Validation of Approximate Dependability Models of a RAID Architecture with Orthogonal Organization*

Juan A. Carrasco
Departament d'Enginyeria Electrònica
Universitat Politècnica de Catalunya
Diagonal 647, plta. 9
08028 Barcelona, Spain
carrasco@eel.upc.es

Abstract

RAID (Redundant Array of Inexpensive Disks) are widely used in storage servers. Level-5 RAID is one of the most popular RAID architectures. Numerical analysis of exact Markovian dependability models of level-5 RAID architectures with orthogonal organization is unfeasible for many realistic model parameters due to the size of the resulting state space. In this paper we develop approximate dependability models for a level-5 RAID architecture with orthogonal organization which have small state spaces. We consider two measures: the steady-state unavailability and the unreliability. The models encompass disk hot spares and imperfect disk reconstruction. Using bounding techniques we analyze the accuracy of the models and show that the models are extremely accurate.

1. Introduction

Increased performance in processors and memory subsystems have made the I/O subsystem the bottleneck of computer systems in many applications. Dependability of the I/O subsystem is also a major concern in many applications such as on-line transaction processing. RAID (Redundant Array of Inexpensive Disks) provide an inexpensive valuable alternative to improve the performance and dependability of the I/O subsystem [11]. RAID can be composed of a large number of disks and incorporate redundancy to achieve high dependability. Several organizations, called levels, are available with different degrees of redundancy. The first RAID taxonomy [10] included levels 1 through 5, but since then, levels 0 and 6 have also been widely accepted [3]. Level 5 is a popular RAID architecture. Many studies about the performance of that architecture are available, e.g. [6, 7, 15]. A number of studies have dealt with the dependability of RAID. Most of them have considered an approximate computation of the mean time to failure [4, 5, 9, 10, 14]. More detailed dependability models [13] have been developed recently using Stochastic Activity Networks [12]. This paper is concerned with dependability modeling of level-5 RAID architectures with orthogonal organizations using continuous-time Markov chains (CTMCs).

Exact CTMC dependability models of level-5 RAID architectures with orthogonal organization are complex and have enormous state spaces when the RAID contains a large number of disks. In this paper we develop approximate CTMC dependability models for such level-5 RAID architectures and analyze their accuracy. We consider unavailability and reliability models. The models generalize the reliability model proposed in [13] by encompassing hot spare controllers and yield small state spaces. The accuracy of the models is assessed analytically by using CTMC models with much greater state spaces yielding bounds for the exact dependability. It is found that the approximate models are very accurate, yielding about 6 significant digits. We want to emphasize that validation of such a degree of accuracy could not have been obtained using simulation. The rest of the paper is organized as follows. In Section 2 we describe the approximate models and discuss intuitively why they should be quite accurate. In Section 3 we describe the bounding CTMC models and using them analyze the accuracy of the approximate models. Section 4 summarizes the main conclusions.

2. Approximate Dependability Models

Figure 1 shows the architecture of the considered level-5 RAID system. The system includes $G \times N$ disks and N

*This work has been supported by the "Comisión Interministerial de Ciencia y Tecnología" (CICYT) under the research grant TAP99-0443-C05-05.

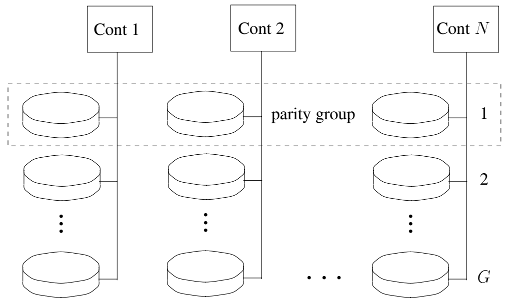

Figure 1. Architecture of the considered level-5 RAID system.

controllers. The disks are organized in G parity groups, each with N disks. Each controller controls a string of G disks. The system also includes C_{H} hot spare controllers and D_{H} hot spare disks. The system is operational if there is access to at least $N-1$ available disks of each parity group. When there is a failed controller all disks of the associated string become unavailable. When a failed disk is replaced by a good one and if all disks of the parity group are available, the parity group starts the reconstruction of data in the replaced disk. The reconstruction process also starts when a disk of a parity group which was not available due to failure of one controller becomes available due to the replacement of the failed controller. All disks of the parity group involved in a reconstruction are "overloaded" and have a higher failure rate. Non-overloaded disks fail with rate λ_{D}. Overloaded disks fail with rate λ_{S}. Controllers fail with rate λ_{C}. The reconstruction process has an exponential duration with rate μ_{DRC}. Failed disks and controllers are replaced, if respective hot spares are available, by a repairman with rates μ_{DRP} and μ_{CRP}, respectively, with priority given to controllers and random selection among the failed components of highest priority. Lacking spares and failed disks and controllers for which there are not spares are replaced with rate μ_{SR} by an unlimited number of repairmen. A reconstruction process is successful with probability P_{R}. Failure in a reconstruction process causes the failure of the system. Finally, when the system is failed, it is returned to its original state, with all disks and hot spares available, by a global repair action which has rate μ_{G}.

The exact previously described model gives CTMCs with very large state spaces even for moderate values of G and N. The reason is that in order to know how many controller failures cause a system failure it is necessary to know whether the unavailable disks are aligned (belong to the same string) or not, since in the first case $N-1$ controller failures will cause system failure, and in the second case all N controller failures will cause system failure. To know whether the unavailable disks are aligned or not in the presence of repairs it is necessary to know how many disks are unavailable in each string, and such a detailed state description gives raise to an exploition in the size of the state space when G and N increase. In this section we will describe approximate models which have much smaller state spaces. The approximation consists in assuming, pessimistically, that if unavailable disks are unaligned they will remain unaligned when one of them becomes available whenever the number of remaining unavailable disks is ≥ 2. Using that approximation it is possible to describe the state of the CTMC using the following state variables: NFD (number of failed disks), NDR (number of disks under reconstruction), NWD (number of disks waiting for reconstruction), NSD (number of hot spare disks), AL (a boolean variable which is true when unavailable disks are aligned and false otherwise), NFC (number of failed controllers), NSC (number of hot spare controllers), and F (a boolean variable which is true when the system is failed and false otherwise). Such a succinct description gives state spaces whose size is independent on N (≥ 2) and grows moderately with G. Figures 2–4 give a formal description of the approximate model for the steady-state unavailability using the model specification language of METFAC-2.1 [1]. State changes are specified by production rules which model actions and have associated rates. In addition, an action may have several responses, which may have associated probabilities. The CTMC is generated from the start state by applying production rules. The `reward` construct assigns reward rates to states of the CTMC. The specification sets the reward rate of the down state (`F == yes`) to 1 and the reward rate of the up states to 0. With that reward rate structure, the generic measure "Expected Steady-State Reward Rate" is the steady-state unavailability of the system. The specification required to compute the unreliability at time t is obtained from the one given in the figures by removing the functional parameter `MUG` and the production rule modeling the global repair action, and associating a reward rate 1 with the system up states. With that reward rate structure, the generic measure "Cumulative Reward Distribution Till Exit of a subset of states" (entry in the absorbing state) is the system unreliability at time t.

