1. Introduction

Solid-state drives (SSDs) are replacing the hard disk drives (HDDs) as the primary storage due to the advantages in speed, reliability and power. In the conventional database application, SSDs are used as the cache for HDDs to balance the performance benefits and costs. Since the bit-cost of NAND flash continuously decreases, the SSDs are proposed as the storage for the database. Since the storage engine (SE), a software component in the database management system (DBMS), controls the data storage, the data information inside the SE is utilized to improve the SSD write performance, which is relatively slow due to the inherent characteristics of the NAND flash memories.

2. SSD Based DBMS

The garbage collection (GC) overhead of the SSD is the write performance bottleneck of the SSD [1]. One approach to reduce the GC overhead is clustering data with similar activities in the same block of the NAND flash. To realize it, upstream information from the SE is utilized. As illustrated in Fig. 1, each layer of the storage stacks of the operating system (OS) has to be modified to pass the hint information from SE to the flash translation layer (FTL) in the SSD controller, conventionally [2]. It introduces additional overhead and engineering efforts. Hence, the OS is bypassed in the proposed SSD-based DBMS. Hint information passes from the SE to the FTL directly.

In this paper, a storage engine assisted solid-state drive (SEA-SSD) has been proposed to improve the storage performance for the database application by co-optimizing the SSD controller and database storage engine. Data with different activities are predicted, classified and aggregated into the same block of the NAND flash memory. From the experimental results, maximum 24% speed boost, 16% power reduction and 19% endurance enhancement are achieved, without requiring a cache layer for the SSD.
dynamic data with the flush list. When new data enters the flush list, it indicates that this data will be flushed to storage in the near future. Hence, the logical page address (LPA) of the data is sent to the SSD controller. When GC is triggered, the data of the hinted LPAs will migrate from Segment_stat to Segment_dyn, described in Algorithm II of Fig. 3.

4. Evaluation

An application-coupled platform has been designed to evaluate the proposed scheme. In this work, the database is directly coupled with the SSD for optimizing the SSD controller. SSD is virtualized for the simulation acceleration. As mentioned earlier, the OS is bypassed in the interaction between the SSD and database. Table I describes the SSD workload settings [2]. Online transaction processing (OLTP) benchmarks, Sysbench [3] and TPC-C [4], are used. From the experimental results, maximum 24% SSD speed boost, 16% power reduction and 19% endurance enhancement are achieved at the SSD capacity of 3 GB (33% over-provisioning) [2]. As illustrated in Fig. 4, the SEA-SSD improves the SSD performance by 7%-24% at a reasonable over-provisioning range of less than 50%.

5. Conclusion

A SE assisted SSD has been proposed to improve the SSD performance for the database application. By obtaining the hint messages from the upstream layer, specifically SE of the database, the data are better classified and more efficiently stored in the SSD. Consequently, the garbage collection overhead of the SSD is reduced and maximum 24% SSD performance improvement is achieved.

Reference


Table I. Benchmark statistics for evaluation [2].

<table>
<thead>
<tr>
<th>MySQL benchmark</th>
<th>SSD workload</th>
<th>Buffer pool size (MB)</th>
<th>Redo log size (MB)</th>
<th>Write size (GB)</th>
<th>Read size (GB)</th>
<th>Write request ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPC-C</td>
<td>T1</td>
<td>256</td>
<td>4</td>
<td>25.1</td>
<td>41.7</td>
<td>37.5%</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>512</td>
<td>8</td>
<td>21.9</td>
<td>18.7</td>
<td>58.2%</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>1024</td>
<td>16</td>
<td>19.3</td>
<td>4.07</td>
<td>82.6%</td>
</tr>
<tr>
<td>Sysbench</td>
<td>S1</td>
<td>256</td>
<td>4</td>
<td>11.1</td>
<td>34.1</td>
<td>24.5%</td>
</tr>
<tr>
<td></td>
<td>S2</td>
<td>512</td>
<td>8</td>
<td>8.93</td>
<td>14.1</td>
<td>38.8%</td>
</tr>
<tr>
<td></td>
<td>S3</td>
<td>1024</td>
<td>16</td>
<td>7.94</td>
<td>0.97</td>
<td>89.1%</td>
</tr>
</tbody>
</table>

Fig. 2. The overview of the SEA-SSD [2].

Fig. 3. SSD management algorithm flow chart [2].

Fig. 4. SEA-SSD write performance evaluation [2].

24% SSD speed boost, 16% power reduction and 19% endurance enhancement are achieved at the SSD capacity of 3 GB (33% over-provisioning) [2].