Analysis of Performance for NAND Flash Based SSDs via Using Host Semantic Information

Jaeho KIM\(^{(a)}\), Member and Jung Kyu PARK\(^{(b)}\), Nonmember

**SUMMARY** The use of flash memory based storage devices is rapidly increasing, and user demands for high performance are also constantly increasing. The performance of the flash storage device is greatly influenced by cleaning operations of Flash Translation Layer (FTL). Various studies have been conducted to lower the cost of cleaning operations. However, there are limits to achieve sufficient performance improvement of flash storages without help of a host system, with only limited information in storage devices. Recently, SCSI, eMMC, and UFS standards provide an interface for sending semantic information from a host system to a storage device. In this paper, we analyze effects of semantic information on performance and lifetime of flash storage devices. We evaluate performance and lifetime improvement through SA-FTL (Semantic Aware Flash Translation Layer), which can take advantage of semantic information in storage devices. Experiments show that SA-FTL improves performance and lifetime of flash based storages by up to 30 and 35%, respectively, compared to a simple page-level FTL.

**key words:** NAND flash memory, Flash Translation Layer (FTL), storage system, performance

1. Introduction

Due to the recent rapid increase in the amount of data, the use of NAND flash memory based storage devices is rapidly spreading from mobile devices to enterprise server storage. Flash storage devices have various advantages such as low power consumption, fast data access, and light weight, but users’ demand for high performance continues.

Flash memory is managed by a software layer called Flash Translation Layer (FTL) in the flash memory controller since it has out-place update property. The main role of FTL is address translation to handle I/O requests from a host system and cleaning operations to reuse invalidated pages. Performance of the flash storage device is greatly influenced by the efficiency of cleaning operations. In order to minimize cleaning cost, many researches such as hot/cold data classification policies [1], [2] and buffer management policy [3] through pattern analysis of I/O requests have been conducted. However, it is difficult to obtain sufficient performance improvement of storage devices without help of the host system, with only limited information in the storage device.

Recently, SCSI, eMMC 4.5, and Universal Flash Storage (UFS 2.0) standards provide semantic information such as SCSI Group Number, Stream ID, Context ID and Data Tag to the interface between a host system and a storage device [4]–[7]. We can expect to improve performance of flash storage devices by using such information. However, studies of the effect of semantic information are insufficient and the effect is not verified. In this paper, we analyze the effect of semantic information of flash storage devices in terms of performance and lifetime. We evaluate performance and lifetime improvement through SA-FTL (Semantic Aware-Flash Translation Layer), which can take advantage of semantic information. Experiments show that SA-FTL improves performance and lifetime by up to 30 and 35%, respectively, compared to a simple page-level FTL.

The rest of the paper is organized as follows. In the next section, we review the flash based storage and present work related to this work. In Sect. 3, we introduce SA-FTL which can utilize semantic information. In Sect. 4, we present evaluation results focussing on performance and lifetime of flash based storage. Finally, we give a summary and conclude in Sect. 5.

2. Background and Related Work

In this section, we first describe internal structure of modern flash based SSDs. Then, we briefly discuss related work.

Today’s SSDs are composed of many NAND flash memory chips connected through channels and ways [8]. Such configuration allows SSDs to achieve high performance through parallel access of those resources. NAND flash memory chips consists of multiple dies and planes, which has multiple blocks, and each block has multiple pages [9]. The basic operations on flash memory are the read and write operations, and these are done in page units.

A unique characteristic of flash memory is that data cannot be overwritten on a used page. In order to overwrite a page, the block containing the page has to be erased first. This erase operation is another order of magnitude slower than a page write operation*. Furthermore, the number of erasures after writing, generally termed the Program/Erasure (P/E) cycle, is limited depending on the manufacturing technology. Today, three types of technologies, namely, SLC, MLC, and TLC, are in wide use and their P/E

\*Note that in flash memory terminology, the write of a page is also referred to as being programmed. Hence, we use the two terms interchangeably.
cycles range roughly in the 100 thousand, 10 thousand, and 1 thousand range, respectively, but also depend heavily on the manufacturers [10].

Clean pages will eventually run out as overwrites are not possible. To rid invalid pages, which hold old data that were logically overwritten, and turn them back to clean pages, a process called garbage collection (GC) is performed. (We use garbage collection and cleaning interchangeably in this paper.) As erases can happen only in block units, GC starts by selecting the block to be erased (victim selection). However, this victim block may possibly hold a mix of valid and invalid pages, so before the erase operation is performed on this block, valid pages in this block must first be moved to unused pages in other blocks. This moving of valid pages is called write amplification (WA) and is a source of overhead. For the cleaning, FTLs reserve some blocks as over-provisioned blocks (OPS) to handle write requests and help make GC efficient. Therefore, flash storage devices are generally divided into data area and OPS area.

Numerous studies have been conducted to improve the performance through reducing the GC cost of flash based SSDs. A number of studies have been conducted at the FTL level. Various FTL schemes improve random write performance by mapping logical addresses to physical addresses [11]–[13]. In addition, performance improvement through hot and cold data classification is also a typical performance improvement technique [13], [14]. Various policies for hot and cold data separation have also been proposed [1], [2]. Meanwhile, Kim et al. propose a buffer management scheme called BPLRU to mitigate GC cost induced by random writes [3]. However, these efforts merely use information inside SSDs that is insufficient for improving performance. To this end, the interface standards between the host and the storage device have recently been defined [5]–[7], [15]. Recently, Multi-stream SSD [16] has been proposed which shows performance improvement by sending host’s hint to SSDs. Multi-stream SSD [16] reduces GC overhead by maintaining multiple streams in an SSD according to the expected lifetime of the data. A similar study has also been conducted to provide differentiated services by sending requests from hosts to storage devices along with classification tags [4]. In contrast, we focus analyzing the effect of semantic information of flash storage devices in terms of performance and lifetime in this paper.