Intuitively, it is clear that the approximate models should be quite accurate. The reason is that they differ from the exact models only following repair from states in which there are non-aligned unavailable disks, which will be rare. However, how good is the approximation? That question will be answered in the following section. We want to note that the approximate model gives CTMCs with small state spaces. To illustrate the point we give in Table 1 the number of states and number of transitions for the model used in the computation of the steady-state unavailability for several values of G, C_{H} and D_{H}.

```
parameters
int
   G,  /* number of parity groups */
   CH, /* number of controller hot spares */
   DH, /* number of disk hot spares */

double
   N,     /* number of disks in parity group */
   LD,    /* failure rate of non-overloaded disk */
   LS,    /* failure rate of overloaded disk */
   LC,    /* controller failure rate */
   MUDRC, /* disk reconstruction rate */
   MUDRP, /* rate at which a failed disk is replaced by a hot spare */
   MUCRP, /* rate at which a failed controller is replaced by a hot spare */
   MUSR,  /* rate at which a lacking spare or a failed component when */
          /* there are not spares is replaced */
   MUG,   /* global repair rate when system is failed */
   PR     /* probability that a disk reconstructions is sucessfull */

state_variables
   NFD,  /* number of failed disks */
   NDR,  /* number of disks under reconstruction */
   NWD,  /* number of disks waiting for reconstruction */
   NSD,  /* number of hot spare disks */
   AL,   /* YES if unavailable disks are aligned */
   NFC,  /* number of failed controllers */
   NSC,  /* number of hot spare controllers */
   F     /* YES if system is failed */

production_rules

/* Failure of a non-overloaded disk in a group which has no disk failed */
/* when no controller is failed  */
if !F && NFC == 0 && NFD+NDR < G action NOD_F with_rate (G-NFD-NDR)*N*LD
   if NFD+NDR == 0 response ALIGNED new_state NFD=NFD+1, AL=yes

   if NFD+NDR > 0 && AL response ALIGNED with_prob 1/N new_state NFD=NFD+1

   if NFD+NDR > 0 && AL response NON_ALIGNED with_prob (N-1)/N new_state NFD=NFD+1, AL=no

   if NFD+NDR > 0 && !AL response NON_ALIGNED new_state NFD=NFD+1
end

/* Failure of a non-overloaded disk in a group which has one disk failed */
/* when no controller is failed */
if !F && NFC == 0 && NFD > 0 action NOD_F with_rate NFD*(N-1)*LD
new_state  F=yes, NFD=0, NDR=0, NWD=0, NSD=0, AL=yes, NFC=0, NSC=0

/* Failure of an overloaded disk when no controller is failed */
if !F && NFC == 0 && NDR > 0 action OD_F with_rate NDR*N*LS
new_state F=yes, NFD=0, NDR=0, NWD=0, NSD=0, AL=yes, NFC=0, NSC=0
```

Figure 2. Formal description of the approximate model using the METFAC-2.1 model specification language.

Table 1. Size of the resulting CTMC for the approximate model for the steady-state unavailability for several values of G, C_{H} and D_{H}.

G	C_{H}	D_{H}	states	transitions
5	1	2	271	1,464
5	2	3	541	3,037
10	1	2	841	5,009
10	2	3	1,681	10,427
20	1	2	2,881	18,249
20	2	3	5,761	38,107

```
/* Failure of a non-aligned disk when one controller is failed */
if !F && NFC == 1 action NOD_F with_rate G*(N-1)*LD
new_state F=yes, NFD=0, NDR=0, NWD=0, NSD=0, AL=yes, NFC=0, NSC=0

/* Failure of an aligned disk when one controller is failed */
if !F && NFC == 1 && NWD > 0 action NOD_F with_rate NWD*LD
new_state NFD=NFD+1, NWD=NWD-1

/* First controller failure */
if !F && NFC == 0 action C_F with_rate N*LC
   if NFD+NDR == 0 response OK new_state  NFC=1, NWD=G

   if NFD+NDR > 0 && AL response OK with_prob 1/N new_state NDR=0, NWD=G-NFD, NFC=1

   if NFD+NDR > 0 && AL response FAILURE with_prob (N-1)/N
   new_state F=yes, NFD=0, NDR=0, NWD=0, NSD=0, AL=yes, NFC=0, NSC=0

   if NFD+NDR > 0 && !AL response FAILURE
   new_state F=yes, NFD=0, NDR=0, NWD=0, NSD=0, AL=yes, NFC=0, NSC=0

end

/* Second controller failure */
if !F && NFC == 1 action C_F with_rate (N-1)*LC
new_state F=yes, NFD=0, NDR=0, NWD=0, NSD=0, AL=yes, NFC=0, NSC=0

/* Global repair */
if F action G_R with_rate MUG
new_state NFD=0, NDR=0, NWD=0, NSD=DH, AL=yes, NFC=0, NSC= CH, F=no

/* Replacement of a failed controller by a hot spare */
if !F && NFC == 1 && NSC > 0 action C_R with_rate MUCRP
new_state NDR=NWD, NWD=0, NFC=0, NSC=NSC-1

/* Replacement of a failed controller from the maintenance service */
if !F && NFC == 1 && NSC == 0 action C_R with_rate MUSR
new_state NDR=NWD, NWD=0, NFC=0

/* Replacement of a failed disk by a hot spare */
if !F && (NFC == 0 || NSC == 0) && NFD > 0 && NSD > 0
action D_R with_rate MUDRP
   if NFC == 0 response REC new_state NFD=NFD-1, NDR=NDR+1, NSD=NSD-1

   if NFC == 1 response WAIT new_state NFD=NFD-1, NWD=NWD+1, NSD=NSD-1
end

/* Replacement of a failed disk from the maintenance service */
if !F && NFD > NSD action D_R with_rate (NFD-NSD)*MUSR
   if NFC == 0 response REC new_state  NFD=NFD-1,NDR=NDR+1

   if NFC == 1 response WAIT new_state NFD=NFD-1, NWD=NWD+1
end
```

Figure 3. Formal description of the approximate model using the METFAC-2.1 model specification language (continuation).

```
/* End of reconstruction of a disk */
if !F && NFC == 0 && NDR > 0 action D_RC with_rate NDR*MUDRC
   if AL response OK with_prob PR new_state NDR=NDR-1

   if !AL && NFD+NDR > 2 response OK with_prob PR new_state NDR=NDR-1

   if !AL && NFD+NDR == 2 response OK with_prob PR new_state NDR=NDR-1, AL=yes

   response FAILURE with_prob 1-PR
   new_state F=yes, NFD=0, NDR=0, NWD=0, NSD=0, AL=yes, NFC=0, NSC=0

end

/* Replacement of hot spare controller */
if !F && NSC < CH action SPARE_C with_rate (CH-NSC)*MUSR
new_state NSC=NSC+1

/* Replacement of hot spare disk */
if !F && NSD < DH action SPARE_D with_rate (DH-NSD)*MUSR
new_state NSD=NSD+1

start_state NFD=0, NDR=0, NWD=0, NSD=DH, AL=yes, NFC=0, NSC=CH, F=no

reward_rate (double)(F == yes)
```

Figure 4. Formal description of the approximate model using the METFAC-2.1 model specification language (continuation).

