Title

Evaluating the Impact of Irrecoverable Read Errors on Disk Array Reliability

Permalink

https://escholarship.org/uc/item/6fq0c1wq

ISBN

978-0-7695-3849-5

Authors

Pâris, Jehan-François
Amer, Ahmed
Long, Darrell DE
et al.

Publication Date

2009-11-01

DOI

10.1109/prdc.2009.66

Peer reviewed

Evaluating the Impact of Irrecoverable Read Errors on Disk Array Reliability

Jehan-François Pâris
Computer Science Dept.
University of Houston
Houston, TX 77204
paris@cs.uh.edu

Ahmed Amer[1]
Computer Science Dept.
University of California
Santa Cruz, CA 95064
a.amer@acm.org

Darrell D. E. Long[2]
Computer Science Dept.
University of California
Santa Cruz, CA 95064
darrell@cs.ucsc.edu

Thomas J. E. Schwarz
Computer Eng. Dept.
Santa Clara University
Santa Clara, CA 95053
tjschwarz@scu.edu

Abstract—We investigate the impact of irrecoverable read errors—also known as bad blocks—on the MTTDL of mirrored disks, RAID level 5 arrays and RAID level 6 arrays. Our study is based on the data collected by Bairavasundaram et al. from a population of 1.53 million disks over a period of 32 months. Our study indicates that irrecoverable read errors can reduce the mean time to data loss (MTTDL) of the three arrays by up to 99 percent, effectively canceling most of the benefits of fast disk repairs. It also shows the benefits of frequent scrubbing scans that map out bad blocks thus preventing future irrecoverable read errors. As an example, once-a-month scrubbing scans were found to improve the MTTDL of the three arrays by at least 300 percent compared to once-a-year scrubbing scans.

Keywords-disk arrays; mirrored disks; RAID arrays.

I. INTRODUCTION

Most efforts aimed at improving the reliability of disk arrays have focused on reducing the impact of disk failures. For instance, RAID level 5 arrays were designed to tolerate one disk failure without losing any data and RAID level 6 arrays were designed to tolerate two simultaneous disk failures. At the same time, much less attention has been given to protecting data against irrecoverable read errors. This situation is changing now due to several factors. First, today's disks are much larger than they were five to ten years ago. As a result, the probability of encountering one or more bad blocks on a given disk is much higher now than it was then. Second, our expectations have changed; much more data are now stored online than ten years ago and we expect these data to survive for several decades. Hence, yearly data loss rates that could have been tolerated ten years ago are now unacceptable. Finally, the cost of storage systems has become sufficiently low to allow us to increase their level of redundancy.

We present here the first inclusive study of the impact of irrecoverable read errors on disk array reliability. It covers mirrored disks, RAID level 5 arrays and RAID level 6 arrays. Unlike previous studies that either relied on manufacturer's data [3] or used a rough estimate of the frequency of irrecoverable read errors [16], our three stochastic models use as inputs the more recent data collected by Bairavasundaram *et al.* from a population of 1.53 million disks over a period of more than two years. Their study offers the first reliable data on the percentage of affected disks and the number of bad blocks per affected disk [1].

Our study indicates that irrecoverable read errors can reduce the mean time to data loss (MTTDL) of the three arrays by up to 99 percent, effectively canceling most of the benefits of fast disk repairs. It also demonstrates the benefits of frequent scrubbing scans that map out bad blocks. For instance, once-a-month scrubbing scans were found to improve the MTTDL of the three arrays by at least 300 and up to 980 percent compared to once-a-year scans.

The remainder of this paper is organized as follows. Section II reviews previous work. Section III introduces our three probabilistic models and presents their results. Section IV discusses the limitations of our approach. Finally, Section V has our conclusions.

