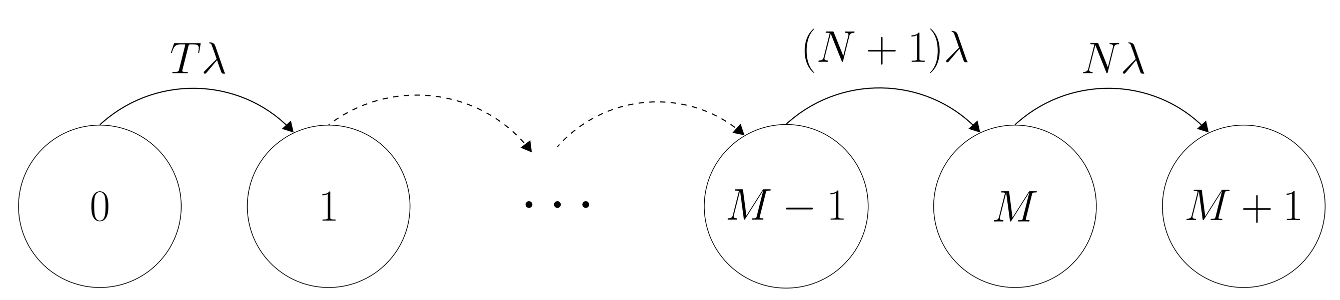
On the Reliability of RAID Systems: An Argument for More Check Drives

Sarah Edge Mann
Program in Applied Mathematics
University of Arizona

Introduction

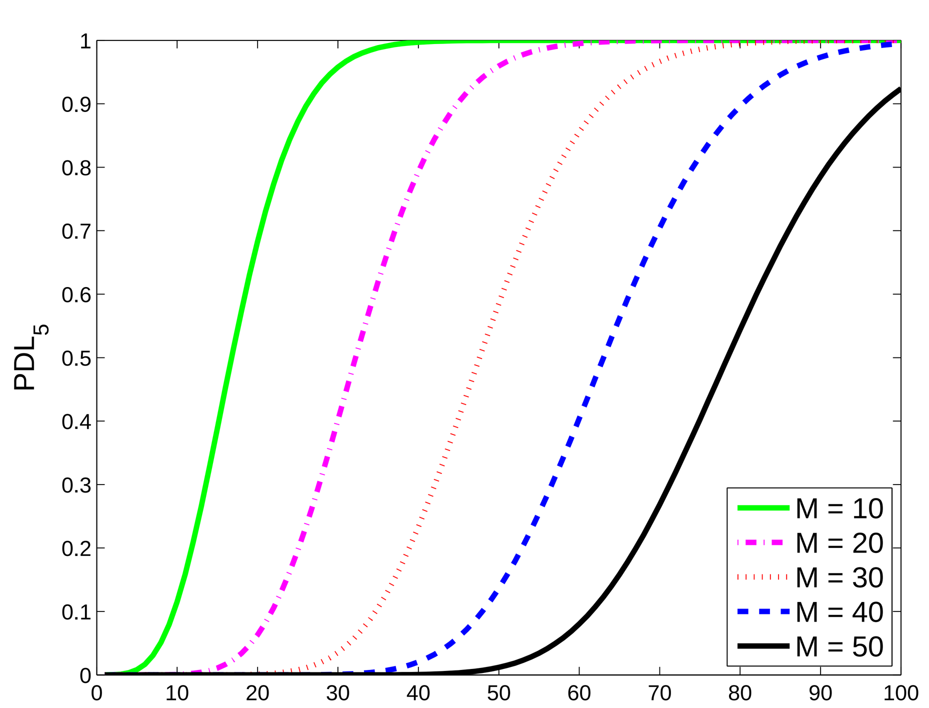
RAID technology has a single primary focus: to apply mathematical techniques to organize data on multiple data storage devices such that in the event of one or more device failures, the original data stored is still available. In this project, we attempt to quantify the increased reliability that is achieved by constructing RAID systems with more robust error correcting codes. We assume that a RAID system consists of N data drives and M check drives containing redundant data for a total of $T = N + M$ drives. In the face of up to M drive failures, all original data may be recovered from the remaining N drives. A simultaneous failure of $M + 1$ or more drives results in unrecoverable data loss. We measure the reliability of a system by the probability of data loss within 5 years of deployment (PDL$_5$), and analyze how N, M, and various reliability models effect the predicted PDL$_5$.

