Analysis of the mean time to data loss of nested disk arrays RAID-01 on basis of a specialized mathematical model

P A Rahman

Department of Automated Technological and Informational Systems, Ufa State Petroleum Technological University, 2, October ave., Sterlitamak, 453118, Russia

E-mail: pavelar@yandex.ru

Abstract. In this scientific paper a specialized Markov chain for a reliability model of the fault-tolerant system with two types of failures, single batch repair and full restore after reaching the failed state is discussed. The generalized formulas for calculation of the system’s stationary availability factor, the mean time to failure and the mean time to repair are also introduced. Application of the special type of the Markov chain in the reliability model of the nested disk array ‘RAID-01’ and calculation of mean time to data loss by using the generalized formulas are also discussed. The calculation examples of mean time to data loss of the ‘RAID-01’ array are also given.

1. Introduction

The modern world deals with complex technical devices and systems, which are used in the industrial manufacturing processes as well as in everyday life. In addition to the primary parameters of technical systems, like performance, capacity, power etc, the reliability parameters are also quite important, because they define efficiency, stability and security of technical systems.

The academic books, dedicated to the reliability theory [1-4], in most cases deal with the well-known simplified reliability model of the repairable systems, based on Markov Birth-Death Chain [5, 6], which allows us to obtain calculation formulas for such complex and integral reliability parameters as a stationary availability factor and mean time to failure. In more sophisticated cases with several types of failures and special kinds of repairs, the traditional Markov Birth-Death Chain cannot be used in reliability models. For these systems, development of a specialized Markov Chain is required. In particular, the reliability model of the nested disk array ‘RAID-01’ (RAID – redundant array of inexpensive disks) [7, 8] cannot be adequately described by Markov Birth-Death Chain and requires development of a specialized Markov chain.

Within the research work in the field of reliability models of data processing, transmission and storage systems [9, 10] the author developed a special type of Markov Chain for the reliability model of the fault-tolerant system with $n$ elements, two types of failures, single batch repair and full restore after reaching the failed state, and derived the generalized formulas for calculation of the stationary availability factor and the mean time to failure.

Finally, the author applied the specialized Markov Chain to the reliability model of the nested disk array ‘RAID-01’ and obtained a special formula for calculation of the mean time to data loss.
2. Specialized Markov chain for the reliability model of the system with two types of failures, single batch repair and full restore after reaching the failed state

Let us introduce a system with a set of operable states \( j = 0 \ldots n \) and one failed state \( F \). State 0 is an initial fully operable state and the states \( j = 1 \ldots n \) are degraded operable states. In degraded states, the system is operable, however for larger \( j \), system’s functioning performance may be lower, and the system is closer to the failed state.

Also, let us introduce the following scheme of transitions between the different states of system:

- From operable state \( j = 0 \ldots n \), the system can pass to the next more degraded operable state \( j + 1 \), if \( j < n \), or to failed state \( F \), if \( j = n \), with given failure rate \( \lambda_j \).
- From the operable state \( j = 1 \ldots n \), the system can pass back directly to initial operable state 0 with given repair rate \( \mu_j \) by using a single batch repair procedure for one or more failed redundant parts of the system.
- In each operable state \( j = 0 \ldots n \), some kind of critical failure may occur in the system, which transfers it from operable state directly to failed state \( F \) with given rate \( \sigma_j \).
- In the failed state a full restore procedure for the system is used, and it returns the system to initial operable state 0 with given rate \( \gamma \).

Now, we may introduce a Markov chain (figure 1), which represents graphically the reliability model, described above:

![Figure 1. Special type of Markov chain for the reliability model of the system with two types of failures, single batch repair and full restore after reaching the failed-state.](image_url)

Accordingly, the system of Kolmogorov-Chapman equations for the stationary case is as follows:

\[
\begin{align*}
    P_0 + P_1 + \ldots + P_n + P_F &= 1; \\
    - (\lambda_0 + \sigma_0)P_0 + \mu_1P_1 + \ldots + \mu_nP_n + \gamma P_F &= 0; \\
    \lambda_0P_0 - (\mu_1 + \lambda_1 + \sigma_1)P_1 &= 0; \\
    \vdots \\
    \lambda_{n-1}P_{n-1} - (\mu_n + \lambda_n + \sigma_n)P_n &= 0; \\
    \sigma_0P_0 + \ldots + \sigma_{n-1}P_{n-1} + (\lambda_n + \sigma_n)P_n - \gamma P_F &= 0.
\end{align*}
\]