II. PREVIOUS WORK

Irrecoverable read errors occur whenever a previously written block cannot be read due either to some malfunction during the write process or later damage to the disk surface. They are also known as bad blocks or latent sector errors because the data loss is not detected until the block is accessed. Irrecoverable read errors are particularly harmful when they occur during the data reconstruction phase of a RAID level 5 array that has one failed disk, as they result in a data loss [8].

Disk scrubbing [16] reduces the likelihood of irrecoverable read errors by periodically reading the contents of a whole disk, detecting bad blocks and reconstructing them in some of the disk spare blocks. It assumes that we have a way to reconstruct the lost data using either an inter-disk or intra-disk parity scheme. Lacking any solid data about the frequency of irrecoverable read errors, Schwarz *et al.* estimated that block failure rates were "about five times higher than disk failure rates" [16].

Baker *et al.* [3] enumerated the multiple threats to data survivability in disk arrays and presented a window of vulnerability model (WOV) that takes into account temporality of faults. They used that model to evaluate the impact of

---

[1] Supported in part by the National Science Foundation under Award Number 0720578.

[2] Supported in part by the Petascale Data Storage Institute under Department of Energy Award DE-FC02-06ER25768.

latent block errors on the MTTDL of mirrored data and based their estimate of the frequency of latent block errors on manufacturers' specification of a 2^{-14} worst-case irrecoverable bit rate.

Bairavasundaram *et al*. collected failure data from a population of 1.53 million disks over a period of 32 months [1]. Their main observations were that "[a] total of 3.45% of 1.53 million disks developed latent sector errors over a period of 32 months" and "[f]or most disk models, more than 80% of disks with latent sector errors have fewer than 50 errors."

Elerath identified major categories of disk failure modes and discussed the impact of latent defects [5]. He observed that read error rates had actually decreased between 2004 and 2007 and reported read-error rates varying between 3.2×10^{-13} and 8×10^{-15} errors/byte. In a more recent study [6], he introduces a formula that provides a good approximation of the expected number of data losses caused by double disk failures for an $N + 1$ RAID array. His model takes into account latent disk errors that reveal themselves during the array rebuilding process and assumes that disk failures obey a Weibull distribution.

III. OUR MODEL

We will base our model on the observations of Bairavasundaram *et al*. because they were collected from a very large disk population and offer us a better insight on the effect of bad blocks on disk array reliability.

Consider for instance the case of a RAID level 5 array with $N + 1$ disks. We know that the array can reconstitute missing data as long as each block stripe participating in the reconstruction process contains only one block that cannot be read. This is why such arrays can tolerate a single disk failure but neither a double disk failure nor the combination of a single disk failure and one or more bad blocks on one of the N remaining disks. Let us now consider what happens when all its disks are operational but two of them contain bad blocks. Thanks to the built-in redundancy of the array, we will be able to reconstitute all the lost data as long as no block stripe contains more than one bad block.

Assuming that a fraction f of the blocks of a disk are bad, the probability of observing more than one bad block in a single stripe will be

$$n_b(1-(1-f)^{N+1}-(N+1)f(1-f)^N),$$

where n_b is the number of stripes in the array (and thus the number of blocks on each disk) and $N + 1$ its number of disks.

Consider now the case of a RAID array consisting of five one-terabyte disks with a block size of four kilobytes. The formula above shows that we would need to observe at least 495 bad blocks on each disk to have a one percent probability of having two bad blocks in the same stripe. Given that Bairavasundaram *et al*. observed an infection rate of 3.54 percent over 32 months and reported that 80 percent of the infected disks had less than 50 bad blocks, we can safely infer that the occurrence of two bad blocks in the same stripe will be an extremely unlikely event.

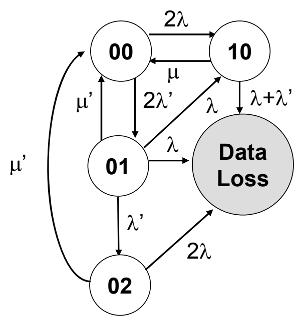

Fig. 1. State-probability transition diagram for a pair of mirrored disks subject to disk failures and irrecoverable read errors.