3. Model Validation

In this section we validate the approximate models described in the previous section in the sense of showing that, for typical model parameter values, they are highly accurate. As already pointed out, the exact models yield CTMCs with have unmanageable state spaces and, thus, validation cannot be accomplished by comparing the approximate models with the exact models. Instead, we will use bounding techniques to get the solution of the exact models with a high enough degree of accuracy. The advantage of using bounding techniques is that the size of the CTMCs which have to be solved is substantially smaller than the size of the CTMCs resulting from the exact models.

3.1 Bounding model for the steady-state unavailability

We start by providing a brief review of the bounding technique we will use to obtain bounds for the steady-state unavailability. The technique uses theoretical results recently obtained [2, 8]. Let $X = \{X(t); t \geq 0\}$ be the CTMC under consideration and denote by S its state space. Let $\lambda_{i,j}$, $i, j \in S$, $j \neq i$ denote the transition rate of X from state i to state j, let $\lambda_i = \sum_{j \in S-\{i\}} \lambda_{i,j}$ denote the output rate of X from state i, and, being B a subset of S, let $\lambda_{i,B} = \sum_{j \in B} \lambda_{i,j}$. The bounding technique requires the definition of a partition $\cup_{0 \leq k \leq N} C_k$ of the state space S and works well when transition rates to the right, i.e. from C_k to $\cup_{k+1 \leq l \leq N} C_l$ are relatively small compared with the transition rates to the left, i.e. from C_k to $\cup_{0 \leq l \leq k-1} C_l$. The technique as reviewed here requires that C_0 has a single state r. Bounds are computed using a CTMC $Y = \{Y(t); t \geq 0\}$ with state space $G \cup \{c_1, c_2, \ldots, c_N\}$, $G = \cup_{k=0}^{K} C_k$, where K is a bounding parameter, typically much smaller than N. The accuracy of the bounds increases with K. Transition rates in Y between states of G are as in X. Y has transition rates from states $i \in G$ to states c_k with rates λ_{i,C_k}. In addition, Y has transition rates $f_{k,l}^{+}$ from c_k to c_l, $l > k$, transition rates g_k^{-} from c_k to c_{k-1}, $k \geq 2$, and a transition rate g_1^{-} from c_1 to r. The transition rate $f_{k,l}^{+}$ has to upper bound λ_{i,C_l}, $i \in C_k$, $l > k$. If all states in C_k have transition to the left we can take a $g_k^{-} > 0$ lower bounding $\lambda_{i,\cup_{0 \leq l \leq k-1} C_l}$, $i \in C_k$. We can also take a $g_k^{-} > 0$ lower bounding $\min_{i \in C_K} q_i/h_i$, where q_i is the probability that X will exit C_k by the left assuming entry in C_k through i and h_i is the mean holding time in C_k assuming entry through i. Using Y, we can obtain a lower and an upper bound for the steady-state reward rate of X by assigning to the states in G the same reward rate as in X and for the states c_k, respectively, a lower bound and an upper bound for the reward rates of X. The steady-state reward rate obtained for Y using the lower bound reward rate in c_k is the desired lower bound and the steady-state reward rate obtained for Y using the upper bound reward rate in c_k is the desired upper bound. For the steady-state unavailability, convenient lower and upper bound reward rates to be assigned to the states c_k are 0 and 1, respectively.

To apply the previous bounding technique we have to select a suitable partition of S. For the bounds to be good with moderate values of K and, thus, with moderate state spaces of Y, the partition should be chosen so that states belonging to subsets C_k with increasing k be progressively rarer; in addition transition rates to the right should be much smaller than transition rates to the left. Let U be the subset of up states of S. $S = U \cup \{f\}$, where f is the single system failed state. For a state $i \in U$, let $N_{\mathrm{FC}}(i)$ be the number of failed controllers, let $N_{\mathrm{FD}}(i)$ be the number of failed disks, let $N_{\mathrm{DR}}(i)$ be the number of disks under reconstruction, let $N_{\mathrm{WD}}(i)$ be the number of disks waiting for reconstruction, let $N_{\mathrm{SC}}(i)$ be the number of hot spare controllers, and let $N_{\mathrm{SD}}(i)$ be the number of hot spare disks, all in i. At first glance, a reasonable partition is $C_0 = \{r\}$, $C_1 = \{f\} \cup \{i \in U \ : \ N_{\mathrm{FC}}(i) + N_{\mathrm{FD}}(i) + N_{\mathrm{DR}}(i) + N_{\mathrm{WD}}(i) + (C_{\mathrm{H}} - N_{\mathrm{SC}}(i)) + (D_{\mathrm{H}} - N_{\mathrm{SD}}(i)) = 1\}$, and $C_k = \{i \in U \ : \ N_{\mathrm{FC}}(i) + N_{\mathrm{FD}}(i) + N_{\mathrm{DR}}(i) + N_{\mathrm{WD}}(i) + (C_{\mathrm{H}} - N_{\mathrm{SC}}(i)) + (D_{\mathrm{H}} - N_{\mathrm{SD}}(i)) = k\}$, $k > 1$. However that "natural" partition has two problems. The first one is that the replacement of a failed disk with a hot spare, which has a fast rate, originates transitions to the right ($D_{\mathrm{H}} - N_{\mathrm{SD}}$ is increased by one, N_{FD} is decreased by one, and either N_{DR} or N_{WD} is increased by one). The second one is that $N_{\mathrm{WD}}(i)$ may have high values with quite a high probability, since it is enough a controller failure to take the system to a state i with $N_{\mathrm{WD}}(i)$ equal to the the number of disks previously unfailed in the string associated with the controller. Thus, in order to have tight bounds we should take for K at least the value $G+1$, which is close to $N = G+C_{\mathrm{H}}+D_{\mathrm{H}}+1$, and thus gives enormous state spaces. The first problem can be solved by taking $D_{\mathrm{H}} - N_{\mathrm{SD}}$ out of the definition of the index k defining the partition when thera are failed disks. In order to solve the second problem, it is necessary to differentiate the disks under reconstruction or waiting for reconstruction as a result of a disk failure, or those which are in those states as a result of a controller failure. Then, let $N_{\mathrm{DRF}}(i)$ and $N_{\mathrm{WDF}}(i)$ be, respectively, the number of disks under reconstruction or waiting for reconstruction as a result of the failure of the disk, and let $N_{\mathrm{DRC}}(i)$ and $N_{\mathrm{WDC}}(i)$ be, respectively, the number of disks under reconstruction or waiting for reconstruction as a result of a controller failure, all in state i. Then, while we have to generate states i with high values of $N_{\mathrm{DRC}}(i)$ and $N_{\mathrm{WDC}}(i)$ we only have to generate states with moderate values of $N_{\mathrm{DRF}}(i)$ and $N_{\mathrm{WDF}}(i)$. Another comment is that disks under reconstruction due to a controller failure are the result of a controller failure and disappear with fast rates and we can think of the reconstruction of those disks as part of the "repair" engaged when the controller fails. Thus, we should count as one pending failure the fact that there are disks under reconstruction due to a controller failure. Doing so requires, in order not to have fast rates to the right, to leave out of the counter defining the partition the number of lacking hot spare controllers. Using those ideas we define the following partition:

$$C_0 = \{r\}\,,$$

and, denoting by I_c the indicator function returning the value 1 when condition c is true and 0 otherwise, for $C_{\rm H}+D_{\rm H} > 0$,

$$\begin{aligned} C_1 = \{f\} & \\ \cup\{i \in U \,:\, & N_{\rm FC}(i) + N_{\rm FD}(i) + N_{\rm DRF}(i) + N_{\rm WDF}(i) \\ & + I_{N_{\rm FC}(i)=0 \,\wedge\, N_{\rm DRC}(i)>0} = 0 \\ & \wedge (C_{\rm H} - N_{\rm SC}(i)) + (D_{\rm H} - N_{\rm SD}(i)) = 1\}\,, \end{aligned}$$

$$\begin{aligned} C_k = & \\ \{i \in U \,:\, & N_{\rm FC}(i) + N_{\rm FD}(i) + N_{\rm DRF}(i) + N_{\rm WDF}(i) \\ & + I_{N_{\rm FC}(i)=0 \,\wedge\, N_{\rm DRC}(i)>0} = 0 \\ & \wedge (C_{\rm H} - N_{\rm SC}(i)) + (D_{\rm H} - N_{\rm SD}(i)) = k\}\,, \end{aligned}$$

$$2 \leq k \leq C_{\rm H} + D_{\rm H}\,,$$

$$\begin{aligned} C_{C_{\rm H}+D_{\rm H}+k} = & \\ \{i \in U \,:\, & N_{\rm FC}(i) + N_{\rm FD}(i) + N_{\rm DRF}(i) + N_{\rm WDF}(i) \\ & + I_{N_{\rm FC}(i)=0 \,\wedge\, N_{\rm DRC}(i)>0} = k\}\,, \end{aligned}$$

$$1 \leq k \leq G+1\,,$$

and, for $C_{\rm H} + D_{\rm H} = 0$,

$$\begin{aligned} C_1 = \{f\} & \\ \cup\{i \in U \,:\, & N_{\rm FC}(i) + N_{\rm FD}(i) + N_{\rm DRF}(i) + N_{\rm WDF}(i) \\ & + I_{N_{\rm FC}(i)=0 \,\wedge\, N_{\rm DRC}(i)>0} = 1\}\,, \end{aligned}$$

$$\begin{aligned} C_k = & \\ \{i \in U \,:\, & N_{\rm FC}(i) + N_{\rm FD}(i) + N_{\rm DRF}(i) + N_{\rm WDF}(i) \\ & + I_{N_{\rm FC}(i)=0 \,\wedge\, N_{\rm DRC}(i)>0} = k\}\,, \end{aligned}$$

$$2 \leq k \leq G+1\,.$$

Using that partition, tight bounds can be obtained by taking $K = C_{\rm H}+D_{\rm H}+M$, where M has a moderate value, say from 3 to 5. In order to generate Y, it is necessary to keep track of $N_{\rm DRF}(i)$, $N_{\rm DRC}(i)$ and $N_{\rm WDF}(i)$, which requires to introduce for each string of the RAID counters of the number of disks under reconstruction due to failure of the disk, under reconstruction due to failure of the controller, waiting for reconstruction due to failure of the disk, and waiting for reconstruction due to failure of the controller. That refinement of the state description enlarges considerably the number of states of X, but is well paid-off by the decrease in size of the state space of Y.

It remains to obtain values for the transition rates $f^+_{k,l}$ and g^-_k of Y. We start with $f^+_{k,l}$. Given the definition of the partition, only failure transitions which do not cause system failure go to the right. Thus, in order to compute $f^+_{k,l}$ we only have to consider failure transition rates, which, as desired, are small. To compute $f^+_{k,l}$, $1 \leq k \leq C_{\rm H} + D_{\rm H}$ we note that states in C_k, $1 \leq k \leq C_{\rm H} + D_{\rm H}$ have transitions to the right only to $C_{C_{\rm H}+D_{\rm H}+1}$. Those transitions may originate from controller failures whose rate is $N\lambda_{\rm C}$ and by disk failures not causing system failure, whose rate is $GN\lambda_{\rm D}$. Thus, for $1 \leq k \leq C_{\rm H} + D_{\rm H}$, we can take

$$f^+_{k,C_{\rm H}+D_{\rm H}+1} = N\lambda_{\rm C} + GN\lambda_{\rm D}\,.$$

States in C_k, $C_{\rm H} + D_{\rm H} + 1 \leq k \leq C_{\rm H} + D_{\rm H} + G$ have transitions to the right only to C_{k+1}. Those transitions may originate from controller failures whose rate is upper bounded by $N\lambda_{\rm C}$ and by disk failures not causing system failure. The rate of those disk failures is upper bounded by $\max_{0\leq k\leq G} k\lambda_{\rm S} + (G-k)N\lambda_{\rm D}$ if no controller is failed and by $\max_{0\leq k\leq G} k\lambda_{\rm D}$ if one controller is failed. Thus, for $C_{\rm H} + D_{\rm H} + 1 \leq k \leq C_{\rm H} + D_{\rm H} + G$, we can take

$$\begin{aligned} f^+_{k,k+1} &= N\lambda_C + \max\Big\{\max_{0\leq k\leq G} k\lambda_{\rm S} + (G-k)N\lambda_{\rm D}, \\ &\qquad\qquad \max_{0\leq k\leq G} k\lambda_{\rm D}\Big\} \\ &= N\lambda_{\rm C} + \max\{G\lambda_{\rm S}, GN\lambda_{\rm D}, G\lambda_{\rm D}\} \\ &= N\lambda_{\rm C} + \max\{G\lambda_{\rm S}, GN\lambda_{\rm D}\} = f^+\,. \end{aligned}$$

