

# Reliability Markov models are becoming unreliable (WIP submission)

Kevin Greenan and Jay J. Wylie

Hewlett-Packard Labs

kmgreen@soe.ucsc.edu, jay.wylie@hp.com

Markov models have traditionally been used to understand the reliability of storage systems. They provide intuition about the sensitivity of storage system reliability to changes in disk failure rates, rebuild rates, sector failure rates, scrubbing rates, and storage capacity. Unfortunately, as we move towards multi-disk fault tolerant storage systems, i.e., storage systems that tolerate two or more disk failures such as RAID 6, *reliability estimates* based on traditional Markov models become *unreliable*. Our concerns go beyond the recent demonstration that Weibull distributions need to be used instead of exponential distributions to accurately determine storage system reliability [1]. We believe that the traditional construction of Markov models is flawed for multi-disk fault tolerant systems, and that their accuracy and utility decreases as the redundancy in the system increases.

In this WIP, we will only discuss one of our concerns: modeling disk rebuild correctly. Two traditional Markov models are used to model two distinct storage rebuild policies. In a *serial rebuild policy*, a storage system rebuilds the first failed disk in its entirety before rebuilding the next failed disk, and so on. In a *concurrent rebuild policy*, a storage system begins rebuilding failed disks as they fail.

Figure 1 illustrates the two traditional Markov models for an  $n$  disk system that tolerates  $m$  disk failures. The label of each state indicates the number of failed disks; state  $m + 1$  is the data loss state. The transitions from left to right are disk failures, with  $\lambda$  being the failure rate. The transitions from right to left are disk rebuilds, with  $\mu$  being the rebuild rate.

For single disk fault tolerant systems, the serial and concurrent rebuild models are identical, and are correct. For multi-disk fault tolerant systems, both rebuild models are incorrect. The same modeling error is made in each case. The rebuild transitions for states 2 through  $m$  are incorrect: they model the rebuild of the disk that failed most recently, whereas reliability is dominated by the rebuild of the disk that failed earliest. In essence, traditional Markov models *reset* the rebuild time for all disks being rebuilt *whenever* another disk fails. The traditional serial rebuild Markov model thus models a rebuild policy

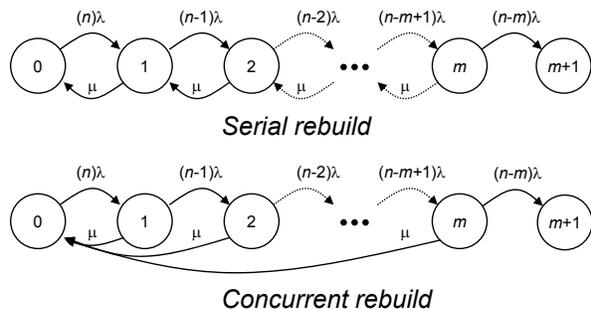


Figure 1: Traditional Markov models for rebuild policies.

in which each subsequent disk failure changes which disk is being rebuilt, and “re-fails” the disk currently being rebuilt. The traditional concurrent rebuild Markov model thus models a rebuild policy in which each subsequent disk failure restarts the rebuild of all failed disks.

The modeling error results in both traditional Markov models producing a similar, conservative reliability calculation. We believe that this explains why Hafner and Rao concluded that different rebuild policies did not lead to noticeably different reliability calculations [3]. We also believe that every additional disk of redundancy compounds the error due to incorrect Markov modeling of disk rebuild. We discuss all of our concerns with reliability Markov models and multi-disk fault tolerant systems in a pending technical report [2].

## References

- [1] J. F. Elerath and M. Pecht. Enhanced reliability modeling of raid storage systems. In *DSN-2007*, pages 175–184. IEEE, June 2007.
- [2] K. Greenan and J. J. Wylie. On the reliability of XOR-based erasure codes. Technical report, HP, February 2008.
- [3] J. L. Hafner and K. Rao. Notes on reliability models for non-MDS erasure codes. Technical Report RJ-10391, IBM, October 2006.