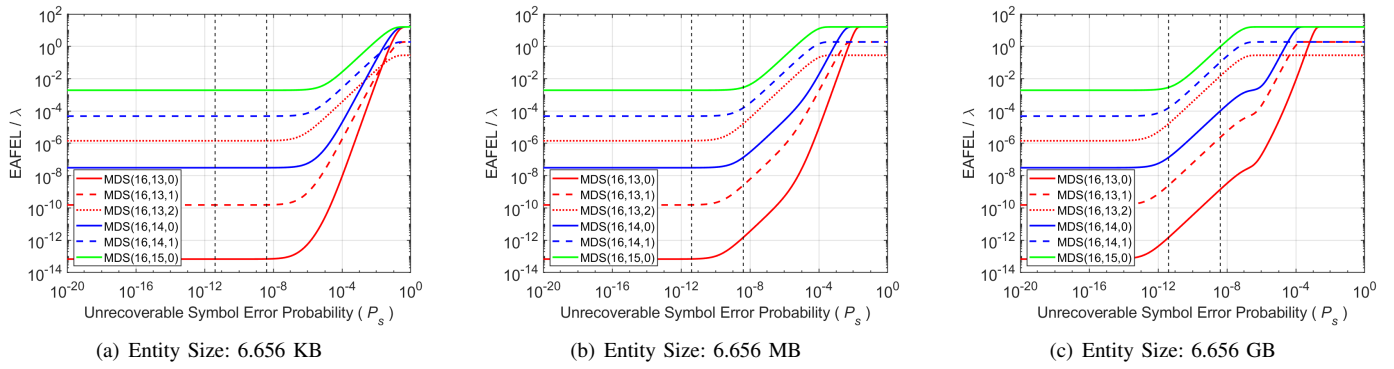(a) Entity Size: 6.656 KB   (b) Entity Size: 6.656 MB   (c) Entity Size: 6.656 GB

Figure 3.   Normalized EAFEL vs. P_s for various MDS(m, l, d) codes; declustered data placement, $n = 64$, $\lambda/\mu = 0.00006$, $c = 20$ TB, and $s = 512$ B.
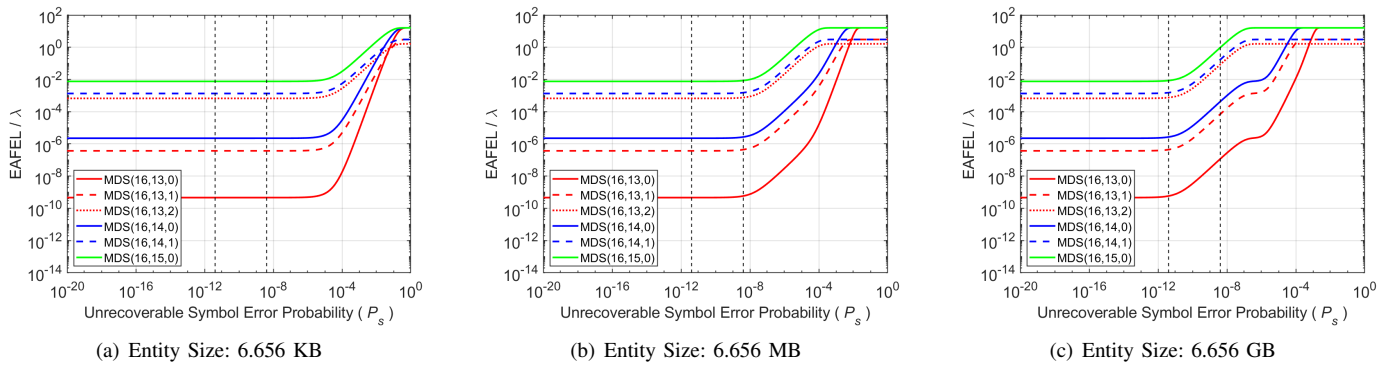

(a) Entity Size: 6.656 KB   (b) Entity Size: 6.656 MB   (c) Entity Size: 6.656 GB

