An Efficient Method for Reliability Evaluation of Data Storage Systems

Ilias Iliadis and Vinodh Venkatesan
IBM Research – Zurich
Email: {ili,ven}@zurich.ibm.com

***Abstract*—The effectiveness of the redundancy schemes that have been developed to enhance the reliability of storage systems has predominantly been evaluated based on the mean time to data loss (MTTDL) metric. This metric has been widely used to compare schemes, to assess tradeoffs, and to estimate the effect of various parameters on system reliability. Analytical expressions for MTTDL are typically derived using Markov chain models. Such derivations, however, remain a challenging task owing to the high complexity of the analysis of the Markov chains involved, and therefore the system reliability is often assessed by rough approximations. To address this issue, a general methodology based on the direct path approximation was used to obtain the MTTDL analytically for a class of redundancy schemes and for failure time distributions that also include real-world distributions, such as Weibull and gamma. The methodology, however, was developed for the case of a single direct path to data loss. This work establishes that this methodology can be extended and used in the case where there are multiple shortest paths to data loss to approximately derive the MTTDL for a broader set of redundancy schemes. The value of this simple, yet efficient methodology is demonstrated in several contexts. It is verified that the results obtained for RAID-5 and RAID-6 systems match with those obtained in previous work. As a further demonstration, we derive the exact MTTDL of a given RAID-51 system and confirm that it matches with the MTTDL obtained from the methodology proposed.**

Keywords–Reliability metric; MTTDL; recovery; RAID.

I. INTRODUCTION

Storage systems experience data losses due to device failures, including disk and node failures. To avoid a permanent loss of data, redundancy schemes were developed that enable the recovery of this data. However, during rebuild operations additional device failures may occur that eventually lead to permanent data losses. There is a variety of redundancy schemes that offer different levels of reliability as they tolerate varying degrees of device failures. Each of these schemes is characterized by an overhead, which reflects the additional operations that need to be performed, and a storage efficiency, which expresses the additional amount of data, referred to as parity, that needs to be stored in the system.

The reliability of storage systems and the effectiveness of redundancy schemes have predominantly been assessed based on the mean time to data loss (MTTDL) metric, which expresses the amount of time that is expected to elapse until the first data is irrecoverably lost [1][2]. During this period, failures cause data to be temporarily lost, which is subsequently recovered owing to the redundancy built into the system.

Analytical expressions for the MTTDL are typically derived using Markov chain models [3], which assume that the times to component failures are independent and exponentially distributed. A methodology for obtaining MTTDL under general non-exponential failure and rebuild time distributions, which therefore does not involve any Markov analysis, was presented in [4]. The complexity of these derivations depends on the redundancy schemes and the underlying system configurations considered. The MTTDL metric has been proven useful for assessing tradeoffs, for comparing schemes, and for estimating the effect of various parameters on system reliability [5][6][7]. Analytical closed-form expressions for the MTTDL provide an accurate account of the effect of various parameters on system reliability. However, deriving exact closed-form expressions remains a challenging task owing to the high complexity of the analysis of the Markov chains involved [8][9]. For this reason, the system reliability is instead often assessed by rough approximations. As the direct MTTDL analysis is typically hard, an alternative is performing event-driven simulations [10][11]. However, simulations do not provide insight into how the various parameters affect the system reliability. This article addresses this issue by presenting a simple, yet efficient method, referred to as *shortest path approximation*, to obtain the MTTDL analytically for a broad set of redundancy schemes. It achieves that by considering the most likely paths that lead to data loss, which are the shortest ones. In contrast to simulations, this method provides approximate closed-form expressions for the MTTDL, thus circumventing the inherent complexity of deriving exact expressions using Markov analysis. Note also that this method was previously applied in the context of assessing system unavailability, in particular for systems characterized by large Markov chains [12]. It turns out that this approach agrees with the principle encountered in the probability context expressed by the phrase *"rare events occur in the most likely way"*. This is also demonstrated in [13], where the reliability level of systems that are highly reliable is essentially determined by the so-called "main event", which is the shortest way of failure appearance, that is, along the minimal monotone paths.

In [4][14-17], it was shown that the direct path approximation yields accurate analytical reliability results. To further investigate the validity of the direct-path-approximation method, we apply it to derive the MTTDL results for RAID-5 and RAID-6 systems and subsequently verify that they match with those obtained in previous works [1][2]. In all these previous works though, there is a single direct path to data loss. In contrast, our article is concerned with the case where there are multiple shortest paths to data loss. In this work, we investigate this issue and establish that the direct-path-

 ISBN: 978-1-61208-402-2

approximation method can be extended and also applied in the case of multiple shortest paths and yield accurate reliability results. In particular, we derive the approximate MTTDL of a RAID-51 system using the shortest path approximation. Subsequently, as a demonstration of the validity of the method proposed, we derive the exact MTTDL for a specific instance of a RAID-51 system and confirm that it matches with the corresponding MTTDL obtained using our method.

The remainder of the paper is organized as follows. Section II reviews the general framework for deriving the MTTDL of a storage system. Subsequently, the notion of the direct path to data loss is discussed in Section III, and the efficiency of the direct path approximation is demonstrated in Section IV. Section V discusses the case of multiple shortest paths to data loss and presents the analysis of a RAID-51 system. Finally, we conclude in Section VI.

