Black-Box Performance Modeling for Solid-State Drives

Shan Li
Department of Electrical and Computer Engineering
The George Washington University
Email: shallia@gwu.edu

H. Howie Huang
Department of Electrical and Computer Engineering
The George Washington University
Email: howie@gwu.edu

***Abstract*—Flash-based Solid-State Drives (SSDs) have become a promising alternative to magnetic Hard Disk Drives (HDDs) thanks to the large improvements in performance, power consumption, and shock resistance. An accurate SSD performance model will provide the important research tools for exploring the design space of flash-based storage systems. While many HDD performance models have been developed, architectural differences prevent these models from being effective for SSDs, mostly because their designs cannot accurately account for many unique SSD characteristics (e.g., low latencies, slow updates, and expensive erases). In this paper, we utilize the black-box modeling technique to analyze and evaluate SSD performance, including latency, bandwidth, and throughput. Such an approach is appealing because it requires minimal *a priori* information about SSDs. We construct and evaluate our models on three commercial SSDs. Although this approach may lead to less accurate predictions for HDDs, we find that a black-box model with a comprehensive set of workload characteristics can achieve the mean relative errors of 20%, 13%, and 6% for latency, bandwidth, and throughput predictions, respectively.**

I. INTRODUCTION

Hard disk drives (HDDs) have been the choice of secondary storage in computer systems for half a century. Unfortunately, the I/O bottleneck remains, because disk access times lag behind CPU and DRAM performance. The situation is poised to change with recent advances of NAND flash-based solid-state drives (SSDs). SSDs present an exciting opportunity to close the I/O performance gap, because they sit in a sweet spot in the middle, with densities comparable to HDDs and I/O performance closer to DRAM. For hard disks and high-performance storage systems, researchers and engineers have utilized performance modeling techniques (e.g., analytical modeling, simulation, and benchmarking) to analyze, develop, and evaluate new algorithms and systems. Likewise for SSDs, these models will help give a good understanding of the state-of-the-art and provide the research tools for exploring the design space for SSDs and developing new storage systems that can better utilize them.

In this paper, we propose to leverage the black-box approach to model the performance of SSDs, including latency, bandwidth, and throughput. A black-box model can be constructed in two steps: 1) benchmark a storage device and collect the training data that consist of the model inputs (workloads characteristics) and outputs (performance metrics); and 2) utilize the statistical methods to quantify the correlations between the inputs and outputs. The rational is that the performance tends to be highly correlated with the workload characteristics. For example, SSD latency and throughput fluctuate when the percentage of write requests, the number of random requests, and the outstanding I/O requests vary. This approach is attractive as it requires no or limited *a priori* information about a storage device. This is very beneficial in the case of SSDs, because for intellectual property reasons, the SSD vendors are reluctant to reveal the internal designs on wear-leveling, caching, queueing, etc. Prior research showed that the black-box modeling can give a reasonable performance prediction for hard disks and arrays [1], [2]. However, accurate performance prediction for a hard disk is challenging, especially for latency [3]. This is not unexpected as a combined result of hardware (moving heads and rotating plates) and software components (e.g., caching and scheduling algorithms) inside a hard disk. In order to develop a good performance model for SSDs, one needs to recognize that vast architectural differences exist between two device families [4]. For example, slow random writes can happen in a solid-state drive due to expensive erases at the block level.

To this end, we start with a basic black-box model with traditional workload characteristics, e.g., read/write ratio and request size. Intuitively, one can speculate that the difficulty of the latency predication will be less severe for flash-based SSDs because there is no longer moving mechanical parts. Nevertheless, a good black-box performance model for an SSD should be able to accommodate the effects that are caused by the unique architectural features of SSDs, e.g., page-level reads and writes, out-of-place updates, and block-level erase operations. Therefore, we take into account these characteristics, specifically, we split the request size into read and write sizes (because of SSD asymmetric read/write performance), and add the access stride (for the effect of request alignments). In our black-box approach, we first collect a large amount of the training data on the workload characteristics and device performance. Our approach then applies the statistical machine learning algorithms for the model fitting, where the performance is predicated as a function of the workload characteristics. We evaluate the models using a wide variety of micro-benchmark workloads on three real-world SSDs. The results are promising - the mean relative errors of an SSD model are as low as 20% for the latency prediction, 13% for bandwidth, and 6% for throughput.

II. Black-Box Performance Modeling for SSDs

In this paper, we begin with a basic black-box model that has been used for hard drives, and extend it to include additional workload characteristics for SSDs. We utilize the statistical machine learning algorithms to capture the correlations between the workload characteristics and observed performance values.