A. *Fundamental assumptions*

Estimating the reliability of a storage system means estimating the probability $R(t)$ that the system will operate correctly over the time interval $[0, t]$ given that it operated correctly at time $t = 0$. Computing that function requires solving a system of linear differential equations, a task that becomes quickly unmanageable as the complexity of the system grows. A simpler option is to focus on the mean time to data loss (MTTDL) of the storage system, which is the approach we will take here.

Our system model consists of an array of disks with independent failure modes. When a disk fails, a repair process is immediately initiated for that disk. Should several disks fail, the repair process will be performed in parallel on those disks. We assume that disk failures are independent events and are exponentially distributed with mean λ. In addition, we require repairs to be exponentially distributed with mean μ. Both hypotheses are necessary to represent each system by a Markov process with a finite number of states.

We assume that each disk is initially in a state where all its defective blocks have been "mapped out" in a way that prevents users from accessing them. We further assume that the apparitions of bad blocks on disks are independent events and are exponentially distributed with mean λ'. This is to say that we are not modeling the formation of individual bad blocks but rather the transition from a state where all defective blocks have been mapped out to a state where the disk has one or more bad blocks holding data. Once a disk has one or more bad blocks, it remains in that state until the missing data are reconstructed and written in one of the spare sectors of the disk. This recovery could be the result of a failed read access or a periodic scrubbing of the whole disk [16]. Combining the effect of these two processes we assume that disks with bad blocks will return to their original state at a rate μ' and that these transitions follow an exponential law.

B. *Mirrored disks*

The simplest redundant data organization consists of replicating data on two disks. As Fig. 1 shows, the pair of disks can be at any time in one out of five possible states, namely:

a) A state where both disks are operational and all their bad blocks have been mapped out, that is, state <00>;
b) A state where both disks are operational and one of them has bad blocks holding data, that is, state <01>;
c) A state where both disks are operational and both of them have bad blocks holding data, that is, state <02>;
d) A state where one disk has failed and the other has no bad blocks holding data, that is, state <10>;
e) A failed state that corresponds to a data loss.

A failure from one of the two disks brings the pair of disks from state <00> to state <10>. The system will remain in that state until:

a) The surviving disk fails and the system incurs a data loss;
b) The surviving disk develops bad blocks and the system incurs a data loss;
c) The failed disk is replaced and the system returns to its original state <00>.

Returning to state <00>, we see that the system will go from state <00> to state <01> when one of the two disks develops bad blocks and then from state <01> to state <02> when the second disk has bad blocks. Observe that state <02> is not a failed state because it is extremely unlikely that the two copies of the same block will be both unreadable. At the same time, a failure of either of the two disks will result in a data loss.

Recall that state <01> has one disk with bad blocks and one disk free of them. A failure of the disk with bad blocks will bring the system to state <10> while a failure of the other disk will result in a data loss.

Note that all repair transitions return the system to its original state.

The Kolmogorov system of differential equations describing the behavior of the pair of disks is

$$\frac{dp_{00}(t)}{dt} = -(2\lambda + 2\lambda')p_{00}(t) + \mu p_{10}(t) + \mu'(p_{01}(t) + p_{02}(t)),$$

$$\frac{dp_{10}(t)}{dt} = -(\lambda + \lambda' + \mu)p_{10}(t) + 2\lambda p_{00}(t) + \lambda p_{01}(t),$$

$$\frac{dp_{01}(t)}{dt} = -(2\lambda + \lambda' + \mu')p_{01}(t) + 2\lambda p_{00}(t),$$

$$\frac{dp_{02}(t)}{dt} = -(2\lambda + \mu')p_{02}(t) + \lambda' p_{01}(t),$$

where $p_{ij}(t)$ is the probability that the system is in state <ij> with the initial conditions $p_{00}(0) = 1$ and $p_{ij}(0) = 0$ otherwise. The Laplace transforms of these equations are