II. Derivation of MTTDL

A. Markov Analysis

Continuous-time Markov chain (CTMC) models reflecting the system operation can be constructed when the device failures and rebuild times are assumed to be independent and exponentially distributed. Under these assumptions, an appropriate CTMC model can be formulated to characterize the system behavior and capture the corresponding state transitions, including those that lead to data loss. Subsequently, using the infinitesimal generator matrix approach and determining the average time spent in the transient states of the Markov chain yields a closed-form expression for the MTTDL of the system [3]. The results obtained by using CTMC models are often approximate because in practice the times to device failure and the rebuild times are not exponentially distributed. To address this issue, a more general analytical method is required.

B. Non-Markov Analysis

Here we briefly review the general framework for deriving the MTTDL developed in [4][14] using an analytical approach that does not involve any Markov analysis, and therefore avoids the deficiencies of Markov models. The underlying models are not semi-Markov, in that the the system evolution does not depend only on the latest state, but also on the entire path that led to that state. In particular, it depends on the fractions of the data not rebuilt when entering each state. In [18] it was demonstrated that a careless evaluation of these fractions may in fact easily lead to erroneous results.

At any point of time, the system can be thought to be in one of two modes: normal mode and rebuild mode. During normal mode, all data in the system has the original amount of redundancy and there is no active rebuild in process. During rebuild mode, some data in the system has less than the original amount of redundancy and there is an active rebuild process that is trying to restore the redundancy lost. A transition from normal to rebuild mode occurs when a device fails; we refer to the device failure that causes this transition as a *first-device* failure. Following a first-device failure, a complex sequence of rebuild operations and subsequent device failures may occur, which eventually leads the system either to an irrecoverable data loss (DL), with the probability of this event denoted by P_{DL}, or back to the original normal mode by restoring all replicas lost. Typically, the rebuild times are much shorter than the times to failure. Consequently, the time required for this complex sequence of events to complete is negligible compared with the time between successive first-device failures, and therefore can be ignored.

Let T_i be the ith interval of a fully operational period, that is, the time interval from the time t that the system is brought to its original state until a subsequent first-device failure occurs. As the system becomes stationary, the length of T_i converges to T. In particular, for a system comprising N devices with a mean time to failure of a device equal to $1/\lambda$, the expected length of T is given by [4]

$$E(T) := \lim_{i\to\infty} E(T_i) = 1/(N\,\lambda)\ . \tag{1}$$

The notation used is given in Table I. Note that the methodology presented here does not involve any Markov analysis and holds for general failure time distributions, which can be exponential or non-exponential, such as the Weibull and gamma distributions.

As the probability that each first-device failure results in data loss is P_{DL}, the expected number of first-device failures until data loss occurs is $1/P_{\mathrm{DL}}$. Thus, by neglecting the effect of the relatively short transient rebuild periods of the system, the MTTDL is essentially the product of the expected time between two first-device-failure events, $E(T)$, and the expected number of first-device-failure events, $1/P_{\mathrm{DL}}$:

$$\mathrm{MTTDL} \approx \frac{E(T)}{P_{\mathrm{DL}}}\ . \tag{2}$$

Substituting (1) into (2) yields

$$\mathrm{MTTDL} \approx \frac{1}{N\,\lambda\,P_{\mathrm{DL}}}\ . \tag{3}$$

III. Direct Path to Data Loss

As mentioned in Section II, during rebuild mode, some data in the system has less than the original amount of redundancy and there is an active rebuild process that aims at restoring the lost redundancy. The direct path to data loss represents the most likely scenario that leads to data loss. This path considers the smallest number of subsequent device failures that occur while the system is in rebuild mode and lead to data loss.

The direct-path-approximation method was applied in [4][14] and led to an analytical approach that does not involve any Markov analysis, and therefore avoids the deficiencies of Markov models. This approach yields accurate results when the storage devices are highly reliable, that is, when the ratio of the mean rebuild time $1/\mu$ (typically on the order of tens of hours) to the mean time to failure of a device $1/\lambda$ (typically on the order of a few years) is very small:

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

TABLE I. Notation of System Parameters

Parameter	Definition
N	Number of devices in the system or in an array group
$1/\lambda$	Mean time to failure for a device
$1/\mu$	Mean time to rebuild

 ISBN: 978-1-61208-402-2

More specifically, this approach considers the system to be in exposure level e when the maximum number of replicas lost by each of the data is equal to e. Let us consider, for instance, a replication-based storage system, where user data is replicated r times. In this case, the system is in exposure level e if there exists data with $r-e$ copies, but there is no data with fewer than $r-e$ copies. Device failures and rebuild processes cause the exposure level to vary over time. Consider the direct path of successive transitions from exposure level 1 to r. In [14], it was shown that P_{DL} can be approximated by the probability of the direct path to data loss, $P_{\text{DL,direct}}$, when devices are highly reliable, that is,

$$P_{\text{DL}} \approx P_{\text{DL,direct}} = \prod_{e=1}^{r-1} P_{e \to e+1}, \tag{5}$$

where $P_{e \to e+1}$ denotes the transition probability from exposure level e to $e+1$. In fact, the above approximation holds for arbitrary device failure time distributions, and the relative error tends to zero as for highly reliable devices the ratio λ/μ tends to zero [4]. The MTTDL is then obtained by substituting (5) into (3). In [16], the direct path methodology is extended to more general erasure codes, which include RAID systems.