Basic Model: A typical black-box model predicts the performance as a function (*F*) of various workload characteristics. This model takes workload characteristics (*wc*) as input parameters and predicts a performance metric (*p*), which can be formally written as in Eq. 1:

$$p = F(wc). \tag{1}$$

The workload is defined as a stream of I/O requests. Typically when modeling a hard drive, the workload can be characterized by read and write ratio (*rw_ratio*) that is defined as the percentage of writes in the request, the request size (*req_size*) that represents the number of bytes transferred to/from the storage device, the number of outstanding I/Os that represents queue depth (*q_dep*), and request randomness (*rand*) that is defined as the percentage of random reads and writes in the I/O request stream. Thus, a single workload *wc* can be represented as a vector of workload characteristics, shown in Eq. 2:

$$wc =< rw\_ratio, q\_dep, req\_size, rand > . \tag{2}$$

In this paper, we focus on three performance metrics: latency (*lat*), bandwidth (*bw*), and throughput in IO per second (*iops*). Thus, the performance *p* can be represented as either of three metrics shown in Eq. 3:

$$p =< lat, bw, iops > . \tag{3}$$

Extended Model: We conduct a number of experiments on various devices in order to examine how different workload characteristics affect the SSD performance. The results show that, for SSDs, the write ratio has a large impact on the performance. In addition, while the size of a write request greatly influences the latency for the SSDs, the read size has a less noticeable impact. Thus, we divide the request size into the read size (*rd_size*) and the write size (*wr_size*), because SSD reads and writes have asymmetric performances. Furthermore, we study different access patterns, including sequential, random, and stride (write stride, *wr_stride*, and read stride, *rd_stride*). We define that an I/O request is a stride access pattern when there exists a fixed offset between the consecutive accesses. The experiments demonstrate that all the devices present different performance behaviors under sequential and random access patterns. In the same time, while the hard disk has the similar stride performance as random access, it turns out that the stride access is very challenging for SSDs. These observations inspire us to compose a model with an extended set of the workload characteristics, which eventually produce more accurate predictions.

In the extended model, a workload *wc* can be written as a vector of workload characteristics shown in Eq. 4:

$$\begin{aligned} wc =< rw\_ratio, q\_dep, wr\_size, rd\_size, \\ wr\_rand, rd\_rand, wr\_stride, rd\_stride > . \end{aligned} \tag{4}$$

Regression Tree: To construct a black-box model, we first need to collect the training data that consist of the workloads characteristics and the corresponding performance of a storage device. In our approach, given the training data as the input, the regression algorithm is applied to calculate a prediction function that maps the workload characteristics (*independent variables*) to the performance metrics (*dependent variables*). Specifically, we construct a tree from the regression function, which is generated by recursively splitting the independent variables into the leaf nodes. A leaf node of the tree provides a prediction of the dependent variables as a function of the independent variables. In this research, we employ the least-square multilinear regression to build our performance model.

III. Evaluation

The experiments are run on the machines with an Intel Atom dual-core 1.6GHz processor, 2GB memory, and Linux kernel 2.6. The training data are collected from four different storage devices: three SSDs including OCZ-Apex (SSD_A), Intel X-25M (SSD_I), Samsung (SSD_S), and one hard drive, Samsung Spinpoint M7 (HDD_S). We use Intel Open Storage Toolkit (OST) [5] as the I/O workload generator. The values for the workload characteristics are selected as follows. The write ratio is from 0%, 25%, 50%, 75%, to 100%, where 0% means read only and 100% write only. The read/write size is from 1KB to 256KB (times of 4), the queue depth from 1 to 64 (times of 4), the randomness for read and write from 0%, 50%, to 100%, and the stride size in the range of 1KB to 256KB. The workloads include three types of access patterns, sequential, random, and stride. Each I/O request is run for one minute. Three performance metrics, latency in millisecond, bandwidth in MB/s, and throughput in IO/s, are measured. To construct a model, we randomly select 2,000 data points out of 12,000 requests and generate a linear regression tree using an open source statistical software GUIDE [6]. In this paper, we evaluate the models with Mean Relative Error (MRE), and R^2 that is a statistical measurement on how well the regression line approximates the real data points. A R^2 value of 1 indicates that the regression line perfectly fits the observed data.