Model 1: No Repair



- Failed hard drives are never repaired
- Model as a discrete state, continuous time Markov process with $M + 2$ states
- State i indicates that i drives have failed
- System is initialized in state 0 (all drives working)
- Assume drives fail independently at a constant failure rate λ per drive
- System moves from state i to $i + 1$ with an *effective failure rate* $\lambda_i = (T - i)\lambda$
- State $M + 1$ is the *failure state*

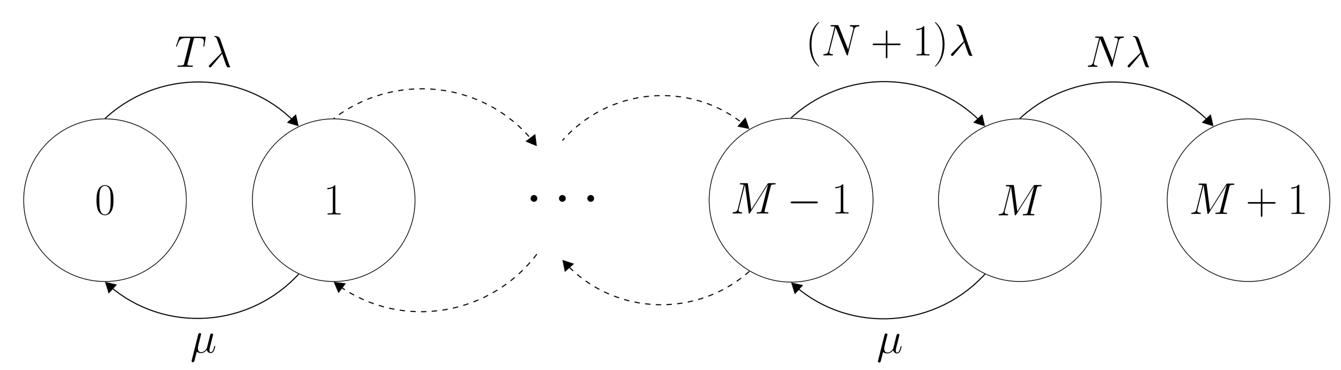
We may numerically calculate PDL$_5$ using the Kolmogorov-Chapman equations and other standard techniques for Markov chains. This figure shows PDL$_5$ under a no repair model as a function of N for five values of M with $\lambda = \frac{1}{10 \text{ years}}$. Notice that to maintain a particular level of reliability (PDL$_5$ value), more check drives are required as the number of data drives increase.



List of Symbols

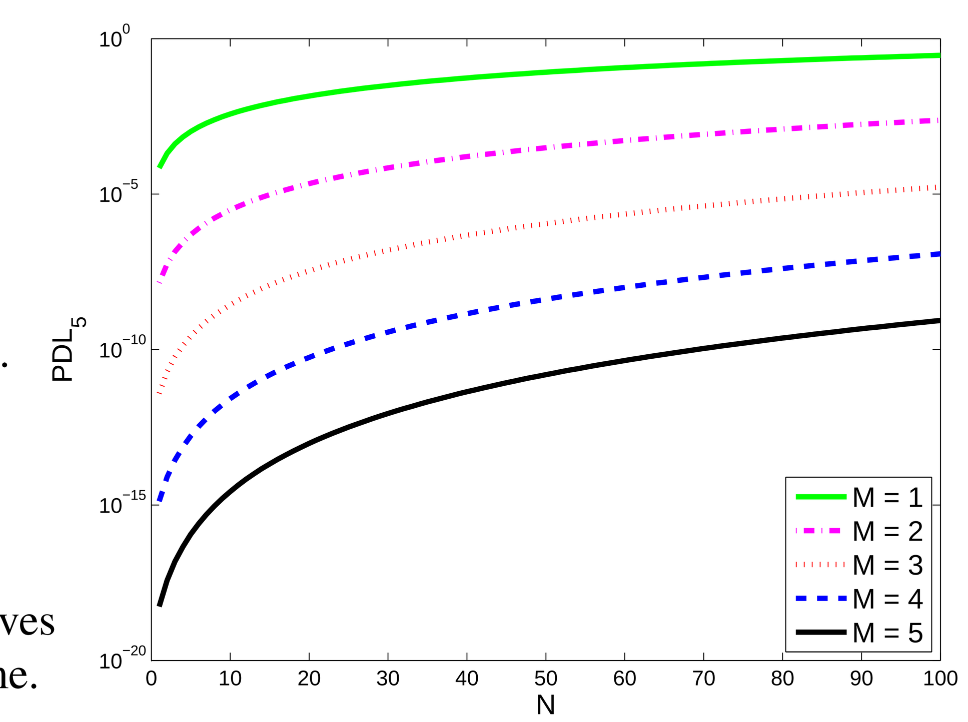
N	Number of data drives	λ	Individual drive failure rate
M	Number of check drives	μ	Drive repair rate
T	Total number of drives	PDL$_5$	Probability of data loss within 5 years