We next proceed with the computation of g^-_k. We start discussing the case $C_{\rm H} + D_{\rm H} > 0$ and $1 \leq k \leq C_{\rm H} + D_{\rm H}$. Within that case we will start discussing the case $k = 1$. State f has transition to the left with rate $\mu_{\rm G}$. The other states of C_1 have one lacking hot spare controller or disk and thus have a transition to the left with rate $\mu_{\rm SR}$. Then, for the case $C_{\rm H} + D_{\rm H} > 0$ we can take

$$g^-_1 = \min\{\mu_{\rm G}, \mu_{\rm SR}\}\,.$$

For $2 \leq k \leq C_{\rm H} + D_{\rm H}$, the states $i \in C_k$ have k hot spare replacements underway and have transition rate to the left of value $k\mu_{\rm SR}$. Thus, we can take

$$g^-_k = k\mu_{\rm SR}\,.$$

It remains to discuss the case $C_{\rm H} + D_{\rm H} + 1 \leq k \leq C_{\rm H} + D_{\rm H} + G + 1$. We will not distinguish in that case the subcases $C_{\rm H} + D_{\rm H} > 0$ and $C_{\rm H} + D_{\rm H} = 0$, except that, for $C_{\rm H} + D_{\rm H} = 0$ and $k = 1$, C_1 includes state f. For $k = C_{\rm H} + D_{\rm H} + G + 1$ all states in C_k have one failed controller and $C_{\rm H}$ lacking controller hot spares. Since there are not available controller hot spares, the failed controller is repaired by the maintenance team at rate $\mu_{\rm SR}$. In addition, lacking hot spares are brought at rate $(C_{\rm H} + D_{\rm H})\mu_{\rm SR}$. Both types of transitions go to the left. Therefore, we can take

$$g^-_{C_{\rm H}+D_{\rm H}+G+1} = (C_{\rm H} + D_{\rm H} + 1)\mu_{\rm SR}\,.$$

It remains to consider the case $C_{\rm H} + D_{\rm H} + 1 \leq k \leq C_{\rm H} + D_{\rm H} + G$. For those values of k, up states i in C_k with $N_{\rm FC}(i) = 0$, $N_{\rm FD}(i) > 0$, $N_{\rm DRF}(i) = 0$, $N_{\rm WDF} = 0$, and $N_{\rm DRC}(i) = 0$ have no fast transition to the left, and taking a lower bound for the minimum of the transition rates to the left would give a poor g_k^-. For those values of k we will take a g_k^- lower bounding $\min_{i \in C_k} q_i/h_i$, where q_i is the probability that C_k will be left by the left starting in state i and h_i is the mean holding time in C_k starting in state i. In order to compute lower bounds for q_i and upper bounds for h_i we will need the following two lemmas.

Lemma 1. *Let Z be a transient CTMC with state space $S \cup \{a_{\rm L}, a_{\rm R}\}$, where $a_{\rm L}$ and $a_{\rm R}$ are absorbing states and all states in S are transient, and $P[Z(0) \in S] = 1$. Let $P_{\rm L} = \lim_{t\to\infty} P[X(t) = a_{\rm L}]$ and $P_{\rm R} = \lim_{t\to\infty} P[X(t) = a_{\rm R}]$ ($P_{\rm L}$ and $P_{\rm R}$ are, respectively, the probabilities that Z will be absorbed in $a_{\rm L}$ and $a_{\rm R}$). Denote by g_i the transition rate from $i \in S$ to $a_{\rm L}$ and denote by f_i the transition rate from $i \in S$ to $a_{\rm R}$. Assume $g_i > 0$, $i \in S$, let $g^- \leq \min_{i\in S} g_i$, $g^- > 0$, and let $f^+ \geq \max_{i \in S} f_i$. Then, $P_{\rm L} \geq g^-/(g^- + f^+)$.*

Proof Let τ_i, $i \in S$ be the mean time to absorption in i of Z. We have

$$P_{\rm L} = \sum_{i \in S} \tau_i g_i \,, \tag{1}$$

$$P_{\rm R} = \sum_{i \in S} \tau_i f_i \,. \tag{2}$$

From (2)

$$P_{\rm R} \leq \left(\sum_{i \in S} \tau_i\right) f^+ \,,$$

$$\sum_{i \in S} \tau_i \geq \frac{P_{\rm R}}{f^+} \,. \tag{3}$$

From (1) and (3)

$$P_{\rm L} \geq \left(\sum_{i \in S} \tau_i\right) g^- \geq P_{\rm R} \frac{g^-}{f^+} \,.$$

Using $P_{\rm L} + P_{\rm R} = 1$

$$P_{\rm L} \geq (1 - P_{\rm L}) \frac{g^-}{f^+} \,,$$

$$P_{\rm L} \left(1 + \frac{g^-}{f^+}\right) \geq \frac{g^-}{f^+} \,,$$

$$P_{\rm L} \geq \frac{\dfrac{g^-}{f^+}}{1 + \dfrac{g^-}{f^+}} = \frac{g^-}{g^- + f^+} \,. \quad \Box$$

Lemma 2. *Let Z be a transient CTMC with state space $S \cup \{a\}$, where a is an absorbing state and all states in S are transient, and $P[Z(0) \in S] = 1$. Let g_i, $i \in S$ denote the transition rate from i to a, assume $g_i > 0$, $i \in S$, and let $g^- \leq \min_{i \in S} g_i$, $g^- > 0$. Let $h = \sum_{i \in S} \int_0^\infty P[Z(t) = i]\, dt$ be the mean time to absorption of Z. Then, $h \leq 1/g^-$.*

Proof Let $p_i(t) = P[Z(t) = i]$, $i \in S$. Let $p(t) = \sum_{i \in S} p_i(t)$. We have

$$\frac{dp}{dt} = -\sum_{i \in S} p_i(t) g_i \,.$$

Define

$$g(t) = \frac{\displaystyle\sum_{i \in S} p_i(t) g_i}{p(t)} \,.$$

We have

$$\frac{dp}{dt} = -g(t)p(t) \,. \tag{4}$$

The solution of (4) ($p(0) = 1$) is

$$p(t) = e^{-\int_0^t g(\tau)\, d\tau} \,.$$

But $g(t) \geq g^-$ and

$$p(t) \geq e^{-g^- t}. \tag{5}$$

Integrating (5)

$$h = \int_o^\infty p(t)\, dt \geq \int_0^\infty e^{-g^- t}\, dt = \frac{1}{g^-} \,. \quad \Box$$

To derive g_k^- for $C_{\rm H} + D_{\rm H} + 1 \leq k \leq C_{\rm H} + D_{\rm H} + G$ we assume $k > 1$ for the case $C_{\rm H} + D_{\rm H} = 0$ and define the following partition of C_k (note that all states in C_k belong to U, i.e. are up states)

$$C_k^1 = \{i \in C_k \;:\; N_{\rm FC}(i) = 0 \;\wedge\; N_{\rm DRC}(i) > 0\} \,,$$

$$C_k^2 = \{i \in C_k \;:\; N_{\rm FC}(i) = 1\} \,,$$

$$C_k^3 = \{i \in C_k \;:\; N_{\rm FC}(i) = 0 \;\wedge\; N_{\rm DRC}(i) = 0 \;\wedge\; N_{\rm DRF}(i) > 0\} \,,$$

$$C_k^4 = \{i \in C_k \;:\; N_{\rm FC}(i) = 0 \;\wedge\; N_{\rm DRC}(i) = 0 \;\wedge\; N_{\rm DRF}(i) = 0\} \,.$$

We will derive lower bounds for $\min_{i\in C_k^l} q_i/h_i$ for each subset C_k^l of the partition and g_k^- will be the minimum of all those lower bounds.