Basic Model: Table I lists the values of R^2 and MRE for the basic models of the four devices. In this model, for all four devices, the bandwidth remains the most difficult to model, e.g., 0.08 for R^2 and 96% for MRE, while both latency and throughput have mean relative errors of smaller than 70% in most cases. The main reason for the high MREs is that four inputs in the basic black-box model cannot completely capture

(a) Latency

Device	R^2	MRE
SSD_A	0.974	12%
SSD_I	0.466	72%
SSD_S	0.775	48%
HDD_S	0.848	55%

(b) Bandwidth

Device	R^2	MRE
SSD_A	0.619	76%
SSD_I	0.394	79%
SSD_S	0.536	77%
HDD_S	0.335	96%

(c) Throughput

Device	R^2	MRE
SSD_A	0.936	12%
SSD_I	0.435	72%
SSD_S	0.499	41%
HDD_S	0.075	36%

TABLE I
PREDICTION ACCURACY OF THE BASIC MODELS

the workload to produce an accurate model. In this model, performance predictions of SSD_I have more than 70% MRE values for latency and throughput, which are even worse than HDD_S.

Extended Model: Using the workload characteristics defined in Eq. 4, we construct the extended models for four disk devices, shown in Table II. In the new models, the MRE values have been reduced, that is, 20% to 29% for latency, 13% to 30% for bandwidth, and 6% to 20% for throughput. In particular, SSD_A has the smallest mean relative errors - 6% for throughput, 13% for bandwidth, and 20% for latency. In summary, compared to the basic models, the extended models for four devices have about 50% improvements in the performance accuracy.

(a) Latency

Device	R^2	MRE
SSD_A	0.979	20%
SSD_I	0.923	21%
SSD_S	0.939	29%
HDD_S	0.941	21%

(b) Bandwidth

Device	R^2	MRE
SSD_A	0.961	13%
SSD_I	0.962	16%
SSD_S	0.873	30%
HDD_S	0.805	29%

(c) Throughput

Device	R^2	MRE
SSD_A	0.967	6%
SSD_I	0.929	14%
SSD_S	0.964	20%
HDD_S	0.720	17%

TABLE II
PREDICTION ACCURACY OF THE EXTENDED MODELS

IV. Related Work

Extensive research has been done for hard disks in analytical modeling [7], [8], simulation [9], and benchmarking [10]. The analytic models and simulators are constructed in a white-box manner by developing an understanding of the internal organization of hard disks. Prior work [1], [2] proposed various black-box performance models for hard drives with minimal knowledge of device specific properties, and [3] designed a relative fitness model to predict the performance difference between two devices. In this work, we focus on severals workload characteristics (e.g., stride pattern, and write size) that are crucial to SSD performance.

V. Conclusions

An accurate performance model for solid-state drives will provide the important research tools for exploring the design space for the flash-based storage systems. In this paper, we study the black-box performance modeling for SSDs and find that a model with an extended set of workload characteristics has the low errors when applied to three commercial SSDs. As part of the future work, we plan to study several additional workload parameters. One interesting thing to consider is the ratio between the computation and I/O time. In addition, we will explore the possibility of applying our black-box models to help design and configure a storage system, preferably in an autonomic manner.

References

[1] M. Wang, K. Au, A. Ailamaki, A. Brockwell, C. Faloutsos, and G. R. Ganger, "Storage device performance prediction with cart models," in *MASCOTS '04: Proceedings of the The IEEE Computer Society's 12th Annual International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems*. Washington, DC, USA: IEEE Computer Society, 2004, pp. 588–595.

[2] L. Yin, S. Uttamchandani, and R. Katz, "An empirical exploration of black-box performance models for storage systems," in *14th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, 2006, pp. 433–440.

[3] M. Mesnier, M. Wachs, R. Sambasivan, A. Zheng, and G. Ganger, "Modeling the relative fitness of storage," *ACM SIGMETRICS Performance Evaluation Review*, vol. 35, no. 1, p. 48, 2007.

[4] F. Chen, D. A. Koufaty, and X. Zhang, "Understanding intrinsic characteristics and system implications of flash memory based solid state drives," in *SIGMETRICS '09: Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems*, 2009, pp. 181–192.

[5] Intel, "Intel-iscsi open storage toolkit."

[6] W.-Y. Loh, "Guide regression tree version 7.9," 2009.

[7] C. Ruemmler and J. Wilkes, "An introduction to disk drive modeling," *Computer*, vol. 27, no. 3, pp. 17–28, 1994.

[8] E. Shriver, A. Merchant, and J. Wilkes, "An analytic behavior model for disk drives with readahead caches and request reordering," in *Proceedings of the 1998 ACM SIGMETRICS joint international conference on Measurement and modeling of computer systems*, 1998, pp. 182–191.

[9] J. Bucy, J. Schindler, S. Schlosser, and G. Ganger, "The DiskSim simulation environment version 4.0 reference manual," Technical Report CMU-PDL-08-101, Carnegie Mellon University, Tech. Rep., 2008.

[10] A. Traeger, E. Zadok, N. Joukov, and C. Wright, "A nine year study of file system and storage benchmarking," *ACM Transactions on Storage (TOS)*, vol. 4, no. 2, p. 5, 2008.