$$sp_{00}^*(s) - 1 = -(2\lambda + 2\lambda')p_{00}^*(s) + \mu p_{10}^*(s) + \mu'(p_{01}^*(s) + p_{02}^*(s)),$$

$$sp_{10}^*(s) = -(\lambda + \lambda' + \mu)p_{10}^*(s) + 2\lambda p_{00}^*(s) + \lambda p_{01}^*(s),$$

$$sp_{01}^*(s) = -(2\lambda + \lambda' + \mu')p_{01}^*(s) + 2\lambda' p_{00}^*(s),$$

$$sp_{02}^*(s) = -(2\lambda + \mu')p_{02}^*(s) + \lambda' p_{01}^*(s).$$

Observing that the mean time to data loss (MTTDL) of the pair of disks is

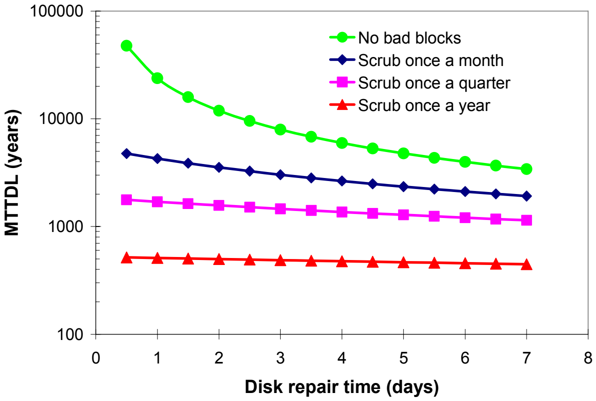

Fig. 2. Expected MTTDL of a pair of mirrored disks subject to disk failures and irrecoverable read errors. Each curve but the topmost corresponds to a specific scrubbing interval.

$$MTTDL = \sum_{i,j} p_{ij}^*(0),$$

we solve the system of Laplace transforms for $s = 0$ and use this result to obtain

$$MTTDL = \frac{6\lambda^2 + 3\lambda\lambda' + \lambda'^2 + 2\lambda\mu + 3\lambda\mu' + \lambda'\mu + \lambda'\mu' + \mu\mu'}{2\lambda(2\lambda^2 + 3\lambda\lambda' + \lambda'^2 + \lambda\mu' + \lambda'\mu)}.$$

When $\mu' \to \infty$, the above expression simplifies into the traditional formula for mirrored disks

$$MTTDL = \frac{3\lambda + \mu}{2\lambda^2}.$$

Fig. 2 displays the MTTDLs achieved by the pair of disks for selected values of the scrubbing interval and repair times that vary between half a day and seven days. We assumed that the disk failure rate λ was one failure every one hundred thousand hours, that is, slightly less than one failure every eleven years, which is consistent with the values reported by Schroeder and Gibson [12, 13] and Pinheiro *et al.* [11]. The bad block formation rate λ' was set to 1.294 percent per year, a value based on the measurements of Bairavasundaram *et al.* [1]. Disk repair times are expressed in days and MTTDLs expressed in years. The topmost curve corresponds to the ideal case of a mirrored pair that will never experience irrecoverable read errors while the three other curves correspond to specific scrubbing intervals varying between once a month and once a year.

As we can see, irrecoverable read errors have a devastating effect on the MTTDL of the mirrored pair. The sole effective countermeasure seems to be increasing the frequency of scrubbing scans from once a year to once a quarter or once a month. For instance, the MTTDL of a mirrored pair that is scrubbed once a year is between 87 and 98 percent shorter than that of a pair that would never experience irrecoverable read errors. Having scrubbing scans once a month instead of once a year would increase its MTTDLs by more than 300 to 800 percent depending on the disk repair time.

