

Wear Leveling Technology

White Paper

January 8, 2009

Version 1.1

Wear Leveling Technology

Introduction

Flash media has many benefits over conventional rotating media such as compact size, lower power consumption, zero seek time, and vibration resistance that makes it ideal for the mobile environment including digital cameras, persona digital assistants, notebooks, and other embedded systems. However, a key limitation of flash media is a low erase-write endurance limit. For an SLC-based flash media, each block will typically wear out after 100K program/erase cycles. If a certain area on the flash media wears out faster the others, it would significantly reduce the lifetime of the whole device; even if their erase counts are far from the program/erase cycle limit. Thus, if the program/erase cycles are fairly distributed across the media, the overall lifespan of the media can be extended substantially, which is what wear leveling being managed to achieve.

The Need for Wear Leveling

Most widely used file systems were written without consideration of the endurance limitations of flash media. For example, the FAT 16 file system reserves the first few hundred sectors of a partition to hold the file allocation table and root directory of the file system. The FAT is a linked list of all sector clusters associated with a particular file. Every time a file is created, deleted, or updated this linked list must be modified. Once the flash sectors associated with the FAT have reached their endurance limit it is expected the whole file system to fail. Buffering changes to the FAT into RAM will help relieve the problem, but overall the FAT sectors will be written much more frequently than the data sectors.

A wear leveling algorithm manages the uneven “wear” on the sectors of a flash media memory by distributing the writes through many sectors of the flash media. Integrated in to the firmware of the flash disk controller, the wear leveling algorithm is transparent to the overlying file system by keeping track of a map between the logical sectors and the physical sectors that mp on the flash media. In the ideal case, a wear-leveling algorithm will result in all the sectors of the flash media reaching their endurance limit nearly simultaneously, maximizing the usable lifetime of the flash media. With the use of aging mechanisms, it is possible to warn the user when endurance limits are reached and to proactively backup the contents before it is lost for good.

Frameworks of Flash Media and SSD

Currently there are mainly two flash memory architectures available: NOR and NAND. Although NOR flash has been a popular choice for applications requiring conventional memory interface, NAND flash has shorter erase and write intervals, more compact storage densities, and economies of scale, as well as up to ten times the endurance of NOR flash, which makes NAND flash very suitable for use in nonvolatile solid-state drives (SSDs), giving much greater endurance than that of NOR flash.

SSD is electrically, mechanically, and application compatible with a conventional hard drive. The difference is that its storage medium is neither magnetic nor optical but the NAND flash memory components. An SSD drive provides faster access time than a hard drive, and it is also more tolerant of ambient temperature and vibration extremes. Also, it is nonvolatile and has significant low power consumption.

Two dominant NAND flash technologies are SLC (single-level cell) and MLC (multilevel cell). The SLC flash memory stores one bit per cell, while each cell of the MLC flash memory contains more than one bit of information (typically 2 bits). The endurance of a block of SLC flash memory is generally specified at 100K erase count, compared to the 10K erase count of an MLC flash block.

Memory Array Organization

The two-plane memory array is made up of cells connected in series to form a NAND structure. It consists of 8,192 erasable 256K-byte blocks. The 4,224-byte page registers are connected to cell arrays, allowing data transfer between I/O buffers and memory during page read and page write operations.

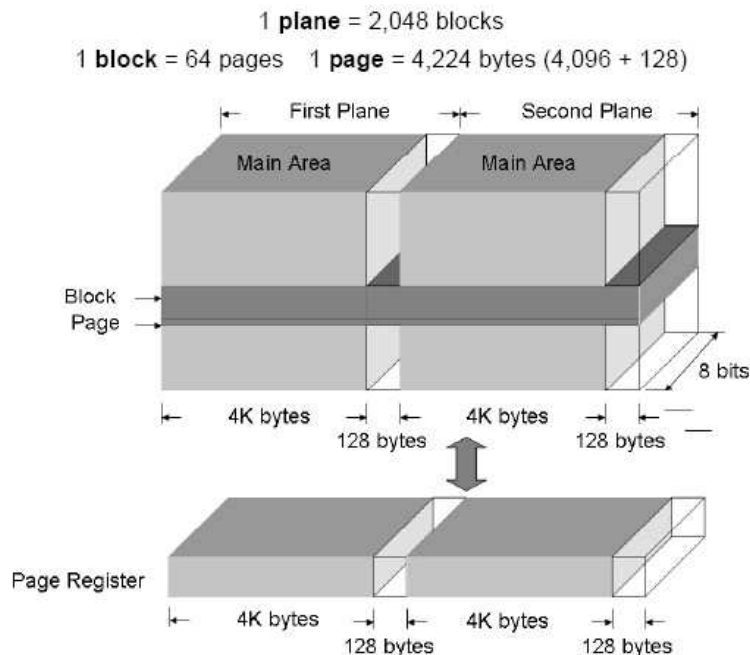


Figure 1: Two-plane NAND flash array organization

The memory array is so arranged that it can perform simultaneous page writes and block erases. The block address map is configured so that two-plane write/erase/read operations can

