Chapter 12
Wear Leveling for PCM Using Hot Data Identification

Inhwan Choi and Dongkun Shin

Abstract Phase change memory (PCM) is the best candidate device among next generation random access memory technologies. PCM has a potential to replace Flash memory due to non-volatility and in-place programmability, and low power consumption. Even though lifetime of PCM is longer than flash memory, wear leveling is needed because of non-uniformity of storage workload or malicious attack. In this paper, we propose a novel wear leveling algorithm for PCM as storage. Proposed algorithm extended the lifetime maximum 16 times and average 14 times in comparison to Segment Swapping algorithm.

Keywords Phase change memory · Wear-leveling · Nonvolatile memory

12.1 Introduction

Mobile consumer devices such as mobile phones and mobile pads require instant availability and low power consumption. However, these two requirements conflict with each other. For instant availability, main memory must preserve data even when the device is not used, thus a large power is consumed to refresh DRAM. For low power consumption, it is better to suspend or shut down the device after the data in main memory is saved in flash memory. However, user must bear a long data restoration time or boot time.

To satisfy these conflicting requirements, phase change memory (PCM) can be a solution. PCM has emerged as a next generation memory device to replace

I. Choi (✉) · D. Shin
School of Information and Communication Engineering,
Sungkyunkwan University, Suwon, 440-746, Korea

K. J. Kim and S. J. Ahn (eds.), *Proceedings of the International Conference on IT Convergence and Security 2011*, Lecture Notes in Electrical Engineering 120,
DOI: 10.1007/978-94-007-2911-7_12,

conventional memory device. PCM has several advantages over DRAM and flash memory. Compared to DRAM, PCM can preserve data without power supply and consumes less power [1]. Compared to NAND flash memory, PCM has shorter read and write latencies and PCM can be overwritten by the unit of a byte while the write operation is performed by the unit of a page in flash memory and the page cannot be overwritten without erasing the corresponding flash memory block. Therefore, PCM is currently expected to be an alternative or auxiliary storage device for DRAM and flash memory.

If PCM is used for storage, the device can read the PCM saved data or the boot image fast. Therefore, PCM allows implementing a mobile device with instant availability and low power consumption.

PCM, however, does have disadvantages. PCM cells endure only a limited number of writes, typically between 10^7 and 10^8. Although PCM is more durable than flash memory, it still does not have enough endurance to be used in main memory. Therefore, a hybrid memory architecture, where a small DRAM cache is backed by a larger capacity PCM memory, is more proper to adopt PCM as a main memory. In addition, we need a wear-leveling technique, which tries to make the writes uniform by remapping heavily written pages to less frequently written pages to improve the lifetime of a PCM system.

Considering the longer latencies of read/write operations on PCM than those of DRAM, it can be better to employ PCM as a storage device rather than a memory device for high performance systems. There were several studies on the exploitation of PCM as storage, where PCM was used as data storage, metadata storage, or a write buffer for NAND flash memory storage. Although the endurance of a PCM cell is higher than that of NAND flash memory, it is also necessary to make an effort to enhance lifetime of PCM since the storage workloads have no uniform write pattern.

Moreover, when PCM is used for a write buffer of flash memory or metadata storage, PCM requires much more endurance than flash memory. A PCM cell can also be worn out sooner than expected by malicious attack. Therefore, a wear-leveling is an indispensable function of PCM file systems.

The wear-leveling techniques have widely been used in NAND flash memory devices, since they have a limited endurance. Generally, the flash translation layer (FTL) performs the wear leveling during the address remapping, which remaps a logical page address into a physical page address to handle the erase-before-write characteristic of flash memory. The wear-leveling uses a table to track write counts associated with each flash block and an indirection table to perform address mapping to achieve uniform wear-out of the flash storage system.

There are inactive and proactive wear-leveling techniques. The inactive wear-leveling tries to balance the program/erase (P/E) counts of flash memory blocks by allocating the block with the minimum program/erase (P/E) count to service write requests. The proactive wear-leveling swaps the physical blocks of hot data and cold data when the P/E count difference between the blocks becomes larger than the threshold value. The basic unit of wear-leveling in FTL is a block because the erase unit is a block.

Generally, FTL can cause additional amount of writes over host requests due to the garbage collection and the proactive wear-leveling. That is called write amplification. The write amplification ratio (WAR) can be formulated as follows:

WAR = data amount written storage/data amount written by host

In this paper, we propose a novel wear-leveling technique called differentiated space allocation (DSA), which can balance the write counts of PCM pages invoking a low WAR. PCM does not require the garbage collection, since PCM cells can be overwritten. Therefore, only the wear leveling technique affects on WAR. To balance the write counts of PCM pages, when a logical page is frequently updated by host system, DSA allocates more physical pages to prevent further increment on the write count of the allocated PCM pages. Therefore, Overall Architecture of DSA Wear Leveling all logical pages have a different number of allocated physical pages depending on the update frequency. Experimental results showed that the proposed technique improved both the lifetime and write amplification ratio of PCM compared with the previous techniques.

12.2 Proposed Algorithm

The main unit of DSA wear-leveling technique is a segment. Therefore, the mapping table translating a LSN into a PSN is required. To reduce the required memory space for the mapping table, we can use a large size of segment. However, for hot segments, DSA manages the write counts of each chunk within segment to mitigate the imbalance within a segment. DSA prevents from increasing the write count of a chunk beyond the threshold value θ by remapping the corresponding logical chunk to another physical chunk if the original physical chunk is updated more than θ.