(1)

Availability factor of system is equal to the sum of probabilities of all operable states:

\[
K_S = \sum_{j=0}^{n} P_j.
\]
Next, we can easily calculate the mean time to repair by taking into consideration the fact that, the system from failed state can transit only to initial operable state with rate $\gamma$:

$$T_R = 1/\gamma.$$  

Finally, mean time to failure can be derived by using the well-known in reliability theory identity for repairable systems $K_S = T_F/(T_F + T_R)$:

$$T_F = K_S/((1 - K_S)).$$

So, to obtain the stationary availability factor and the mean time to failure, we need to solve the equations system and get all stationary probabilities.

Within the scientific research work, the author obtained the general analytic formulas (for any given $n > 0$) for calculation of the stationary availability factor and the mean time to failure:

$$M = \sum_{q=0}^{n} \left( \frac{1}{\lambda_q} \prod_{j=1}^{n-q} \left( 1 + \frac{\mu_{q+j} + \sigma_{q+j}}{\lambda_{q+j}} \right) \right);$$

$$D = 1 + \sum_{q=0}^{n} \frac{\sigma_q}{\lambda_q} \prod_{j=1}^{n-q} \left( 1 + \frac{\mu_{q+j} + \sigma_{q+j}}{\lambda_{q+j}} \right);$$

$$K_S = \frac{\gamma M}{\gamma M + D}; \quad T_F = \frac{M}{D}; \quad T_R = \frac{1}{\gamma}.$$  

We should mention that in our later discussion we will also need a special case with $\lambda_n = 0$, in case of which a division by zero problem is occurred in calculation of coefficients $M$ and $D$. However, algebraic conversions show that, in fraction $M / D$, parameter $\lambda_n = 0$ can be eliminated. Thus, for case $\lambda_n = 0$, the author derived the specially converted formulas to avoid the division by zero in calculation of the coefficients $M$ and $D$:

$$M = 1 + (\mu_n + \sigma_n) \sum_{q=0}^{n-1} \frac{1}{\lambda_q} \prod_{j=1}^{n-q-1} \left( 1 + \frac{\mu_{q+j} + \sigma_{q+j}}{\lambda_{q+j}} \right);$$

$$D = (\mu_n + \sigma_n) \sum_{q=0}^{n-1} \frac{\sigma_q}{\lambda_q} \prod_{j=1}^{n-q-1} \left( 1 + \frac{\mu_{q+j} + \sigma_{q+j}}{\lambda_{q+j}} \right);$$

$$K_S = \frac{\gamma M}{\gamma M + D}; \quad T_F = \frac{M}{D}; \quad T_R = \frac{1}{\gamma}.$$  

3. Specialized Markov chain in the reliability model of nested disk array ‘RAID-01’

The nested disk array ‘RAID-01’ is a data storage system, which combines high performance in the inner level due to data striping in ‘RAID-0’ technology and high reliability in the outer level due to the data mirroring in ‘RAID-1’ technology. In the inner level, there are two independent $n$-disk ‘RAID-0’ arrays. Data in each $n$-disk ‘RAID-0’ array is distributed between blocks, located on the different disks. In the outer level, the two ‘RAID-0’ arrays are united to an array of arrays by using the ‘RAID-1’ technology (figure 2). Data in the ‘RAID-1’ array is synchronized between two ‘RAID-0’ arrays.

The ‘RAID-01’ array in total contains $2n$ disks ($2n \geq 4$). User data capacity is equal to 50% of total disk space because of data mirroring between two ‘RAID-1’ arrays.

In the best case, the ‘RAID-01’ array is operable even if $n$ disks fail, if all of them belong to the one of the ‘RAID-0’ arrays. In the worst case, even failure of two disks can cause the whole ‘RAID-01’ failure and data loss, if one of the disks belongs to first ‘RAID-0’ array, and the other one belongs to second ‘RAID-0’ array.
Figure 2. The structure of the nested disk array ‘RAID-01’.