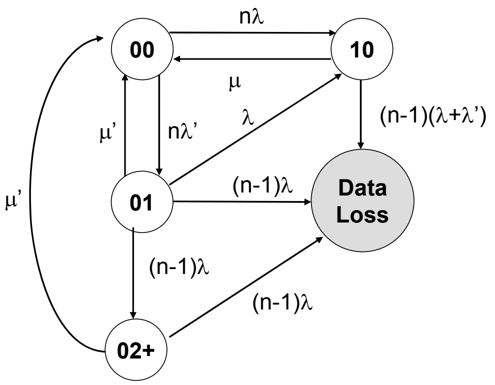

Fig. 3. State-probability transition diagram for a RAID level 5 array consisting of *n* disks, all subject to disk failures and irrecoverable read errors.

We also note that the MTTDLs of mirrored pairs are much less affected by the disk repair time than they would have been in the absence of irrecoverable read errors.

C. RAID level 5

Let us now consider the case of a RAID level 5 array with *n* disks [4, 7, 10, 14, 15]. As Fig. 3 shows, the array can be at any time in one out of five possible states, namely:

a) A state where all *n* disks are operational and all their bad blocks have been mapped out, that is, state <00>;

b) A state where all *n* disks are operational and one of them has bad blocks, that is, state <01>;

c) A state where all *n* disks are operational and two or more of them have bad blocks, that is, state <02+>;

d) A state where one of the *n* disks has failed and none of the $n - 1$ other disks has bad blocks, that is, state <10>;

e) A failed state that corresponds to a data loss.

A failure from one of the *n* disks brings the pair of disks from state <00> to state <10>. The system will remain in that state until:

a) One of the $n - 1$ surviving disks develops bad blocks and the system incurs a data loss;

b) The failed disk is replaced and the system returns to its original state <00>.

Returning to state <00>, we see that the array will go from state <00> to state <01> when one of its *n* disks develops bad blocks. From state <01>, the array will:

a) Go to state <10> if the disk that has bad blocks fail;

b) Incur a data loss if any of the $n - 1$ other disks fail;

c) Go to state <02+> if more disks develop bad blocks;

d) Return to its original state once all bad blocks have been mapped out.

Observe that state <02+> comprises all configurations where all *n* disks are operational and two or more of them have bad blocks It is not a failed state because it is rather unlikely that the array will have two or more bad blocks in the same stripe. An array in state <02+> will remain in that state until:

a) One of its *n* disks fails and the system experiences a data loss or

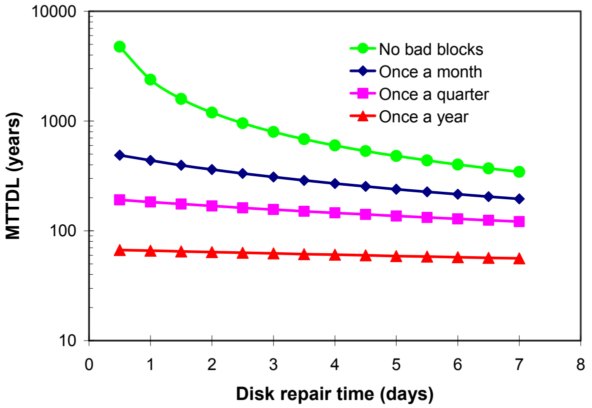

Fig. 4. Expected MTTDL of a RAID level 5 array consisting of five disks, all subject to disk failures and irrecoverable read errors. Each curve but the topmost corresponds to a specific scrubbing interval.

b) The bad blocks are mapped out and the system returns to its original state.

Using the same techniques as in the previous case, we compute the MTTDL of the array for $n = 5$, and obtain a quotient of two very large polynomials of degree 3 for the numerator and of degree 4 for the denominator.