Figure 4.   Normalized EAFEL vs. P_s for various MDS(m, l, d) codes; clustered data placement, $n = 64$, $\lambda/\mu = 0.00006$, $c = 20$ TB, and $s = 512$ B.
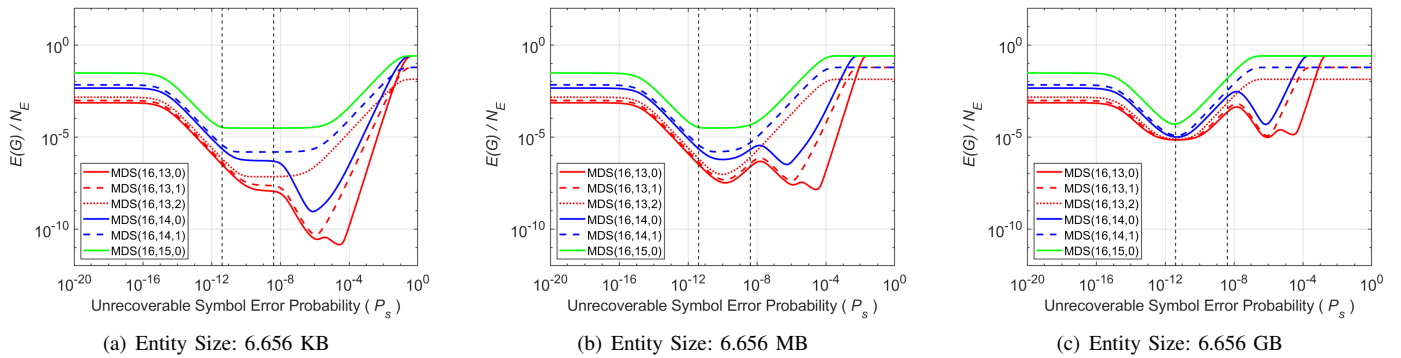

(a) Entity Size: 6.656 KB   (b) Entity Size: 6.656 MB   (c) Entity Size: 6.656 GB

Figure 5.   Normalized $E(G)$ vs. P_s for various MDS(m, l, d) codes; declustered data placement, $n = 64$, $\lambda/\mu = 0.00006$, $c = 20$ TB, and $s = 512$ B.
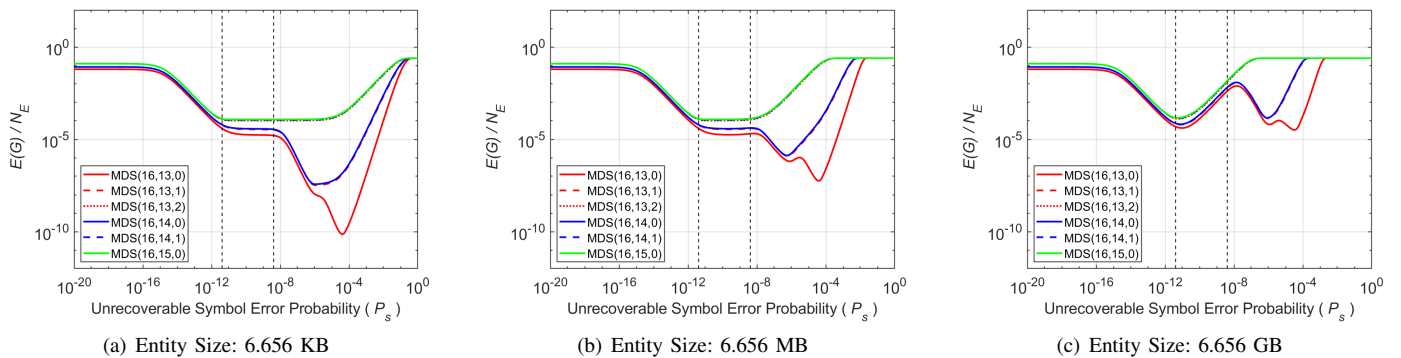

(a) Entity Size: 6.656 KB   (b) Entity Size: 6.656 MB   (c) Entity Size: 6.656 GB

Figure 6.   Normalized $E(G)$ vs. P_s for various MDS(m, l, d) codes; clustered data placement, $n = 64$, $\lambda/\mu = 0.00006$, $c = 20$ TB, and $s = 512$ B.

and EAFEL reliability metrics, the reliability level achieved by the declustered data placement scheme is higher than that of the clustered one.

The normalized expected number $E(G)/N_E$ of lost entities, given that a data loss has occurred, relative to the total number of entities is obtained from (17) and shown in Figures 5 and 6 for the declustered and clustered data placement schemes, respectively. In contrast to the P_{DL}, EAFEL, and $E(Q)$ metrics that increase monotonically with P_s, we observe that $E(G)$ does not do so. The reason for that is the following. From Figure 2, it follows that for $P_s \gg 10^{-15}$, data loss is more likely to be due to sector errors than to device failures. Given that sector errors result in a negligible number of lost entities compared with the substantial number of lost entities caused by device failures, when P_s increases over the value of 10^{-15}, the conditional number of lost entities decreases. Clearly, this is reversed for high values of P_s, and the conditional number of lost entities increases.

