SSD-RAID

Roman Shor

May 22, 2016

Introduction

As the capacity of flash memory increases the reliability problem of flash devices is becoming increasingly serious. Limited number of P/E cycles of flash memory is decreasing according to enlarged storage capacities. BER rapidly increase with P/E cycles progress.

As a result schemes using internal RAID in SSD devices are proposed supplementary to the ECC as a way increasing reliability.

When using RAID technologies to obtain high reliability, side effect of parity pages updates leads to rapid increase in the P/E cycles and thus lifespan of flash memory is shortened. Also adding parity pages takes additional redundant space of the device, which can either be taken from the OP and lead to inefficient GC or be conserved in additional space meaning higher redundancy.

These are the main problems in adopting RAID technology in SSD.

Observations

During my research of the proposed schemes I have noticed two orthogonal improvements in RAID implementation.

1. First improvement and the most promising one, described in [3] is to reduce WA by allowing elastic striping and anywhere parity , meaning when page is updated it is written in a new stripe, which doesn't require parity update with every page update and thus reducing the number of redundant writes, without harming the reliability. This scheme also uses cache to allow writing long stripes, meaning page parity is accumulated and stored in some internal cache and physically written only when enough pages are written to fill a stripe, or some time threshold is passed. The cache method is first described in [2] where it allows to postponed the parity page update.

2. Second improvement is concentrated on other aspect of applying RAID scheme in SSD, in [4] the focus is on the changing BER in flash based SSDs during their lifespan. It is commonly believed and was shown in recent works [6, 5][6, 5] that BER rapidly increases with the growth of P/E cycles. So the solution to the high redundancy rate of the space taken by the RAID scheme proposed in [4], was to adapt the size of the stripe according to the current state of the device, meaning according to the P/E cycle of the device. So the UER (uncorrectable error rate) is constant throughout the whole lifespan of the SSD.

There were also other studies like [1] which concentrated on avoiding SSD uniform aging in order to avoid uncorrectable failures. In this scheme the main idea is to differentiate the load in between the RAID nodes (SSDs) so there will be only one SSD close to it's expiration P/E cycle, thus lowering the probability of multiple node failure.

Summary

In summary there are still a lot of unresolved problems in applying RAID as an internal or external reliability tool for SSDs.

The main problems remain the high overhead of space required to save the parity data combining with OP, impose low rate and high price for bit of information.

And the higher WA due to the constant need to update parity pages.

References

[1] Mahesh Balakrishnan, Asim Kadav, Vijayan Prabhakaran, and Dahlia Malkhi. Differential raid: Rethinking raid for ssd reliability. *Trans. Storage*, 6(2):4:1–4:22, July 2010.

[2] S. Im, S. Im, S. Im, S. Im, D. Shin, D. Shin, D. Shin, and D. Shin. Flash-aware raid techniques for dependable and high-performance flash memory ssd. *IEEE Transactions on Computers*, 60(1):80–92, Jan 2011.

[3] J. Kim, J. Lee, J. Choi, D. Lee, and S. H. Noh. Improving ssd reliability with raid via elastic striping and anywhere parity. In *2013 43rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pages 1–12, June 2013.

[4] Sehwan Lee, Bitna Lee, Kern Koh, and Hyokyung Bahn. A lifespan-aware reliability scheme for raid-based flash storage. In *Proceedings of the 2011 ACM Symposium on Applied Computing*, SAC '11, pages 374–379, New York, NY, USA, 2011. ACM.

[5] Justin Meza, Qiang Wu, Sanjev Kumar, and Onur Mutlu. A large-scale study of flash memory failures in the field. *SIGMETRICS Perform. Eval. Rev.*, 43(1):177–190, June 2015.

[6] Bianca Schroeder, Raghav Lagisetty, and Arif Merchant. Flash reliability in production: The expected and the unexpected. In *14th USENIX Conference on File and Stor-*

age Technologies (FAST 16), pages 67–80, Santa Clara, CA, February 2016. USENIX Association.