Model 2: Individual Drive Repair



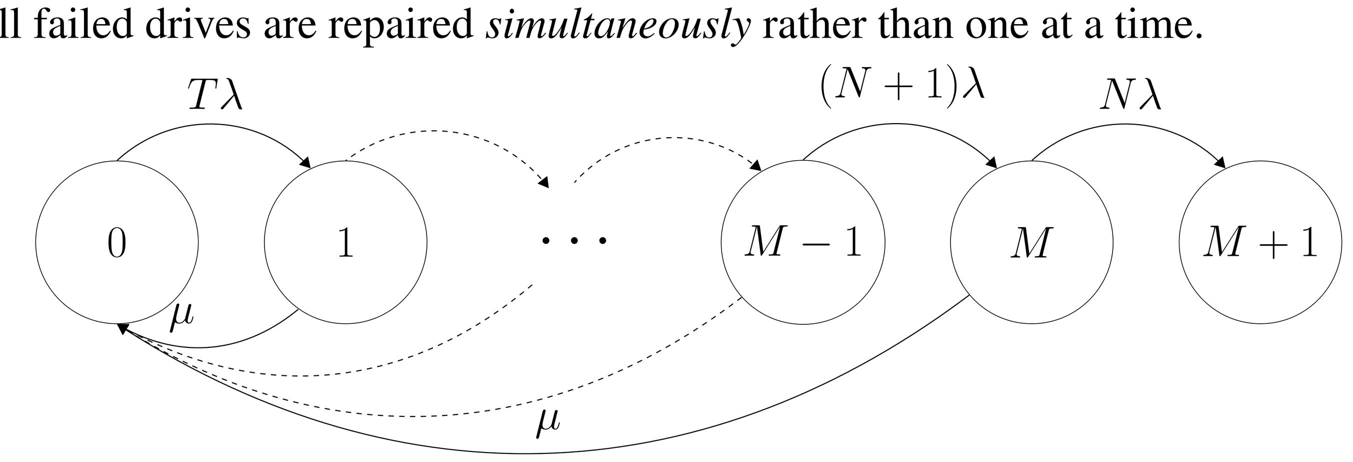
- Model of drive failure is the same as in the no repair model
- Failed drives are repaired *one at a time*
- When a drive is repaired the system moves from state i to state $i - 1$
- Assume drives are repaired at a constant rate μ, independent of the number of failed drives
- System moves from state i to $i - 1$ with *effective repair rate* $\mu_i = \mu$

This figure shows PDL$_5$ as a function of N for five values of M with $\lambda = \frac{1}{10 \text{ years}}$ and $\mu = \frac{1}{6 \text{ hours}}$. Notice that these curves are spaced evenly apart for PDL$_5$ on a logarithmic scale. This indicates that a RAID system under the individual repair model with $M+1$ check drives is *exponentially* better than a RAID system with M check drives and all other parameters the same.



Model 3: Simultaneous Repair

- Build on individual drive repair model except ...
- All failed drives are repaired *simultaneously* rather than one at a time.