Figure 12.1 shows the overall architecture of DSA wear leveling. We assumed a segment consists of four chunks. Each chunk is represented by the pair of (segment number, chunk offset). Each LSN is mapped to a PSN. DSA manages chunk-level write counts only for recently-used segments. Therefore, the required memory space is not large. In addition, DSA does not maintain the write counts of all segments whereas the segment swapping manages write counts of all segments. If the write count of a chunk of a recently-used segment exceeds θ, a new physical chunk from the reserved segment pool (RSP) is allocated for the corresponding logical chunk. The RSP is an overprovision area, the capacity of which is hidden to host.

For example, in Fig. 12.1, the LSN 1 is mapped to the PSN 23 and the logical chunk (1, 2) is a hot chunk with the write count of 150. When the write count (N_{PW}) of the physical chunk (23, 2) becomes same to the value of θ, 100, DSA allocates the physical chunk (100, 2) for the logical chunk (1, 2). The hot chunk mapping is written at the hot chunk mapping table. If N_{PW} of the additionally allocated chunk becomes θ, the chunk becomes an expired chunk and another chunk is allocated from the RSP. In Fig. 12.1, four physical chunks are exhausted by the logical chunk (0, 0) and the physical chunk (120, 0) is finally

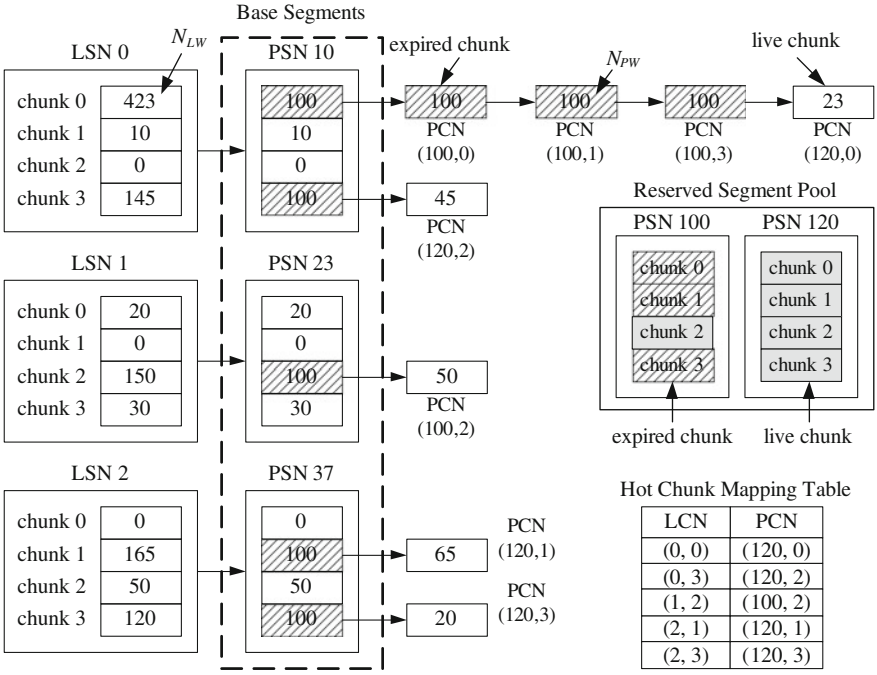

Fig. 12.1 Overall architecture of DSA wear leveling

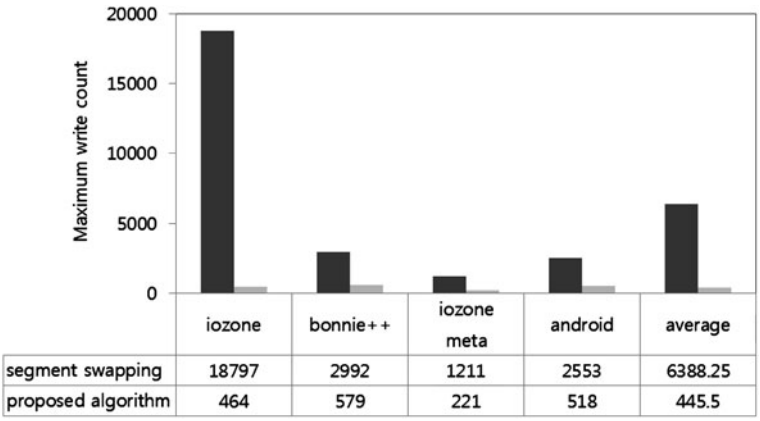

Fig. 12.2 Best results of segment swapping vs. proposed algorithm

has the up-to-date value. The physical chunks are prevented from being written over the value of θ thus the difference between the write counts of physical chunks is smaller than θ.

12.3 Experiments

In order to estimate the performance of DSA algorithm and compare with the segment swapping technique [2], we implemented a PCM simulator, which counts the number of update operations.

Figure 12.2 compares the maximum write counts of two wear-leveling techniques under the best configurations of each technique. The DSA technique outperforms 5 times at minimum and 14 times on average compared with the segment swapping technique.

12.4 Conclusions

In this paper, we proposed a novel wear leveling algorithm for PCM as storage. Because proposed algorithm performs analyzing write pattern and dispersing selected hot region, effective wear leveling was possible with minimum amount of write amplification. The results showed that proposed algorithm make lifetime of PCM average 14 times longer than Segment Swapping.

References

1. International Technology Roadmap for Semiconductors (ITRS) (2009) [Online]. Available: http://public.itrs/net
2. Zhou P, Zhao B, Yang J, Zhang Y (2009) A durable and energy efficient main memory using phase change memory technology. In: Proceedings international symposium on computer architecture (ISCA), pp 14–23, June 2009