Assume $C_k^1 \neq \emptyset$ and consider $i \in C_k^1$. Let $j = N_{\mathrm{DRC}}(i)$. State i may not have fast transition to the left if $j > 1$. Consider the partition of $C_k^1 \cup_{1\leq l\leq G} C_{k,l}^1$, where $C_{k,l}^1$ includes the states of C_k^1 with $N_{\mathrm{DRC}}(i) = l$. States in subset $C_{k,l}^1$, $l > 1$ have only transitions to the left, to the right and to $C_{k,l-1}^1$, the transition rates to the right are upper bounded by f^+ and the transition rates to the left or to $C_{k,l-1}^1$ are lower bounded by $l\mu_{\mathrm{DRC}}$. States in the subset $C_{k,1}^1$ have only transitions to the left and to the right, the transition rates to the right are upper bounded by f^+ and the transition rates to the left are lower bounded by μ_{DRC}. Consider any state $s \in C_{k,l}^1$, let $q_{s,l}$ be the probability that X will make a jump to the left or, if $l > 1$, to $C_{k,l-1}^1$ starting in state s, and let $h_{s,l}$ be the mean holding time in $C_{k,l}^1$ starting in s. Using Lemmas 1–2 we have

$$q_{s,l} \geq \frac{l\mu_{\mathrm{DRC}}}{l\mu_{\mathrm{DRC}} + f^+} \, ,$$

$$h_{s,l} \leq \frac{1}{l\mu_{\mathrm{DRC}}} \, .$$

Obviously, q_i is lower bounded by the product of the lower bounds for $q_{s,j}, q_{s,j-1}, \ldots, q_{s,1}$ and h_i is upper bounded by the sum of the upper bounds for $h_{s,j}, h_{s,j-1}, \ldots, h_{s,1}$. This gives

$$q_i \geq \prod_{l=1}^{j} \frac{l\mu_{\mathrm{DRC}}}{l\mu_{\mathrm{DRC}} + f^+} \, ,$$

$$h_i \leq \sum_{l=1}^{j} \frac{1}{l\mu_{\mathrm{DRC}}} \, .$$

Thus, taking into account that j may vary between 1 and G, assuming $C_k^1 \neq \emptyset$, we have

$$\min_{i\in C_k^1} \frac{q_i}{h_i} \geq \min_{1\leq j\leq G} \frac{\prod_{l=1}^{j} \frac{l\mu_{\mathrm{DRC}}}{l\mu_{\mathrm{DRC}} + f^+}}{\sum_{l=1}^{j} \frac{1}{l\mu_{\mathrm{DRC}}}} = \frac{\prod_{l=1}^{G} \frac{l\mu_{\mathrm{DRC}}}{l\mu_{\mathrm{DRC}} + f^+}}{\sum_{l=1}^{G} \frac{1}{l\mu_{\mathrm{DRC}}}} \, . \tag{6}$$

Assume now $C_k^2 \neq \emptyset$ and consider $i \in C_k^2$. If $N_{\mathrm{SC}}(i) > 0$ the failed controller will be replaced by a hot spare at rate μ_{CRP}; if $N_{\mathrm{SC}}(i) = 0$ the failed controller will be repaired by the maintenance team at rate μ_{SR}. Summarizing, we will have a rate to C_k^1 or to the left from i and all states within C_k^2 reachable from $i \geq \mu_{\min 1} = \min\{\mu_{\mathrm{CRP}}, \mu_{\mathrm{SR}}\}$. Let q_i^2 be the probability that starting in i X will make a jump to C_k^1 or to the left and let h_i^2 be the mean holding time in C_k^2 starting in i. Using Lemmas 1–2 we have

$$q_i^2 \geq \frac{\mu_{\min 1}}{\mu_{\min 1} + f^+} \, ,$$

$$h_i^2 \leq \frac{1}{\mu_{\min 1}} .$$

Assume that there is a jump from i to C_k^1 and let s be the entry state. Let $j = N_{\mathrm{DRC}}(s)$. Let q_s^1 be the probability that there will be an exit to the left starting in s and let h_s^1 be the mean holding time in C_k^1 starting in s. Following the previous developments we have

$$q_s^1 \geq \prod_{l=1}^{j} \frac{l\mu_{\mathrm{DRC}}}{l\mu_{\mathrm{DRC}} + f^+} \, ,$$

$$h_s^1 \leq \sum_{l=1}^{j} \frac{1}{l\mu_{\mathrm{DRC}}} \, .$$

It is clear that q_i is lower bounded by the product of the lower bound for q_i^2 and the lower bound for q_s^1, which achieves its minimum for $j = G$. Also, h_i is upper bounded by the sum of the upper bound for h_i^2 and the upper bound for h_s^1, which achieves its maximum for $j = G$. Then, assuming $C_k^2 \neq \emptyset$, we have

$$\min_{i\in C_k^2} \frac{q_i}{h_i} \geq \frac{\frac{\mu_{\min 1}}{\mu_{\min 1} + f^+} \prod_{l=1}^{G} \frac{l\mu_{\mathrm{DRC}}}{l\mu_{\mathrm{DRC}} + f^+}}{\frac{1}{\mu_{\min 1}} + \sum_{l=1}^{G} \frac{1}{l\mu_{\mathrm{DRC}}}} \, . \tag{7}$$

Assume $C_k^3 \neq \emptyset$ and consider $i \in C_k^3$. Let $j = N_{\mathrm{DRF}}(i)$. All states in C_k reachable within C_k from i have a transition rate to the left $\geq j\mu_{\mathrm{DRC}}$. The transition rates to the right are upper bounded by f^+. Then, using Lemmas 1–2 we have

$$q_i \geq \frac{j\mu_{\mathrm{DRC}}}{j\mu_{\mathrm{DRC}} + f^+} \, ,$$

$$h_i \leq \frac{1}{j\mu_{\mathrm{DRC}}} .$$

Noting that j varies between 1 and G, we obtain, assuming $C_k^3 \neq \emptyset$

$$\begin{aligned} \min_{i\in C_k^3} \frac{q_i}{h_i} &\geq \min_{1\leq j\leq G} \frac{j\mu_{\mathrm{DRC}}}{j\mu_{\mathrm{DRC}} + f^+} \, j\mu_{\mathrm{DRC}} \\ &= \frac{\mu_{\mathrm{DRC}}}{\mu_{\mathrm{DRC}} + f^+} \, \mu_{\mathrm{DRC}} \, . \end{aligned} \tag{8}$$