Also, in the interval $[4.096 \times 10^{-12}, 4.096 \times 10^{-9}]$ of practical importance for P_s, and by contrast to MTTDL and EAFEL, employing lazy rebuild does not affect $E(G)$ 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, $E(G)$ is lower for the declustered data placement scheme, but it is not significantly affected for the clustered one. However, Figures 5(c) and 6(c) reveal that this difference is reduced when entities are very large.

V. CONCLUSIONS

We considered the Expected Annual Fraction of Entity Loss (EAFEL) metric, which assesses the durability of distributed and cloud storage systems and reflects losses at an entity, say file or object, level. This metric, together with the traditional MTTDL metric, provide a useful profile of the size and frequency of data losses. The methodology developed for deriving the Mean Time to Data Loss (MTTDL) and the Expected Annual Fraction of Data Loss (EAFDL) reliability metrics was extended to derive analytically the EAFEL metric for erasure-coding redundancy schemes. Closed-form expressions capturing the effect of unrecoverable latent errors and lazy rebuild were obtained for the symmetric, clustered and declustered data placement schemes. We established that, for realistic unrecoverable sector error rates, MTTDL is adversely affected by the presence of latent errors, whereas EAFEL is adversely affected only when entities are large. It was also shown that the declustered data placement scheme offers superior reliability. 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.

An extension of the analytical framework developed to consider also the cases of fixed-size entities that do not contain an integer number of codewords and of variable-size entities is a subject of further investigation.

APPENDIX A

Proof of Proposition 1.

Upon entering exposure level u ($u \geq d+1$), there are C_u most-exposed codewords to be recovered. As an entity contains

J codewords, the number E_u of entities to be recovered is

$$E_u \approx \frac{C_u}{J}, \quad \text{for } u = d+1, \ldots, \tilde{r}-1. \tag{67}$$

Furthermore, according to Eq. (98) of [21], the probability of recovering an entity is $q_u^{f_{\mathrm{cor}} J}$, where q_u is the probability of restoring a codeword and is determined by (21), and f_{cor} accounts for the correlation of latent errors and is determined by (22). Consequently, the probability $\tilde{q}_u$ that an entity is lost is determined by (43).

Let us define the variable $W_{u,i}$ as follows:

$$W_{u,i} = \begin{cases} 1, & \text{when entity } E_{u,i} \text{ is lost} \\ 0, & \text{when entity } E_{u,i} \text{ is recovered, } i = 1, \ldots, E_u, \end{cases} \tag{68}$$

which implies that

$$E(W_u) = E(W_{u,i}) = P(\text{entity } E_{u,i} \text{ is lost}) = \tilde{q}_u, \quad \text{for } i = 1, \ldots, E_u. \tag{69}$$

Let us also define by Y_{U} the number of lost entities at exposure level u during the rebuild process of the C_u codewords. Then it holds that,

$$Y_{\mathrm{U}} = \sum_{i=1}^{E_u} W_{u,i}, \tag{70}$$

which implies that

$$E(Y_{\mathrm{U}}|C_u) = E\left(\sum_{i=1}^{E_u} W_{u,i}\right) = \sum_{i=1}^{E_u} E(W_{u,i})$$
$$\overset{(69)}{=} \sum_{i=1}^{E_u} E(W_u) = E_u \, \tilde{q}_u \overset{(67)}{\approx} \frac{C_u}{J} \, \tilde{q}_u. \tag{71}$$

Substituting (10) into (71) yields

$$E(Y_{\mathrm{U}}|\vec{\alpha}_{u-1}) \approx \frac{C}{J} \left(\prod_{j=1}^{u-1} V_j \, \alpha_j\right) \tilde{q}_u. \tag{72}$$

Subsequently, the expected number $E(Y_{\mathrm{UF}_u}|R_{d+1}, \vec{\alpha}_{u-1})$ of lost entities 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

$$E(Y_{\mathrm{UF}_u}|R_{d+1}, \vec{\alpha}_{u-1}) = P_u(R_{d+1}, \vec{\alpha}_{u-1}) \, E(Y_{\mathrm{U}}|\vec{\alpha}_{u-1}), \tag{73}$$

where $P_u(R_{d+1}, \vec{\alpha}_{u-1})$ is 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}. This probability is determined by [13, Eq. (47)] for $u = d+1, \ldots, \tilde{r}-1$:

$$P_u(R_{d+1}, \vec{\alpha}_{u-2}) \approx (\lambda b_{d+1} R_{d+1})^{u-d-1} \prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} (V_i \, \alpha_i)^{u-1-i}. \tag{74}$$

Note that $P_u(R_{d+1}, \vec{\alpha}_{u-1}) = P_u(R_{d+1}, \vec{\alpha}_{u-2})$, which implies that $P_u(R_{d+1}, \vec{\alpha}_{u-1})$ is not dependent on α_{u-1}.