Let us now observe the offered reliability model of data storage system, based on the nested disk array ‘RAID-01’ with \(2n\) identical disks (the same capacity, manufacturer and model of all disks). The disk failure rate is \(\lambda\). Disks can fail independently.

The initial state of the data storage system is 0 (all disks are operable). From this state, in case of failure one of the \(2n\) disks, system passes to the state 1 (one of non-fatal failure occurred). In this case one of the ‘RAID-0’ arrays, in which failed disk is located, fails, however, due to the second operable ‘RAID-0’ array, the whole data storage system still operable.

From the state 1, in case of failure of one of the \(2n-1\) disks, the system passes either to the state 2 (two of non-fatal failure occurred), if failed disk is the one of the \(n-1\) disks, which belong to the previously failed ‘RAID-0’ array, or to failed state F if failed disk belongs to the another ‘RAID-0’.

From state 2, in case of failure of one of the \(2n-2\) disks, the system passes either to the state 3 (three of non-fatal failures occurred), if failed disk is the one of the \(n-2\) disks, which belong to the previously failed ‘RAID-0’ array, or to failed state F, if failed disk belongs to another ‘RAID-0’.

Accordingly, in state \(n\) (\(n\) of non-fatal failures occurred) all \(n\) disks in of the ‘RAID-0’ arrays are failed, thus failure of any remaining \(n\) disks, all of which in this case belong to the another ‘RAID-0’ array, will transfer the system to failed state F, because in this case both of the ‘RAID-0’ arrays fail.

Next, let us assume, that in any of state \(j = 1...n\), when all failed disks belong to one of the ‘RAID-0’ arrays, the ‘RAID-1’ technology launches a single batch rebuild process between the two ‘RAID-0’ arrays after replacement of all failed disks in the failed ‘RAID-0’ array. After completion of the rebuild process system passes to initial state 0. The ‘RAID-1’ array rebuild rate is \(\mu\). If a disk failure occurs during the batch rebuild process and failed disk belongs to the previously failed ‘RAID-0’ array, then the rebuild process will be restarted after replacement of the failed disk. Rebuild rate in all states \(j = 1...n\) has a same value and does not depend on number of the replaced disks, because rebuild process involves two whole ‘RAID-0’ arrays. Moreover, for simplification of the reliability model, we will consider the disk replacement time as negligible (assuming that the unlimited amount of the hot-spare disks is available).

We must mention, that during the rebuild process, data from the \(n\) disks of the operable ‘RAID-0’ array are copied to the \(n\) disks of the failed ‘RAID-0’ array, and we will take into consideration an additional disk read error rate \(\varepsilon\), which should be added to the disk failure rate. A read error as well as failure of any disk on the operable ‘RAID-0’ array during the rebuild process will transfer the system directly to failed state F.

Next, let us take into consideration critical array controller errors with rate \(\sigma\), which transfer the system from any operable state \(j = 0...n\) directly to failed state F.
Finally, let us assume that the data storage system does not include any additional hardware and software to provide total backup of the user data, thus, in case of reaching the failed state all data will be lost. So, we will consider the rate of system full recovery rate, which is zero, $\gamma = 0$.

Now, we may introduce next Markov chain (figure 3), which represents graphically the reliability model, described above:

![Figure 3. Special type of reliability model for the nested disk array ‘RAID-01’.](image)

This model is a particular case of the specialized Markov chain, discussed above, with the following substitutions of the source reliability parameters:

$$
\begin{align*}
\lambda_0 &= 2n\lambda; \\
\lambda_j &= (n-j)\lambda; \\
\sigma_0 &= \sigma; \\
\sigma_j &= \sigma + n(\lambda + \varepsilon); \\
\gamma &= 0; \\
\mu_j &= \mu;
\end{align*}
$$