Note that this analysis can also be applied to assess reliability, in terms of the MTTDL, for systems modeled using a CTMC. For instance, in [5], a RAID-5 system that was modeled using a CTMC was analyzed by both a Markov analysis and an approach similar to the general framework. This fact is used in Section IV to compare the MTTDL of RAID systems obtained using the direct path approximation in the context of the general framework with the corresponding MTTDL obtained using Markov analysis of CTMCs. This approach is simpler, in that it circumvents the inherent complexity of deriving exact MTTDL expressions using Markov analysis. In Section V, we demonstrate that the direct-path-approximation method can also be extended and applied in the case of multiple shortest paths. We establish this for a system modeled using a CTMC, and conjecture that this should also hold in the case of non-Markovian systems.

IV. Comparison of Markov Analysis and Direct Path Approximation

A common scheme used for tolerating device (disk) failures is the redundant array of independent disks (RAID) [1][2]. The RAID-5 scheme arranges devices in groups (arrays), each with one redundant device, and can tolerate one device failure per array. Similarly, the RAID-6 scheme arranges devices in arrays, each with two redundant devices, and can tolerate up to two device failures per array. Considering that an array consists of N devices, the storage efficiency of a RAID-5 and RAID-6 system is $(N-1)/N$ and $(N-2)/N$, respectively.

It turns out that the MTTDL of systems comprised of highly reliable devices can be approximated by using the *direct path approximation*. We proceed to demonstrate this by presenting two specific examples, the RAID-5 and RAID-6 systems. In both cases, the RAID array is assumed to contain N devices, and the numbered states of the corresponding Markov models represent the number of failed devices. The DL state represents a data loss due to a device failure that occurs when the system is in the critical mode of operation. A RAID array is considered to be in *critical mode* when an additional device failure can no longer be tolerated. Thus, RAID-5 and RAID-6 arrays are in critical mode when there are $N-1$ devices and $N-2$ devices in operation, that is, when they operate with one device and two devices failed, respectively.

A. MTTDL *for a RAID-5 Array*

The Markov chain model for a RAID-5 array is shown in Fig. 1. When the first device fails, the array enters critical mode, which corresponds to the transition from state 0 to state 1. As initially there are N devices in operation, the mean time until the first failure is equal to $1/(N\lambda)$, and the corresponding transition rate is its inverse, that is, $N\lambda$. Subsequently, the critical mode ends owing to either a successful completion of the rebuild or another device failure. The former event is represented by the state transition from state 1 to state 0 with a rate of μ, given that the mean rebuild time is equal to $1/\mu$. The latter event leads to data loss and is represented by the state transition from state 1 to state DL with a rate of $(N-1)\lambda$ given that in critical mode there are $N-1$ devices in operation.

The exact MTTDL, denoted by $\text{MTTDL}_{\text{RAID-5}}$, is obtained from [5, Eq. (45)] by setting $P_{\text{uf}}^{(1)} = 0$ as follows:

$$\text{MTTDL}_{\text{RAID-5}} \;=\; \frac{\mu + (2N-1)\lambda}{N\,(N-1)\,\lambda^2}\,. \tag{6}$$

Note that when $\lambda \ll \mu$, the first term of the numerator in (6) can be ignored, such that the $\text{MTTDL}_{\text{RAID-5}}$ can be approximated by $\text{MTTDL}_{\text{RAID-5}}^{(\text{approx})}$ as follows:

$$\text{MTTDL}_{\text{RAID-5}}^{(\text{approx})} \;\approx\; \frac{\mu}{N(N-1)\,\lambda^2}\,. \tag{7}$$

This result was obtained in [1] by using an approach that is essentially the direct path approximation. Next, we present this derivation for completeness. The transition from state 0 to state 1 represents the first device failure. The direct path to data loss involves a subsequent device failure prior to completing the rebuild process and returning to state 0. This corresponds to the state transition from state 1 to state DL, with the corresponding probability $P_{1 \to DL}$ given by

$$P_{\text{DL}} = P_{\text{DL,direct}} = P_{1 \to DL} = \frac{(N-1)\,\lambda}{\mu + (N-1)\,\lambda}\,. \tag{8}$$

Substituting (8) into (3) yields

$$\text{MTTDL}'_{\text{RAID-5}} \;\approx\; \frac{\mu + (N-1)\,\lambda}{N(N-1)\,\lambda^2}\,. \tag{9}$$

Note that the approximation given in (7) now follows immediately from (9) by using (4) and therefore neglecting the second term of the numerator.

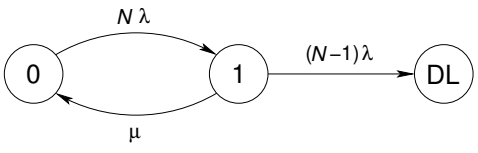

Figure 1. Reliability model for a RAID-5 array.