Substituting (74) and (72) into (73) yields

$$E(Y_{\mathrm{UF}_u}|R_{d+1}, \vec{\alpha}_{u-1}) \approx (\lambda b_{d+1} R_{d+1})^{u-d-1} \left[\prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} (V_i \, \alpha_i)^{u-i}\right]$$
$$\cdot \frac{C}{J} \left(\prod_{j=1}^{d} V_j\right) \tilde{q}_u. \tag{75}$$

Given that the elements of $\vec{\alpha}_{u-1}$ are independent random variables approximately distributed according to (9), such that $E(\alpha_i^k) \approx 1/(k+1)$ for $i \geq d+1$, we have

$$E\left(\prod_{i=d+1}^{u-1} \alpha_i^{u-i}\right) = \prod_{i=d+1}^{u-1} E(\alpha_i^{u-i}) \approx \prod_{i=d+1}^{u-1} \frac{1}{u-i+1} = \frac{1}{(u-d)!} .$$

(76)

Unconditioning (75) on $\vec{\alpha}_{u-1}$ and using (76) yields

$$E(Y_{\mathrm{UF}_u}|R_{d+1}) \approx (\lambda b_{d+1} R_{d+1})^{u-d-1} \left(\prod_{i=d+1}^{u-1} \frac{\tilde{n}_i}{b_i} V_i^{u-i}\right) \frac{1}{(u-d)!}$$
$$\cdot \frac{C}{J} \left(\prod_{j=1}^{d} V_j\right) \tilde{q}_u .$$

(77)

Unconditioning (77) on R_{d+1}, and using (7), (12), and (29) yields (42).

□

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] 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.

[10] 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.

[11] 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.

[12] 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.

[13] I. Iliadis, "Effect of lazy rebuild on reliability of erasure-coded storage systems," in Proc. 15th Int'l Conference on Communication Theory, Reliability, and Quality of Service (CTRQ), Apr. 2022, pp. 1–10.

[14] 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.

[15] I. Iliadis and V. Venkatesan, "Rebuttal to 'Beyond MTTDL: A closed-form RAID-6 reliability equation'," ACM Trans. Storage, vol. 11, no. 2, Mar. 2015, pp. 1–10.

[16] ——, "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.

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

[18] 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.

[19] ——, "Data loss in RAID-5 and RAID-6 storage systems with latent errors," Int'l J. Adv. Software, vol. 12, no. 3&4, Dec. 2019, pp. 259–287.

[20] ——, "Reliability of erasure-coded storage systems with latent errors," Int'l J. Adv. Telecommun., vol. 15, no. 3&4, Dec. 2022, pp. 23–41.

[21] ——, "Reliability evaluation of erasure-coded storage systems with latent errors," ACM Trans. Storage, vol. 19, no. 1, Jan. 2023, pp. 1–47.

[22] 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.

[23] ——, "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.

[24] 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.

[25] Amazon Web Services, "Amazon Simple Storage Service (Amazon S3)," 2022. [Online]. Available: http://aws.amazon.com/s3/ [retrieved: December 7, 2022]

[26] D. Borthakur et al., "Apache Hadoop goes realtime at Facebook," in Proc. ACM Int'l Conference on Management of Data (SIGMOD), Jun. 2011, pp. 1071–1080.

[27] R. J. Chansler, "Data availability and durability with the Hadoop Distributed File System," ;login: The USENIX Association Newsletter, vol. 37, no. 1, Feb. 2012, pp. 16–22.

[28] K. Shvachko, H. Kuang, S. Radia, and R. Chansler, "The Hadoop Distributed File System," in Proc. 26th IEEE Symposium on Mass Storage Systems and Technologies (MSST), May 2010, pp. 1–10.

[29] 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.

[30] A. Chiniah and A. Mungur, "On the adoption of erasure code for cloud storage by major distributed storage systems," EAI Endorsed Transactions on Cloud Systems, vol. 7, no. 21, 2022, pp. 1–11.

[31] V. Venkatesan and I. Iliadis, "Effect of codeword placement on the reliability of erasure coded data storage systems," IBM Research Report, RZ 3827, Aug. 2012.

[32] Y. Li, X. Chen, N. Zheng, J. Hao, and T. Zhang, "An exploratory study on software-defined data center hard disk drives," ACM Trans. Storage, vol. 15, no. 3, May 2019, pp. 1–22.

[33] Seagate, exos x20, data sheet. [Online]. Available: https://www.seagate.com/products/enterprise-drives/exos-x/x20/ [retrieved: March, 2023]

[34] Backblaze drive stats for 2021. [Online]. Available: https://www.backblaze.com/blog/backblaze-drive-stats-for-2021/ [retrieved: March, 2023]