Finally, assume $C_k^4 \neq \emptyset$ and consider $i \in C_k^4$. State i has no fast transition to the left. We have $N_{\rm FD}(i) > 0$. If there are hot spare disks available, disk repair will occur at a rate lower bounded by $\mu_{\rm DRP}$. If there are hot spare disks available, disk repair will occur at rate lower bounded by $\mu_{\rm SR}$. Thus, disk repair will occur at rate lower bounded by $\mu_{\min 2} = \min\{\mu_{\rm DRP}, \mu_{\rm SR}\}$. Upon repair, a state in C_k^3 will be entered and a disk reconstruction will start. Summarizing, states in C_k^4 reachable within C_k^4 have only transitions to C_k^3, to the left, and to the right. The transition rate to C_k^3 or to the left of those states is lower bounded by $\mu_{\min 2}$. The transition rate to the right is upper bounded by f^+. Then, denoting by q_i^4 the probability that, starting in i, C_k^4 will be exited by C_k^3 or the left, and by h_i^4 the mean holding time in C_k^4 starting in i, using Lemmas 1–2 we have

$$q_i^4 \geq \frac{\mu_{\min 2}}{\mu_{\min 2} + f^+} \, ,$$

$$h_i^4 \leq \frac{1}{\mu_{\min 2}} .$$

Assume that there is a jump to C_k^3 and let s be the entry state in that subset. We have $N_{\rm DRF}(s) = 1$ and the states in C_k^3 reachable within C_k^3 from s have a transition rate to the left $\geq \mu_{\rm DRC}$. The transition rates to the right are upper bounded by f^+. Then, denoting by q_s^3 the probability that given entry by s C_k^3 will be exited by the left and by h_s^3 the mean holding time in C_k^3 given entry by s, using Lemmas 1–2 we have

$$q_s^3 \geq \frac{\mu_{\rm DRC}}{\mu_{\rm DRC} + f^+} \, ,$$

$$h_s^3 \leq \frac{1}{\mu_{\rm DRC}} \, .$$

Now, q_i is lower bounded by the product of the lower bound for q_i^4 and the lower bound for q_s^3 and h_i is upper bounded by the sum of the upper bound for h_i^4 and the upper bound for h_s^3. Then, we have

$$q_i \geq \frac{\mu_{\min 2}}{\mu_{\min 2} + f^+} \frac{\mu_{\rm DRC}}{\mu_{\rm DRC} + f^+} \, ,$$

$$h_i \leq \frac{1}{\mu_{\min 2}} + \frac{1}{\mu_{\rm DRC}} \, .$$

The, assuming $C_k^4 \neq \emptyset$

$$\min_{i \in C_k^4} \frac{q_i}{h_i} \geq \frac{\mu_{\min 2}}{\mu_{\min 2} + f^+} \frac{\mu_{\rm DRC}}{\mu_{\rm DRC} + f^+} \frac{\mu_{\min 2}\mu_{\rm DRC}}{\mu_{\min 2} + \mu_{\rm DRC}} \, . \tag{9}$$

Finally, using (6), (7), (8), and (9), we can take, for $C_{\rm H} + D_{\rm H} + 1 \leq k \leq G + C_{\rm H} + D_{\rm H}$,

$$\begin{aligned} g_k^- = \; & \min \left\{ \frac{\prod_{l=1}^{G} \frac{l\mu_{\rm DRC}}{l\mu_{\rm DRC} + f^+}}{\sum_{l=1}^{G} \frac{1}{l\mu_{\rm DRC}}}, \right. \\ & \frac{\frac{\mu_{\min 1}}{\mu_{\min 1} + f^+} \prod_{l=1}^{G} \frac{l\mu_{\rm DRC}}{l\mu_{\rm DRC} + f^+}}{\frac{1}{\mu_{\min 1}} + \sum_{l=1}^{G} \frac{1}{l\mu_{\rm DRC}}}, \\ & \frac{\mu_{\rm DRC}}{\mu_{\rm DRC} + f^+} \mu_{\rm DRC}, \\ & \left. \frac{\mu_{\min 2}}{\mu_{\min 2} + f^+} \frac{\mu_{\rm DRC}}{\mu_{\rm DRC} + f^+} \frac{\mu_{\min 2}\mu_{\rm DRC}}{\mu_{\min 2} + \mu_{\rm DRC}} \right\} \end{aligned}$$

$$= \min \left\{ \frac{\frac{\mu_{\min 1}}{\mu_{\min 1} + f^+} \prod_{l=1}^{G} \frac{l\mu_{\rm DRC}}{l\mu_{\rm DRC} + f^+}}{\frac{1}{\mu_{\min 1}} + \sum_{l=1}^{G} \frac{1}{l\mu_{\rm DRC}}} \, , \; \frac{\mu_{\min 2}}{\mu_{\min 2} + f^+} \frac{\mu_{\rm DRC}}{\mu_{\rm DRC} + f^+} \frac{\mu_{\min 2}\mu_{\rm DRC}}{\mu_{\min 2} + \mu_{\rm DRC}} \right\} .$$

It remains to discuss the case $C_{\rm H} + D_{\rm H} = 0$, $k = 1$. In that case, C_k includes state f. That state has a transition to the left with rate $\mu_{\rm G}$ and no other transition. Therefore, we have $q_f = 1$ and $h_f = 1/\mu_{\rm G}$, which yields $q_f/h_f = \mu_{\rm G}$. Therefore, in that case it is enough to take for g_1^- the minimum of the above expression and $\mu_{\rm G}$.

We have completed the description of the bounding CTMC Y. Specification of Y using the modeling language of METFAC-2.1 depends on N and is cumbersome. Thus, for $N = 10$ the model specification file had 2,223 lines. In addition, the CTMCs Y are large. Table 2 gives the number of states and transitions of Y for some set of the parameters G, N, $C_{\rm H}$, $D_{\rm H}$ and M on which the structure of Y depends used in the validation experiments.

3.2 Bounding model for the unreliability

Obtaining bounds for the unreliability is conceptually much simpler than for the steady-state unavailability. The bounding CTMC Y has state space $G \cup \{f, a\}$, where f is the absorbing system failed state, a is another absorbing state and $G = \cup_{0 \leq k \leq M} C_k$, where $C_k = \{i \in U : N_{\rm FC}(i) + N_{\rm FD}(i) + N_{\rm DRF}(i) + N_{\rm WDF}(i) + I_{N_{\rm FC}(i)=0 \wedge N_{\rm DRC}(i)>0} = k\}$. Transition rates within G and from G to f in Y are as in X; in addition, Y has transition rates from states $i \in G$ to a with rates $\lambda_{i,U-G}$. A lower bound for the unreliability at

Table 2. Size of bounding model Y for the steady-state unavailability for several sets of values of G, N, C_{H}, D_{H} and M.

G	N	C_{H}	D_{H}	M	states	transitions
5	5	1	2	4	25,816	188,795
5	10	1	2	4	292,066	2,336,490
10	5	1	2	4	68,721	522,655
10	10	1	2	4	823,371	6,782,950
15	5	1	2	4	111,626	856,515
15	10	1	2	4	1,354,676	11,229,410
20	5	1	2	4	154,531	1,119,375
20	19	1	2	4	1,885,981	15,675,870

Table 5. Analysis of the accuracy of the approximate model for the unreliability at $t = 8{,}760\,\mathrm{h}$.

G	N	approximate	exact	rel. error
5	5	0.016062752	0.016062750	1×10^{-7}
5	10	0.038646150	0.038646143	2×10^{-7}
10	5	0.030989562	0.030989550	4×10^{-7}
10	10	0.072143536	0.072143484	7×10^{-7}
15	5	0.045511165	0.045511134	7×10^{-7}
15	10	0.010383939	0.010383926	1×10^{-6}
20	5	0.059700616	0.059700557	1×10^{-6}
20	10	0.013409427	0.013409403	2×10^{-6}

time t is $P[Y(t) = f]$; an upper bound is $P[Y(t) \in \{f, a\}]$. The size of Y is slightly smaller than the size of the Y used for the steady-state unavailability with the same value of M.