B. MTTDL *for a RAID-6 Array*

The Markov chain model for a RAID-6 array is shown in Fig. 2. The first device failure is represented by the transition from state 0 to state 1. As initially there are N devices in operation, the mean time until the first failure is $1/(N\lambda)$, and the corresponding transition rate is its inverse, that is, $N\lambda$. The system exits from state 1 owing to either a successful completion of the rebuild or another device failure. The former event is represented by the state transition from state 1 to state 0 with a rate of μ. The latter event is represented by the state transition from state 1 to state 2 with a rate of $(N-1)\lambda$. Subsequently, the system exits from state 2 owing to either a successful completion of the rebuild or another device failure. The former event is represented by the state transition from state 2 to state 0 with a rate of μ, given that the mean rebuild time is equal to $1/\mu$. The latter event leads to data loss and is represented by the state transition from state 2 to state DL with a rate of $(N-2)\lambda$ given that in critical mode there are $N-2$ devices in operation.

The exact MTTDL, denoted by $\text{MTTDL}_{\text{RAID-6}}$, is obtained from [5, Eq. (45)] by setting $\mu_1 = \mu_2 = \mu$ and $P_{\text{uf}}^{(\text{r})} = P_{\text{uf}}^{(2)} = 0$ as follows:

$$\text{MTTDL}_{\text{RAID-6}} \;=\; \frac{\mu^2 + 3(N-1)\lambda\mu + (3N^2 - 6N + 2)\,\lambda^2}{N\,(N-1)\,(N-2)\,\lambda^3}\;. \tag{10}$$

Note that when $\lambda \ll \mu$, the last two terms of the numerator of (10) can be neglected and thus $\text{MTTDL}_{\text{RAID-6}}$ can be approximated by $\text{MTTDL}_{\text{RAID-6}}^{(\text{approx})}$ as follows:

$$\text{MTTDL}_{\text{RAID-6}}^{(\text{approx})} \;\approx\; \frac{\mu^2}{N(N-1)(N-2)\,\lambda^3}\;, \tag{11}$$

which is the same result as that reported in [2].

We now proceed to show how the approximate MTTDL of the system can be derived in a straightforward manner by applying the direct-path-approximation technique. The transition from state 0 to state 1 represents the first device failure. The direct path to data loss involves two subsequent device failures prior to completing the rebuild process and returning to state 0. This corresponds to the state transitions from state 1 to state 2 and from state 2 to state DL, with the corresponding probabilities $P_{1\to 2}$ and $P_{2\to DL}$ given by

$$P_{1\to 2} = \frac{(N-1)\,\lambda}{\mu + (N-1)\,\lambda}\;. \tag{12}$$

and

$$P_{2\to DL} = \frac{(N-2)\,\lambda}{\mu + (N-2)\,\lambda}\;. \tag{13}$$

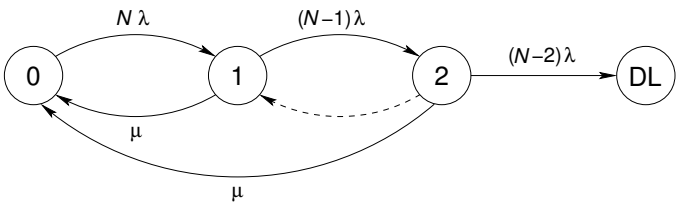

Figure 2. Reliability model for a RAID-6 array.

Thus, the probability of data loss, that is, the probability that from state 1 the system goes to state DL prior to reaching state 0, is equal to

$$\begin{aligned} P_{\text{DL}} = P_{\text{DL,direct}} &= P_{1\to 2}\,P_{2\to DL} \\ &= \frac{(N-1)\,\lambda}{\mu + (N-1)\,\lambda}\;\cdot\;\frac{(N-2)\,\lambda}{\mu + (N-2)\,\lambda} \qquad (14) \\ &\approx \frac{(N-1)(N-2)\,\lambda^2}{\mu^2}\;, \qquad (15) \end{aligned}$$

where the approximation is obtained by using (4) and therefore neglecting the second terms of the denominators in (14).

We now verify that substituting (15) into (3) yields the approximation given in (11).

Remark 1. If the transition from state 2 to state 0 was not to state 0, but was instead to state 1, as shown in Fig. 2 by the dashed arrow, the expression for $P_{2\to DL}$ given by (13) would still hold. However, in this case it would hold that $P_{\text{DL}} > P_{\text{DL,direct}}$ as, in addition to the direct path $1 \to 2 \to \text{DL}$, there are other possible paths $1 \to 2 \to 1 \to 2 \to \cdots \to 1 \to 2 \to \text{DL}$ to data loss. In [14] it was shown that, for highly reliable systems, the direct path dominates the effect of all other possible paths and therefore its probability, $P_{\text{DL,direct}}$, approximates well the probability of all paths, P_{DL}, that is,

$$P_{\text{DL}} \approx P_{\text{DL,direct}} = P_{1\to 2}\,P_{2\to DL} \approx \frac{(N-1)(N-2)\,\lambda^2}{\mu^2}\;. \tag{16}$$

In this case, the MTTDL is given by

$$\text{MTTDL}'_{\text{RAID-6}} \;=\; \frac{(3N^2 - 6N + 2)\,\lambda^2 + 2(N-1)\lambda\mu + \mu^2}{N\,(N-1)\,(N-2)\,\lambda^3}\;, \tag{17}$$

which, as expected, is less than that given in (10). Despite this difference, the approximation given in (11) still holds because (16) is the same as (15),.