When $\mu' \to \infty$, that expression simplifies into the traditional formula for RAID level 5 arrays

$$MTTDL = \frac{9\lambda + \mu}{20\lambda^2} = \frac{(2n-1)\lambda + \mu}{n(n-1)\lambda^2}.$$

Fig. 4 displays the MTTDLs achieved by the disk array for selected values of the scrubbing interval and repair times that vary between half a day and seven days. We assumed the same disk failure rate λ and the same bad block formation rate λ' as in the previous case. As before, we notice the devastating effect of irrecoverable read errors and the benefits of shorter scrubbing intervals. Having scrubbing scans once a month instead of once a year can increase the MTTDL of the array from 248 to 630 percent depending on the disk repair time.

D. RAID level 6

Let us now consider the case of a RAID level 6 with *n* disks [2]. As Fig. 5 shows, the array can be at any time in one out of eight possible states.

As before, state <00> represents the initial state of the system when the *n* disks are operational and all their bad blocks are mapped out. The first disk failure will bring the array to state <10>. A second disk failure will bring it to state <20> and a third disk failure will result in a data loss.

Formation of bad blocks on a single disk will respectively bring the array from state <00> to state <01> from state <10> to state <11> and from state <20> to the failed state. Similarly, the apparition of bad blocks on a second disk would respectively move the array from state <01> to state <02+> and from state <11> to state <12+>.

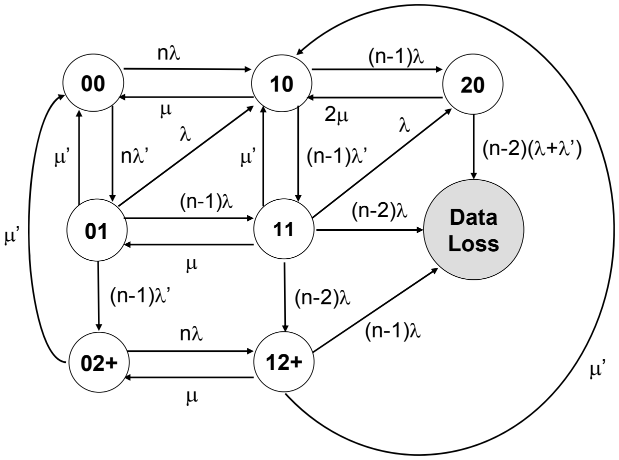

Fig. 5. State-probability transition diagram for a RAID level 6 array consisting of *n* disks, all subject to disk failures and irrecoverable read errors.

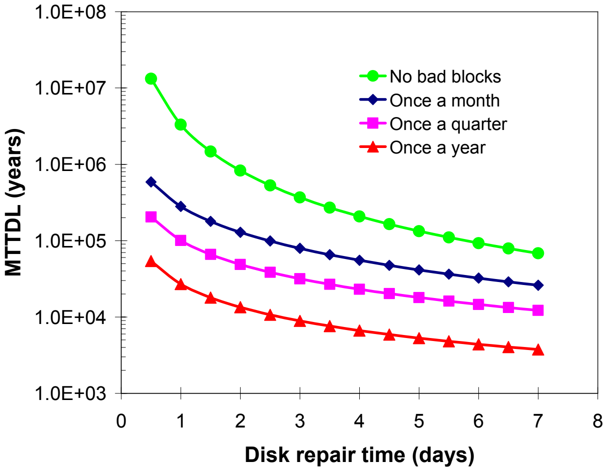

Fig.6. Expected MTTDL of a RAID level 6 array consisting of six disks, all subject to disk failures and irrecoverable read errors. Each curve but the topmost corresponds to a specific scrubbing interval.

The apparition of bad blocks on a third disk will leave the system in either state <02+> or <12+> as we do not distinguish between array configurations having bad blocks on two disks and those having bad blocks on three or more disks.

From state <01>, the array can:

a) Go to state <10> if the disk that has bad blocks fail;
b) Go to state <11> if any of the $n - 1$ other disks fail;
c) Go to state <02+> if more disks develop bad blocks;
d) Return to its original state once all its bad blocks have been mapped out.