3. SA-FTL

In this section, we describe Semantic Aware-Flash Translation Layer (SA-FTL), which exploits semantic information received from host system in SSDs. First of all, we note that classification between metadata and user data managed by host file system are defined as the semantic information of host system. The types of metadata managed by the file system include superblock, group descriptors, inodes, journals, etc. Access area of logical block address (LBA) for the metadata forms particular bands. This means that the metadata has strong spatial locality resulting in reducing the cost of GC in flash storage. We show this in Sect. 4.

SA-FTL follows the structure of a general page mapping FTL and divides both data and OPS space into user and meta areas. Figure 1 shows overall layout of an SSD with SA-FTL. The size of the user and meta areas in the both data and OPS spaces is dynamically adjusted according to the amount of data from the host. The total amount of user and meta data is shown in Sect. 4. The location where each write request will be written is determined through the semantic information. In the cleaning operation, even when the valid data is copied, the data of the two separated areas are not mixed. An important point of this scheme is that there is no mix of user and meta data.

4. Evaluation

In this section, we firstly explain our experimental environment. Then, we discuss evaluation results.

4.1 Environment of Evaluation

This section quantitatively analyzes the effect of the semantic information on the performance of flash storage devices through SA-FTL. For comparison, we use Page-FTL, a general page mapping FTL that does not receive any semantic information from a host system. We measure and analyze I/O response time, number of P/E cycles, and ratio of valid pages moved per block on cleaning operations. To evaluate SA-FTL and Page-FTL, we used an SSD simulator (DiskSim with SSD Extension), which consists of 8 chips and 4096 blocks per chip [17]. Table 1 shows the parameters of the SSD simulator. We used the Fileserver and Varmail workloads of Filebench and also used Postmark benchmark. Ext4 file system is used on host system and the semantic information is generated by tagging ext4 metadata and user data classification in I/O requests. Table 2 shows the characteristics of the I/O workload. Figure 2 shows the access pattern of metadata and user data requests, showing sector number of each request over time. As shown in Fig. 2, the access patterns of metadata are gathered in specific sector numbers, while the access patterns of user data are distributed in a much wider area than the metadata. It means that metadata has higher spatial locality than the user data. In particular, in case of Postmark shown in Fig. 2 (c), we see that I/O access area of user data is distributed across almost
all of the sectors.

4.2 Results of Evaluation

Figure 3 (a) shows the average response time of I/O requests for all workload according to OPS (Over Provisioned Space) size. OPS is reserved space to process cleaning operations in flash storage device. The smaller the size of the ops, the more the number of cleaning operations, which results in a performance degradation. In case of OPS 10% in Fig. 3 (a), the performance difference between the two FTLs is almost zero. On the other hand, for OPS 5% and 3% in Fig. 3 (a), the performance improvement of the SA-FTL increases as the OPS space becomes smaller. In case of OPS 3%, the performance of SA-FTL on Fileserver, Varmail, and Postmark workloads is improved by 10%, 20%, and 30%, respectively. Since SA-FTL separates metadata with high spatial locality from user data, cost of cleaning operations is reduced. As the OPS becomes smaller, the number of cleaning occurrences increases. Therefore, the performance improvement of SA-FTL is greatest in the smallest OPS.

Figure 3 (b) shows the number of P/E cycles that occurred during execution of the benchmarks. The number of P/E cycles indicates number of block erases during cleaning operations. Therefore, the smaller number of P/E cycles, the better efficiency of cleaning operations. The results of P/E cycles shown in Fig. 3 (b) also show that SA-FTL performs better than the existing page-FTL. For the 3% OPS in Fig. 3 (b), the number of P/E decreases by 21%, 28% and 35% in Fileserver, Varmail and Postmark workloads, respectively.

Particularly, the largest performance improvement on Postmark workloads shown in both Fig. 3 (a) and 3 (b) is due to the fact that the sector access range of user data is much wider area than the metadata. The results show that performance improvement can be obtained by using the semantic information to distinguish metadata which has high spatial locality from user data.

4.3 Effect of Ratio of Valid Page

Figure 4 shows average valid page ratio of victim blocks (hereafter referred to as ‘u’) with respect to the OPS size
The performance of flash storage devices is greatly influenced by cleaning operations, a write request is amplified to n writes. Generally, degree of amplification is defined as WAF (Write Amplification Factor) \([18]\), which depends on \(u\) as in the following Eq. 1 \([19]\).

\[
WAF(u) = \frac{u \cdot N_p}{(1-u) \cdot N_p}
\]

5. Conclusion

The performance of flash storage devices is greatly influenced by cleaning efficiency of FTLs. Many existing studies have attempted to improve performance by using only limited information in flash storage devices without help of the host system, but such methods have limitations for sufficient improvements. Recently, SCSI, eMMC, and UFS standards provide semantic information of host system, and it is expected to improve performance of flash storage device by using it. Therefore, we present and analyze the effect of the semantic information on performance and lifetime of flash storage in this paper. We show that performance and lifetime improvements through SA-FTL, which can utilize the semantic information from the host.

References