V. Multiple Shortest Paths to Data Loss

We now consider redundancy schemes for which there are multiple shortest paths to data loss. Following the analysis presented in [14] for the direct path approximation, we conjecture that, for highly reliable systems, the shortest paths dominate the effect of all other possible paths and therefore the sum of their corresponding probabilities, $P_{\text{DL,shortest}}$, approximates well the probability of all paths, P_{DL}, that is,

$$P_{\text{DL}} \approx P_{\text{DL,shortest}}\;. \tag{18}$$

A. A RAID-51 Array

We proceed by considering a RAID-51 system, which is a RAID-5 array with mirroring. The contents of failed devices are recovered by their mirrors, and if this is not possible, they are recovered through the corresponding RAID-5 arrays. The configuration comprises D pairs of mirrored devices, where each pair contains two devices with identical content. It therefore consists of two identical RAID-5 arrays, for a total of $N\,(=2D)$ devices. This configuration was considered in [9], referred to as RAID 5+1, with the corresponding Markov model shown in [9, Fig. 7(a)]. It is redrawn in Fig. 3 with the parameters λ and μ corresponding to the parameters μ and ν of the initial figure, respectively. Also, the DL states correspond

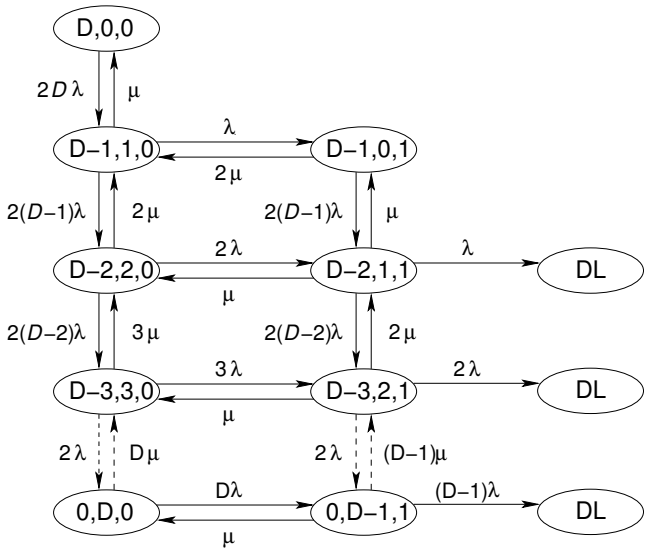

Figure 3. Reliability model for a RAID-51 array.

to the 'Failure' states, and the state tuples (x, y, z) indicate that there are x pairs with both devices in operation, y pairs with one device in operation and one device failed, and z pairs with both devices failed. Also, some typos regarding the transition rates were corrected.

An exact evaluation of the MTTDL associated with this Markov chain model appears to be a very challenging, if not infeasible, task. Thus, in [9] a rough approximation was obtained by first deriving the failure and repair rates for a mirrored pair of devices, and then substituting these values into expression (7) for a single RAID-5 system. The MTTDL is obtained in [9, Eq. (11)] as follows:

$$\text{MTTDL} \approx \frac{\mu^3}{4D(D-1)\lambda^4} . \qquad (19)$$

B. MTTDL *Evaluation Using the Shortest Path Approximation*

The transition from $(D, 0, 0)$ to state $(D-1, 1, 0)$ represents the first device failure. As initially there are $2\,D$ devices in operation, the mean time until the first failure is $1/(2\,D\lambda)$, and the corresponding transition rate is its inverse, $2\,D\lambda$.

The most likely path to data loss is the shortest path from state $(D-1, 1, 0)$ to a DL state, which in this case comprises two such paths, as shown in Fig. 4: the upper path $(D-1, 1, 0) \rightarrow (D-1, 0, 1) \rightarrow (D-2, 1, 1) \rightarrow$ DL and the lower path: $(D-1, 1, 0) \rightarrow (D-2, 2, 0) \rightarrow (D-2, 1, 1) \rightarrow$ DL. Each of these paths involves three subsequent device failures.