From state <11>, the array can:

a) Go to state <20> if the disk that has bad blocks fails;
b) Incur a data loss if any of the $n - 2$ other disks fails;
c) Go to state <12+> if one or more disks develop bad blocks;
d) Return to its original state once all its bad blocks have been mapped out.

Observe that our model assumes that a disk failure occurring when the array is in state <02+> will always bring that array into state <12+>. As a result, we neglect the less frequent case where the array has bad blocks on exactly two disks and one of them fails, thus bringing the array into state <11>. This observation does not apply to state <12+> since the loss of any of its $n - 1$ operational disks will result in a data loss.

Replacing the failed disks will bring the array from state <20> to state <10> and from state <10> to its original state. In the same way, scrubbing scans will bring the array from state <02+> to state <00> and from state <12+> to state <10>.

Using the same techniques as in the two previous cases, we compute the MTTDL of a RAID level 6 array with the same capacity as the RAID level 5 array in our previous example. It will comprise six disks instead of five to accommodate the additional parity information [2]. We obtain a quotient of two very large polynomials of degree 6 for the numerator and of degree 7 for the denominator.

Fig. 6 displays the MTTDLs achieved by the disk array for selected values of the scrubbing interval and repair times that vary between half a day and seven days. We observe again the devastating effects of irrecoverable disk errors and the benefits of frequent scrubbings. Scheduling scrubbing scans once a month instead of once a year can increase the MTTDL of the array from 597 to 987 percent depending on the disk repair time.

A major difference with the two previous cases is the more pronounced effect of the disk repair time. This is because RAID level 6 arrays are only affected by irrecoverable read errors after they have lost two disks. Quickly replacing the first failed disk greatly reduces the risk of experiencing transitions from states <11> or <12+> to the failure state.

IV. DISCUSSION

In order to be able to use stochastic models with finite numbers of states, we had to introduce some assumptions that are not true for real systems. First, we assumed that failure occurrences and repair times were exponentially distributed. This is not true for real disk populations as failures tend to be distributed according to a Weibull distribution and repair time distributions have much smaller coefficients of variation than the exponential distribution. Second, we assumed constant failure rates λ and λ' over the lifetime of the array, while actual failure rates tend to decrease over the first few months of the disk lifetime and increase again after a few years.

Other simplifying assumptions were the result of a lack of data. First, the failure rates λ and λ' that we selected were average rates estimated over very large heterogeneous populations of disks comprising both enterprise class and nearline class disks. Second, we did not take into account the impact of correlated failures on the MTTDL of the array.

This could have been handled by making the two failure rates λ and λ' functions λ(*m*) and λ'(*m*) of the number *m* of previous failures. We could not do this due to insufficient data.

A last issue concerns our choice of MTTDL to represent the reliability of disk arrays. MTTDLs characterize fairly well the behavior of disk arrays that would remain in service until they fail without being ever replaced for any reason other than a device failure. This is rarely the case as disk arrays are typically replaced after five to seven years, that is, well before they experience any failure. MTTDLs do not take into account this relatively short lifetime, and tend to overestimate the probability of a data loss over this lifetime. This effect remains negligible as long as the time to repair an individual disk is at least one thousand times shorter than its MTTF [9].

V. CONCLUSION

We have presented the first comprehensive study of the impact of irrecoverable read errors on the MTTDL of mirrored disks, RAID level 5 arrays and RAID level 6 arrays using the data collected by Bairavasundaram *et al*. from a population of 1.53 million disks over a period of 32 months. Our study has shown the dramatic impact of these errors, which can reduce the mean time to data loss (MTTDL) of the three arrays by up to 99 percent. It also shows the dramatic benefits of more frequent scrubbing scans for preventing future irrecoverable read errors. As an example, once-a-month scrubbing scans were found to improve the MTTDL of the three arrays by at least 300 percent and up to 900 percent compared to once-a-year scrubbing scans.

More work needs to be done to investigate the benefits of adding extra levels of redundancy either inside each disk [16] or among multiple arrays [17].