3.3 Analysis of the accuracy of the approximate model

In order to asses the accuracy of the approximate models we used the bounding techniques with the minimum M required to have an accuracy of 8 digits in the solution of the exact models. Those solutions were then compared with the solutions given by the approximate models and the relative errors were computed. For the analysis we took $C_{\mathrm{H}} = 1$, $D_{\mathrm{H}} = 2$, $\lambda_{\mathrm{D}} = 10^{-5}$, $\lambda_{\mathrm{S}} = 2 \times 10^{-5}$, $\lambda_{\mathrm{C}} = 5 \times 10^{-5}$, $\mu_{\mathrm{DRC}} = 1, \mu_{\mathrm{DRP}} = 4, \mu_{\mathrm{CRP}} = 4, \mu_{\mathrm{SR}} = 0.25, \mu_{\mathrm{G}} = 0.25$, $P_{\mathrm{R}} = 0.999$, and varied the parameters G and N. For G we considered the values 5, 10, 15, and 20; for N we considered the values 5 and 10. All rates have been given in h^{-1}. The models were solved on a workstation with a Sun-Blade-1000 processor and 4 GB of memory.

Table 3 gives the results obtained for the steady-state unavailability. We give the value obtained with the approximate model, the exact value (up to 8 significant digits) and the relative error. Tables 4 and 5 gives the results obtained for the unreliability for, respectively, $t = 1$ h and $t = 8{,}760$ h (1 year). The relative errors increase with both G and N and are significantly smaller for the unreliability at $t = 1$ h. The accuracy of the approximate models is better than 5 significant digits in all cases examined. For larger values of G and N we can expect an increase in the relative errors but the accuracy of the approximate models would still remain excellent for all reallistic values of G and N.

4. Conclusions

We have developed approximate dependability models for a level-5 RAID architecture with orthogonal organization. The approximate models are parametric on both the number of parity groups and the number of strings of the architecture, have a succinct specification, yield CTMCs with small state spaces, and their accuracy is excellent. Validation of the degree of accuracy exhibited by the approximate models would have been impossible using simulation and has been carried out using bounding techniques recently developed. Although such a techniques allow to compute rigorous bounds for the dependability measures, they require much more cumbersome model specifications than the approximate models and require the solution of CTMCs of much larger state spaces. Thus, the approximate models developed in the paper are an attractive alternative for the evaluation of the dependability of level-5 RAID architectures.

References

[1] J. A. Carrasco, "Markovian Dependability/Performability Modeling of Fault-Tolerant Systems," in *Reliability Engineering Handbook*, Springer-Verlag, H. Pham, ed., to appear.

[2] J. A. Carrasco, "Bounding steady-state availability models with group repair and phase type repair distributions," *Performance Evaluation*, vol. 35, no. 4, May 1999, pp. 193–214.

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

[4] G. A. Gibson, *Redundant Disk Arrays: Reliable, Parallel Secondary Storage*, MIT Press, 1992.

[5] R. H. Katz, G. A. Gibson, and D. A. Patterson, "Disk system architectures for high performance computing," *Proceedings of the IEEE*, vol. 77, no. 12, December 1989, pp. 1842–1858.

[6] E. K. Lee and R. H. Katz, "An analytic performance model of disk arrays," in *Proc. 1993 ACM SIGMETRICS Conference*

Table 3. Analysis of the accuracy of the approximate model for the steady-state unavailability.

G	N	approximate	exact	rel. error
5	5	7.3951238×10^{-6}	7.3951230×10^{-6}	1×10^{-7}
5	10	1.7998847×10^{-5}	1.7998843×10^{-5}	2×10^{-7}
10	5	1.4376174×10^{-5}	1.4376169×10^{-5}	3×10^{-7}
10	10	3.4194378×10^{-5}	3.4194352×10^{-5}	8×10^{-7}
15	5	2.1271556×10^{-5}	2.1271541×10^{-5}	7×10^{-7}
15	10	5.0065851×10^{-5}	5.0065783×10^{-5}	1×10^{-6}
20	5	2.8111183×10^{-5}	2.8111154×10^{-5}	1×10^{-6}
20	10	6.5747700×10^{-5}	6.5747575×10^{-5}	2×10^{-6}

Table 4. Analysis of the accuracy of the approximate model for the unreliability at $t = 1$ h.

G	N	approximate	exact	rel. error
5	5	4.5149870×10^{-7}	4.5149870×10^{-7}	$< 1 \times 10^{-7}$
5	10	1.1181868×10^{-6}	1.1181868×10^{-6}	$< 1 \times 10^{-7}$
10	5	8.7737251×10^{-7}	8.7737251×10^{-7}	$< 1 \times 10^{-7}$
10	10	2.1225892×10^{-6}	2.1225892×10^{-6}	$< 1 \times 10^{-7}$
15	5	1.3023726×10^{-6}	1.3023726×10^{-6}	$< 1 \times 10^{-7}$
15	10	3.1245661×10^{-6}	3.1245661×10^{-6}	$< 1 \times 10^{-7}$
20	5	1.7265787×10^{-6}	1.7265787×10^{-6}	$< 1 \times 10^{-7}$
20	10	4.1244730×10^{-6}	4.1244730×10^{-6}	$< 1 \times 10^{-7}$

on Measurement and Modeling of Computer Systems, 1993, pp. 98–109.

[7] E. K. Lee and R. H. Katz, "The performance of parity placements in disk arrays," *IEEE Trans. on Computers*, vol. 42, no. 6, June 1993, pp. 651–664.

[8] S. Mahévas and G. Rubino, "Bounding asymtotic dependability and performance measures," in *Proc. 2nd IEEE Int. Performance and Dependability Symposium*, September 1996, pp. 176–186.

[9] S. W. Ng, "Crosshatch disk array for improved reliability and performance," in *Proc. 1994 IEEE Int. Symp. on Computer Architecture*, 1994.

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

[11] D. A. Patterson, P. M. Chen, G. A. Gibson, and R. H. Katz, "Introduction to redundant arrays of inexpensive disks (RAID)," in *Proc. IEEE Computer Society International Conference (COMPCON)*, 1989, pp. 112–117.

[12] W. H. Sanders, W. D. Obal, M. A. Qureshi, and K. K. Widjanarko, "The UltraSAN modeling environment," *Performance Evaluation*, vol. 24, no. 1, October 1995.

[13] V. Santoja, M. Alonso, J. Molero, J. J. Serrano, P. Gil, and R. Ors, "Dependability Models of RAID using Stochastic Activity Networks," *Dependable Computing-EDCC-2*, Lecture Notes in Computer Science, 1996.

[14] M. Schulze, G. Gibson, R. H. Katz, and D. A. Patterson, "How reliable is a RAID?,", in *Proc. IEEE Computer Society International Conference (COMPCON)*, 1989, pp. 118–123.

[15] D. Stodolsky, M. Holland, W. V. Courtright II, and G. A. Gibson, "Parity-logging disk arrays," *ACM Transactions on Computer Systems*, vol. 12, no. 3, August 1994, pp. 206–235.