After the first device has failed, there are $D-1$ pairs with both devices in operation, and one pair, say PR_1, with one device in operation and one device failed, which corresponds to the transition from state $(D, 0, 0)$ to state $(D-1, 1, 0)$. The rebuild of the failed device consists of recovering its data to a new spare device by copying the contents of its mirror to it, that is, of the device in operation in PR_1. Then, the next event can be either a successful completion of the rebuild or another device failure. The former event is represented by the

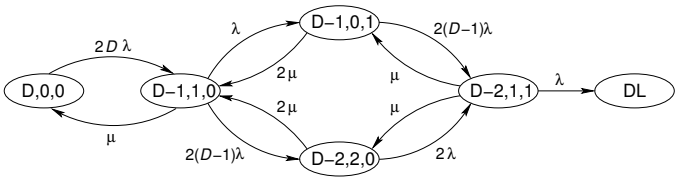

Figure 4. Shortest-path reliability model for a RAID-51 array.

state transition from state $(D-1, 1, 0)$ to state $(D, 0, 0)$ with a rate of μ. For the latter event, two cases are considered:

Case a) The second device that fails is the device in operation concerning pair PR_1, which corresponds to the transition from state $(D-1, 1, 0)$ to state $(D-1, 0, 1)$, as both devices of pair PR_1 have failed, and all other $D-1$ pairs remain intact. The transition rate is λ, which is the failure rate of the last failed device. The contents of the devices of pair PR_1 are recovered through the corresponding RAID-5 arrays. As both devices of pair PR_1 are under rebuild, the transition rate from state $(D-1, 0, 1)$ back to state $(D-1, 1, 0)$ is $2\,\mu$. If, however, prior to the completion of any of the two rebuilds another device of the remaining $2\,(D-1)$ fails, then there will be $D-2$ pairs with both devices in operation, one pair, say PR_2, with one device in operation and one device failed, and pair PR_1 with both devices failed. This corresponds to the transition from state $(D-1, 0, 1)$ to state $(D-2, 1, 1)$ with a corresponding transition rate equal to $2\,(D-1)\,\lambda$. Note that in [9, Fig. 7(a)] this transition rate is erroneously indicated as $(2\,D-1)\,\mu$ instead of $2\,(D-1)\,\mu$.

Case b) The second device that fails is one of the $2\,(D-1)$ devices in the $D-1$ pairs, say a device concerning PR_2. This corresponds to the transition from state $(D-1, 1, 0)$ to state $(D-2, 2, 0)$, as both pairs PR_1 and PR_2 have one device in operation and one device failed, and all other $D-2$ pairs remain intact. The corresponding transition rate is equal to $2\,(D-1)\,\lambda$. Note that in [9, Fig. 7(a)] this transition rate is erroneously indicated as $(2\,D-1)\,\mu$ instead of $2\,(D-1)\,\mu$. The contents of the failed devices are recovered from their corresponding mirrors. As both devices of the two pairs PR_1 and PR_2 are under rebuild, the transition rate from state $(D-2, 2, 0)$ back to state $(D-1, 1, 0)$ is $2\,\mu$. If, however, prior to the completion of any of the two rebuilds another device of the two remaining devices in operation in PR_1 and PR_2 fails (say, that of pair PR_1), then there will be $D-2$ pairs with both devices in operation, one pair (PR_2) with one device in operation and one device failed, and one pair (PR_1) with both devices failed. This corresponds to the transition from state $(D-2, 2, 0)$ to state $(D-2, 1, 1)$, with a corresponding transition rate $2\,\lambda$.

At state $(D-2, 1, 1)$, the device in pair PR_2 failed is recovered by its mirror. However, the corresponding failed device in pair PR_1 cannot be recovered because the RAID-5 array has suffered two device failures. In contrast, the other failed device in pair PR_1 can by recovered because the corresponding RAID-5 array has suffered only one device failure.

The completion of the rebuild of the failed device in pair PR_2 corresponds to the transition from state $(D-2, 1, 1)$ to state $(D-1, 0, 1)$, with a transition rate of μ. The completion of

the rebuild of the failed device in pair PR_1 through the RAID capability corresponds to the transition from state $(D-2,1,1)$ to state $(D-2,2,0)$, with a transition rate of μ. Note that in [9, Fig. 7(a)] this transition rate is erroneously indicated as $2\,\mu$ instead of μ. If, however, prior to the completion of any of these rebuilds, the device still in operation of pair PR_2 fails, this leads to data loss, as there will be two pairs failed, with each of the RAID-5 arrays having two devices failed. This corresponds to the transition from state $(D-2,1,1)$ to state DL, with a corresponding rate of λ.

The probabilities of the transitions discussed above are given by

$$P_{(D-1,1,0)\to(D-1,0,1)} = \frac{\lambda}{\mu+(2D-1)\,\lambda}\,, \tag{20}$$

$$P_{(D-1,0,1)\to(D-2,1,1)} = \frac{2(D-1)\,\lambda}{2\,\mu+2(D-1)\,\lambda}\,, \tag{21}$$

$$P_{(D-1,1,0)\to(D-2,2,0)} = \frac{2(D-1)\,\lambda}{\mu+(2D-1)\,\lambda}\,, \tag{22}$$

$$P_{(D-2,2,0)\to(D-2,1,1)} = \frac{2\,\lambda}{2\,\mu+2\,\lambda}\,, \tag{23}$$

and

$$P_{(D-2,1,1)\to\mathrm{DL}} = \frac{\lambda}{2\,\mu+\lambda}\,. \tag{24}$$

Consequently, the probability of the upper path to data loss, P_u, is given by

$$\begin{aligned} P_u =& P_{(D-1,1,0)\to(D-1,0,1)}\,P_{(D-1,0,1)\to(D-2,1,1)}\,P_{(D-2,1,1)\to\mathrm{DL}} \\ =& \frac{\lambda}{\mu+(2D-1)\,\lambda}\cdot\frac{2(D-1)\,\lambda}{2\,\mu+2(D-1)\,\lambda}\cdot\frac{\lambda}{2\,\mu+\lambda}\,, \end{aligned} \tag{25}$$