VI. REFERENCES

[1] L. Bairavasundaram, G. Goodson, S. Pasupathy, J. Schindler, "An analysis of latent sector errors in disk drives," *Proc. ACM SIGMETRICS 2007 International Conference on Measurement and Modeling of Computer Systems*, San Diego, CA, 2007.

[2] W. A. Burkhard and J. Menon, "Disk array storage system reliability," *Proc. 23rd International Symposium on Fault-Tolerant Computing*, Toulouse, France, pp. 432–441, June 1993.

[3] M. Baker, M. Shah, D. S. H. Rosenthal, M. Roussopoulos, P. Maniatis, T. J. Giuli, and P. Bungale, "A fresh look at the reliability of long-term storage," *Proc. First EuroSys Conference* (EuroSys 2006), Leuven, Belgium, Apr. 2006.

[4] P. M. Chen, E. K. Lee, G. A. Gibson, R. Katz, and D. Patterson, "RAID: High-performance, reliable secondary storage," *ACM Computing Surveys*, 26(2):145–185, 1994.

[5] J. G. Elerath, "Hard disk drives: The good, the bad and the ugly!" *ACM Queue* 5(6):28–37, Sep./Oct., 2007.

[6] J. G. Elerath, "A simple equation for estimating the reliability of an N+1 redundant array of independent disks (RAID)," *Proc. 39th nternational Conference on Dependable Systems and Networks* (DSN 2009), Estoril, Portugal, June 2009.

[7] G. A. Gibson, "Redundant disk arrays: Reliable, parallel secondary storage," *Ph.D. thesis*, University of California at Berkeley, 1990.

[8] R. Harris, "Why RAID 5 stops working in 2009," ZDnet, http://blogs.zdnet.com/storage/?p=162, July 18, 2007 (Retrieved on May 27, 2009)

[9] J.-F. Pâris, T. J. E. Schwarz, D. D. E. Long and A. Amer, "When MTTDLs are not good enough: Providing better estimates of disk array reliability," *Proc. 7th International Information and Telecommunication Technologies Symposium* (I2TS '08), Foz do Iguaçu, PR, Brazil, pp. 140–145, Dec. 2008.

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

[11] E. Pinheiro, W.-D. Weber and L. A. Barroso, "Failure trends in a large disk drive population," *Proc. 5th USENIX Conference on File and Storage Technologies* (FAST '07), San Jose, CA, Feb. 2007.

[12] B. Schroeder and G. A. Gibson, "A large-scale study of failures in high-performance computing systems," *Proc. 40th Annual IEEE/IFIP International Conference on Dependable Systems and Networks* (DSN 2006), Philadelphia, PA, pp. 249–258 , June 2006.

[13] B. Schroeder and G. A. Gibson, "Disk failures in the real world: What does an MTTF of 1,000,000 hours mean to you?" *Proc. 5th USENIX Conference on File and Storage Technologies* (FAST '07), San Jose, CA, pp. 1–16, Feb. 2007.

[14] M. Schulze, G. Gibson, R. Katz and D. Patterson, "How reliable is a RAID?" *Proc. Spring COMPCON '89 Conference*, San Francisco, CA, pp. 118–123, Mar. 1989.

[15] T. J. E. Schwarz and W. A. Burkhard. "RAID organization and performance," *Proc. 12th International Conference on Distributed Computing Systems*, Yokohama, Japan, pp. 318–325 June 1992.

[16] T. J. E. Schwarz, Q. Xin, E. L. Miller, D. D. E. Long, A. Hospodor and S. Ng, "Disk scrubbing in large archival storage systems," *Proc. 12th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems* (MASCOTS '04), pages 409-418, Oct. 2004.

[17] A. Wildani, T. J. E. Schwarz, E. L. Miller and D. D. E. Long, "Protecting against rare event failures in archival systems," *Proc. 17th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems* (MASCOTS '09), London, UK, Sep. 2009, *to appear*.