As the system full recovery rate is zero, $\gamma = 0$, mean time to recovery is infinite and stationary availability factor is zero for the discussed ‘RAID-01’ array. However, the mean time to data loss for ‘RAID-01’ is equal to the mean time to failure in discussed above special type of Markov chain. So, after substitution of the source reliability parameters to the discussed above calculation formula (3), with taking into consideration $\lambda_n = 0$, we obtain the following formula for calculation of the mean time to data loss of the ‘RAID-01’ array:

$$
\begin{align*}
M &= 1 + (\mu + \sigma + n(\lambda + \varepsilon)) \times \\
&\quad \times \sum_{q=0}^{n-1} \frac{1}{((2 - \min(1, q))n - q)\lambda} \times \prod_{j=1}^{n-1-q} \left(1 + \frac{\mu + \sigma + n(\lambda + \varepsilon)}{(n-q-j)\lambda}\right); \\
D &= \sigma + n(\lambda + \varepsilon) + (\mu + \sigma + n(\lambda + \varepsilon)) \times \\
&\quad \times \sum_{q=0}^{n-1} \frac{\sigma + \min(1, q)n(\lambda + \varepsilon)}{((2 - \min(1, q))n - q)\lambda} \times \prod_{j=1}^{n-1-q} \left(1 + \frac{\mu + \sigma + n(\lambda + \varepsilon)}{(n-q-j)\lambda}\right); \\
T_{DL} &= M / D.
\end{align*}
$$

4. The calculation examples for the mean time to data loss of the ‘RAID-01’ array

A ‘RAID-01’ array with total number of disks $2n$ is given. The disk failure rate is $\lambda = 1/120000$ hour$^{-1}$, the additional disk read error rate during rebuild process is $\varepsilon = 1/112$ hour$^{-1}$, the ‘RAID-1’ array rebuild rate is $\mu = 1/9$ hour$^{-1}$ and the array controller critical error rate is $\sigma = 1/1200000$ hour$^{-1}$.

Let us calculate mean time to data loss of the ‘RAID-01’ array for the total number of disks $2n = 2, 4, 6, 8, 10$ and $12$. 


Calculations by the formula (5) give us the following values, presented in table 1:

**Table 1. Mean times to data loss of the ‘RAID-01’ array for different total number of disks.**

| n  | T<sub>DL</sub> (hours) |
|----|------------------------|
| 4  | 184042                 |
| 6  | 95106                  |
| 8  | 58827                  |
| 10 | 40574                  |
| 12 | 30060                  |
| 14 | 23413                  |
| 16 | 18918                  |

Calculation results show that the mean time to data loss of the ‘RAID-01’ array is rapidly decreased with the growth of the total number of disks.

5. **Conclusion**

Within the scope of this scientific paper, a special type of Markov chain for the reliability model of the fault-tolerant system with two types of failures, single batch repair and full restore after reaching the failed state is discussed.

The generalized formulas for calculation of the system’s stationary availability factor, the mean time of failure and the mean time to repair are also introduced. Finally, application of a special type of Markov chain in the reliability model of the nested disk array ‘RAID-01’ and calculation of the mean time to data loss by using the generalized formulas are also presented.

Obtained scientific results were used by the author in designing of the fault-tolerant data storage systems based on the ‘RAID-01’ arrays for Moscow Power Engineering Institute, Nuclear Power Plant ‘Balakovo’ and several other enterprises.

**Acknowledgements**

Author thanks S. N. Khorkov, chief network administrator of Moscow Power Engineering Institute (MPEI), for technical support and professor I. I. Ladygin, MPEI, for scientific support.

**References**

[1] Ostreykovsky V A 2003 *Reliability Theory* (Moscow: Vyshaya Shkola)
[2] Ushakov I A 2008 *Reliability Theory Course* (Moscow: Drofa)
[3] Rausand M and Holyand A 2009 *System Reliability Theory* (John Wiley & Sons)
[4] Abd-El-Barr M 2007 *Design and Analysis of Reliable and Fault-Tolerant Computer Systems* (London: Imperial College Press)
[5] Kijma M 1997 *Markov Processes for Stochastic Modeling* (London: Chapman & Hall)
[6] Anderson W J 1991 *Continuous-Time Markov Chains* (Springer-Verlag)
[7] Khurshidov A S 2001 *The Essential Guide to Computer Data Storage* (Prentice Hall)
[8] Schmidt K 2006 *High-availability and disaster recovery* (Springer-Verlag)
[9] Rahman P A and Bobkova E Yu 2016 *IOP Conf. Ser.: Mater. Sci. Eng.* 124 012023
[10] Rahman P A, Muraveva E A and Sharipov M I 2016 *Key Eng. Mat.* 685 805-810