and that of the lower path to data loss, P_l, is given by

$$\begin{aligned} P_l =& P_{(D-1,1,0)\to(D-2,2,0)}\,P_{(D-2,2,0)\to(D-2,1,1)}\,P_{(D-2,1,1)\to\mathrm{DL}} \\ =& \frac{2(D-1)\,\lambda}{\mu+(2D-1)\,\lambda}\cdot\frac{2\,\lambda}{2\,\mu+2\,\lambda}\cdot\frac{\lambda}{2\,\mu+\lambda}\,, \end{aligned} \tag{26}$$

By considering (4), (25) and (26) yield the following approximations:

$$P_u \;\approx\; \frac{\lambda}{\mu}\cdot\frac{2(D-1)\,\lambda}{2\,\mu}\cdot\frac{\lambda}{2\,\mu} \;=\; \frac{(D-1)\,\lambda^3}{2\mu^3} \tag{27}$$

and

$$P_l \;\approx\; \frac{2(D-1)\,\lambda}{\mu}\cdot\frac{\lambda}{\mu}\cdot\frac{\lambda}{2\,\mu} \;=\; \frac{(D-1)\,\lambda^3}{\mu^3}\,. \tag{28}$$

The probability of the shortest paths to data loss, $P_{\mathrm{DL,shortest}}$, is the sum of P_u and P_l, which by using (18), (27), and (28), yields

$$P_{\mathrm{DL}} \;\approx\; P_{\mathrm{DL,shortest}} = P_u + P_l \;\approx\; \frac{3(D-1)\,\lambda^3}{2\mu^3}\,. \tag{29}$$

Substituting (29) into (3), and considering $N=2D$, yields the approximate MTTDL of the RAID-51 system, $\mathrm{MTTDL}^{(\mathrm{approx})}_{\text{RAID-51}}$ given by

$$\mathrm{MTTDL}^{(\mathrm{approx})}_{\text{RAID-51}} \;\approx\; \frac{\mu^3}{3D(D-1)\,\lambda^4}\,. \tag{30}$$

Remark 2. Note that the prediction given by (30) is higher than that obtained in [9], which is given by (19). At first glance, this seems to be counterintuitive. The approximation in [9] considers only failures of mirrored pair of devices, which corresponds to the upper path to data loss. As this neglects the lower path, one would expect the prediction in [9] to be higher, not lower. The reason for this counterintuitive result is the fact that considering additional paths on the one hand may increase the number of paths that lead to data loss, but on the other hand it may also increase the number of the paths that do not lead to data loss, therefore delaying the occurrence of data loss. For instance, when the lower path is neglected, the probability $P_{(D-2,1,1)\to\mathrm{DL}}$ of the transition from state $(D-2,1,1)$ to state DL is equal to $\lambda/(\lambda+\mu)$, which is greater than the corresponding one given by (24), if also the lower path is considered.

C. *Exact* MTTDL *Evaluation for D=3*

An exact evaluation of the reliability of a RAID-51 system through the MTTDL associated with the corresponding Markov chain model shown in Fig. 3 appears to be a very challenging, if not infeasible, task for arbitrary D. We therefore proceed by considering a RAID-51 system with $D=3$. The corresponding Markov chain model is shown in Fig. 5. The exact MTTDL of this system, denoted by $\mathrm{MTTDL}^{(D=3)}_{\text{RAID-51}}$, is obtained by using the infinitesimal generator matrix approach and determining the average time spent in the transient states of the Markov chain [3] Because of space limitations, we only provide the final result:

$$\mathrm{MTTDL}^{(D=3)}_{\text{RAID-51}} = \frac{2+20\frac{\lambda}{\mu}+93(\frac{\lambda}{\mu})^2+287(\frac{\lambda}{\mu})^3+677(\frac{\lambda}{\mu})^4+939(\frac{\lambda}{\mu})^5+630(\frac{\lambda}{\mu})^6}{12\;\lambda^4\;\mu^{-3}\;[3+18\frac{\lambda}{\mu}+35(\frac{\lambda}{\mu})^2+30(\frac{\lambda}{\mu})^3]}\,. \tag{31}$$

Note that when $\lambda \ll \mu$, $\mathrm{MTTDL}^{(D=3)}_{\text{RAID-51}}$ can be approximated by $\mathrm{MTTDL}^{(D=3,\mathrm{approx})}_{\text{RAID-51}}$ as follows:

$$\mathrm{MTTDL}^{(D=3,\mathrm{approx})}_{\text{RAID-51}} \;\approx\; \frac{\mu^3}{18\,\lambda^4}\,, \tag{32}$$