be performed by dividing the array into planes 0 and 1. During operations, data bits are kept in the cells of a NAND flash device. The data is written by page of cells. To update partial data in a page, an erase process on the block including the page must be executed before overwriting the page with its new contents.

Memory Management

For NAND flash memory, blocks are the smallest unit for erase operations, while reads and writes being performed on a page basis. Prior to writing pages, the block erase operation is required. To manage the NAND flash device, an emulating layer is implemented allowing the file systems to be built on its top with Flash Translation Layer (FTL) protocol and NAND Flash Translation Layer protocol. These layers can be implemented either by software on a host or firmware on the flash controller. Refer to Figure 2 for the NAND flash-based file systems.

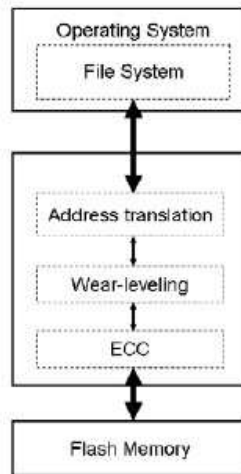


Figure 2: Wear leveling in NAND flash-based file systems

FTL translates page-level addresses, while the NFTL performs address translations at the block level. Both mechanisms are required for address translation and garbage collection, allowing the host to access NAND flash devices in the same way as disk drives.

Address translation means mapping a logical block address (LBA) to the physical block address (PBA) by the translation table so that the host performs its reads and writes to LBAs only. Garbage collection is to reclaim pages of invalid data by copying valid pages into a free area and erasing their residing blocks.

The Wear Leveling Algorithm

Dynamic Wear Leveling

Wear leveling allows erase counts of blocks to be evenly distributed over the storage media so that write/erase endurance for the entire flash drive could be increased. Dynamic wear leveling is an algorithm by which the controller in the SSD recycles blocks with small erase counts in the flash array. Whenever a sector is written, the entire block must be moved since erase operations are at a block level. A spare block is identified in the spare sector pool; the old data is copied and appended with the new sector to be written. The logical to physical sector mapping is updated and the old block is added to spare sector pool. If an error is detected during program or erases operation, the entire block is marked as bad on the bad block table and a spare block is used

instead. Although dynamic wear leveling has great improvement on wear leveling, the endurance increase is constrained by its nature. Updates and recycling of blocks and pages only happen to those blocks that are free or occupied by frequently updated data.

Global Wear Leveling

The objective of global wear leveling is to prevent any frequently updated data from staying at the static area so that wear leveling could be evenly applied to all blocks. Static areas contain any data that does not change, and are ignored by dynamic wear leveling. Such static data may include operating system files, table look-ups, executable files, and etc. Global wear leveling frequently replaces blocks in this area with block in the hot area, and thus each block in all areas has the same probability to be used.

Conclusion

Apacer's wear-leveling scheme for NAND flash SSDs is achieved via buffer management and Apacer-specific global wear leveling. They both ensure that the lifetime of the flash media can be increased, and the drive performance is optimized as well.

Use of wear leveling together with the error-correction code can greatly improve the reliability and significantly prolong the overall lifespan of the storage media. This delivers industrial-grade quality of Apacer solid-state drive for our customers.

Revision History

| Revision | Date | Description | Remark |
|----------|-----------------|-----------------|--------|
| 1.0 | May 20, 2008 | Initial Release | |
| 1.1 | January 8, 2009 | Wording refined | |

Apacer Technology Inc.

9/F, No. 100, Hsin Tai Wu Rd.
Hsichih, Taipei County 221, Taiwan
Tel: +886-2-2696-1666 Fax: +886-2-2696-1668
www.apacer.com

Copyright © 2009 Apacer Technology Inc. All Rights Reserved.
Information in this document is subject to change without prior notice.
Apacer and the Apacer logo are trademarks or registered trademarks of Apacer Technology Inc.
Other brands, names, trademarks or registered trademarks may be claimed as the property of
their respective owners.