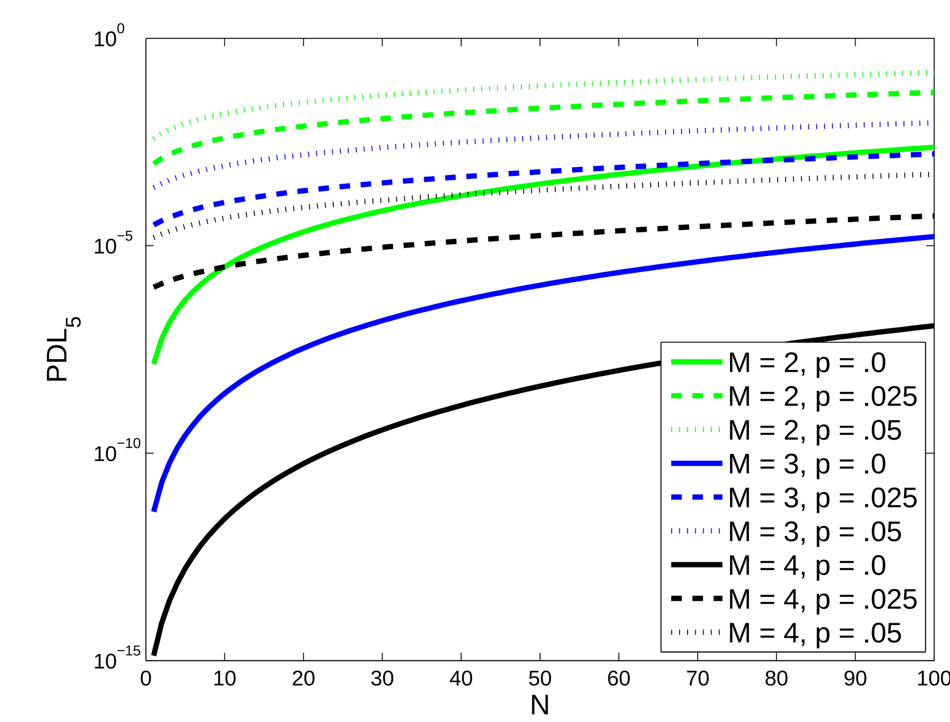
The effects on predicted reliability of the RAID as a result of this change to the model are negligible. For example, with $M = 5$, the PDL$_5$ for the simultaneous repair model is 0 - 3% lower than for the individual repair model, and the effect grows linearly with N. The same relationship holds for other values of M with a smaller constant of proportionality for smaller M. For large M this effect might be significant but the reliability model is not sensitive to this modification.

Model 4: Imperfect Repair

Consider the effects of human mistakes made while servicing RAID systems:

- Build on simultaneous repair model but use different effective failure and repair rates
- Probability p that in servicing failed drives a working drive is damaged and the failed drives are not repaired
- Probability $1 - p$ that failed drives will successfully be repaired
- Effective failure rate: $\lambda_0 = T\,\lambda$, $\lambda_j = (T - j)\,\lambda + \mu\,p$ for $j > 0$
- Effective repair rate: $\mu_j = \mu\,(1 - p)$

This figure shows the effects on PDL$_5$ of considering imperfect repair. Notice that even for p small, imperfect repair decreases the reliability of the system by several orders of magnitude. Doubling p decreases the reliability by at least one further order of magnitude. For larger M the effect is more pronounced with a decrease in reliability of as much as 10 orders of magnitude.



Other Effects Modeled

- Sector errors: a small portion of a hard drive becomes unreadable and additional protection is required to recover lost data.
- Delay of service: hard drives are repaired infrequently.
- Rebuild time: drives take some minimum amount of time to repair during which time other drives may fail.
- Silent data corruption: data may be written or read incorrectly. Additional data redundancy allows this to be detected and corrected.

Conclusions

- For a fixed number of data drives N, increasing the number of check drives M dramatically increases the reliability of the RAID.
- For a fixed total number of drives T, a single large RAID system is more reliable than two RAID systems with the same data rate N/T.
- The reliability of RAID systems is often overstated. To ensure persistence of important data, more check drives are required than are typically deployed.

This work was done in collaboration with Dr. Marek Rychlik and Dr. Michael H. Anderson, CEO, with support from StreamScale, Inc and the National Science Foundation. For more on RAID reliability, see our paper at `http://arxiv.org/abs/1202.4423`