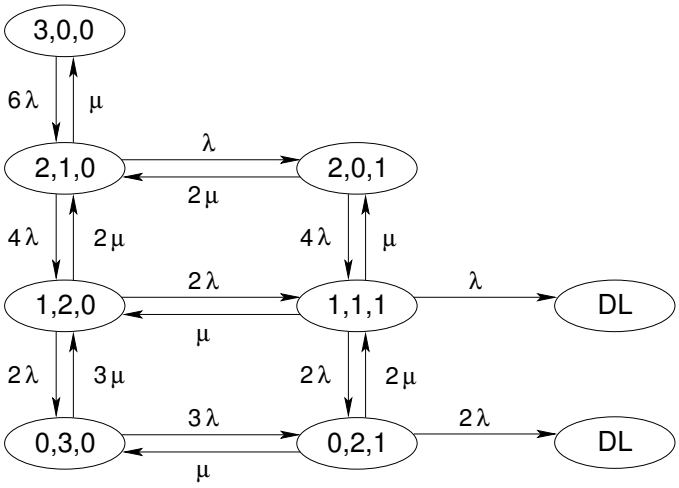

Figure 5. Reliability model for a RAID-51 array with $D=3$.

which is the same result as that predicted by (30) for $D = 3$, and therefore confirms its validity.

VI. CONCLUSIONS

We considered the mean time to data loss (MTTDL) metric, which assesses the reliability level of storage systems. This work presented a simple, yet efficient methodology to approximately assess it analytically for highly reliable systems and a broad set of redundancy schemes. We extended the direct path approximation to a more general method that considers all shortest paths that lead to data loss. We subsequently applied this method to obtain a closed-form expression for the MTTDL of a RAID-51 system. We also considered a specific instance of a RAID-51 system, then derived the corresponding exact MTTDL, and subsequently confirmed that it matches that obtained from the shortest-path-approximation method. As the direct path approximation accurately predicts the reliability of non-Markovian systems with a single shortest path, we conjecture that the shortest-path-approximation method would also accurately predict the reliability of non-Markovian systems with multiple shortest paths.

Application of the shortest-path-approximation methodology developed to derive the MTTDL for systems using other redundancy schemes, such as erasure codes, is a subject of future work.

REFERENCES

[1] D. A. Patterson, G. Gibson, and R. H. Katz, "A case for redundant arrays of inexpensive disks (RAID)," in Proceedings of the ACM SIGMOD International Conference on Management of Data, 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] K. S. Trivedi, Probabilistic and Statistics with Reliability, Queueing and Computer Science Applications, 2nd ed. New York: Wiley, 2002.

[4] V. Venkatesan and I. Iliadis, "A general reliability model for data storage systems," in Proceedings of the 9th International Conference on Quantitative Evaluation of Systems (QEST), Sep. 2012, pp. 209–219.

[5] 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, May 2008, pp. 1–42.

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

[7] 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, Jul. 2011, pp. 1–42.

[8] K. Rao, J. L. Hafner, and R. A. Golding, "Reliability for networked storage nodes," IEEE Trans. Dependable Secure Comput., vol. 8, no. 3, May 2011, pp. 404–418.

[9] Q. Xin, E. L. Miller, T. J. E. Schwarz, D. D. E. Long, S. A. Brandt, and W. Litwin, "Reliability mechanisms for very large storage systems," in Proceedings of the 20th IEEE/11th NASA Goddard Conference on Mass Storage Systems and Technologies (MSST), Apr. 2003, pp. 146–156.

[10] Q. Xin, T. J. E. Schwarz, and E. L. Miller, "Disk infant mortality in large storage systems," in Proceedings of the 13th Annual IEEE/ACM International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Sep. 2005, pp. 125–134.

[11] A. Wildani, T. J. E. Schwarz, E. L. Miller, and D. D. E. Long, "Protecting against rare event failures in archival systems," in Proceedings of the 17th Annual IEEE/ACM International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Sep. 2009, pp. 1–11.

[12] M. Bouissou and Y. Lefebvre, "A path-based algorithm to evaluate asymptotic unavailability for large markov models," in Proceedings of the 48th Annual Reliability and Maintainability Symposium, Jan. 2002, pp. 32–39.

[13] I. B. Gertsbakh, "Asymptotic methods in reliability theory: A review," Advances in Applied Probability, vol. 16, no. 1, Mar. 1984, pp. 147–175.

[14] V. Venkatesan, I. Iliadis, C. Fragouli, and R. Urbanke, "Reliability of clustered vs. declustered replica placement in data storage systems," in Proceedings of the 19th Annual IEEE/ACM International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Jul. 2011, pp. 307–317.

[15] V. Venkatesan, I. Iliadis, and R. Haas, "Reliability of data storage systems under network rebuild bandwidth constraints," in Proceedings of the 20th Annual IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Aug. 2012, pp. 189–197.

[16] V. Venkatesan and I. Iliadis, "Effect of codeword placement on the reliability of erasure coded data storage systems," in Proceedings of the 10th International Conference on Quantitative Evaluation of Systems (QEST), Sep. 2013, pp. 241–257.

[17] ——, "Effect of latent errors on the reliability of data storage systems," in Proceedings of the 21th Annual IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Aug. 2013, pp. 293–297.

[18] I. Iliadis and V. Venkatesan, "Expected annual fraction of data loss as a metric for data storage reliability," in Proceedings of the 22nd Annual IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Sep. 2014, pp. 375–384.

 ISBN: 978-1-61208-402-2