

A Learning-based Approach Towards Automated Tuning of SSD Configurations

Daixuan Li and Jian Huang
Systems Platform Research Group
University of Illinois at Urbana-Champaign

Abstract

Thanks to the mature manufacturing techniques, solid-state drives (SSDs) are highly customizable for applications today, which brings opportunities to further improve their storage performance and resource utilization. However, the SSD efficiency is usually determined by many hardware parameters, making it hard for developers to manually tune them and determine the optimal SSD configurations.

In this paper, we present an automated learning-based framework, named LearnedSSD, that utilizes both supervised and unsupervised machine learning (ML) techniques to drive the tuning of hardware configurations for SSDs. LearnedSSD automatically extracts the unique access patterns of a new workload using its block I/O traces, maps the workload to previously workloads for utilizing the learned experiences, and recommends an optimal SSD configuration based on the validated storage performance. LearnedSSD accelerates the development of new SSD devices by automating the hardware parameter configurations and reducing the manual efforts. We develop LearnedSSD with simple yet effective learning algorithms that can run efficiently on multi-core CPUs. Given a target storage workload, our evaluation shows that LearnedSSD can always deliver an optimal SSD configuration for the target workload, and this configuration will not hurt the performance of non-target workloads.

1 Introduction

Flash-based solid-state drive (SSDs) have become the backbone of modern storage infrastructures in various computing platforms, as they offer orders-of-magnitude better performance than hardware-disk drives (HDDs), while their cost is approaching to that of HDDs [8, 32, 36, 54, 55, 62, 68]. Thanks to the development of manufacturing and shrinking process technology [1], the industry has been able to rapidly produce SSD devices with different hardware configurations.

Although SSD devices are becoming highly customizable to meet the ever-increasing demands on storage performance and capacity for new applications [19, 55], identifying optimal device configurations is on the critical path of SSD development. This is because the SSD hardware configurations are usually determined by the requirements from applications and customers [20, 41], and these configurations involve many components in the storage controller, such as flash chip specifications, chip layout, block/page sizes, device

buffer sizes, and so on. In order to deliver optimal performance for applications in a new generation of storage device development, storage vendors usually use typical application workloads as their benchmarks to aid them to determine the device configurations. However, an SSD device has hundreds of parameters in its configurations, and these parameters usually have dependencies (i.e., the update of one parameter may affect other parameters), making it hard for hardware engineers to tune the device configurations and identify the optimal ones in a short time. This significantly hurts the productivity of new SSD device development [20].

Furthermore, there is an increasing demand for customized storage devices from various computing platforms and applications. This is for two reasons. First, computing platforms always wish to deploy the best-fit storage devices for their workloads, such that they can achieve the maximum performance. This is especially true for cloud platforms that require highly customized SSDs to support their cloud services, such as Database-as-a-Service [3] and web services [2]. As applications such as cloud storage services are evolving quickly, we need to revolutionize the configuration tuning procedure to shorten the lifecycle of producing new SSDs.

Second, our study shows that storage workloads can be categorized with learning algorithms, which provides the evidence that it is feasible to customize storage devices for a specific workload type (see Figure 2), especially considering the SSD manufacturing techniques become mature today. However, there is still a long-standing gap between application demands and SSD device configurations. And our community lacks a framework that can instantly transfer application demands into device configurations of SSDs.

In this paper, we develop an automated framework named LearnedSSD, which exploits both supervised and unsupervised machine learning (ML) techniques to drive the hardware configurations for new SSDs. Given a storage workload, LearnedSSD will recommend an optimal SSD configuration that delivers optimized storage performance. It leverages the linear regression techniques to expose the device configurations that have the strongest correlation to the storage performance. To present reasonable device configurations, we formulate different types of hardware parameters in the SSD, transfer them into the vectors in the ML model, and utilize learning techniques to explore the optimization space and identify the near-best options, with specified constraints

such as SSD capacity, interfaces (NVMe or SATA), and flash memory types.

To reduce the execution time of learning an optimal SSD configuration while ensuring the learning accuracy, we develop pruning algorithms to identify the most important hardware parameters in SSDs. LearnedSSD also maintains a configuration database named ConfDB that stores the learned workloads and SSD configurations. For a new workload, LearnedSSD will extract its features and compare them with the records in the ConfDB using similarity comparison networks. If LearnedSSD identifies a similar workload in its ConfDB, it will recommend the corresponding SSD configuration directly, such that we can utilize the previously learned experience. Otherwise, LearnedSSD will learn a new SSD configuration for the workload, and add them into its ConfDB for future references. As many storage workloads share similar data access patterns and can be categorized into a general type (see Figure 2), LearnedSSD can assist developers to identify the most critical parameters for a type of storage workloads, and recommend an optimal SSD configuration.

Our study with LearnedSSD leads to interesting insights. We summarize them into learning rules that can aid developers to prioritize their optimization strategies when producing new SSDs. For instance, (1) not all SSD parameters are equal, the layout arrangement of flash chips is an important factor for storage performance; (2) with different targets and configuration constraints, the tuning procedure of device configurations are different, and for each parameter, its correlation with SSD performance is also different; (3) not all parameters are sensitive to storage performance, and some of them can be configured as the same as commodity SSDs today.

To evaluate the efficiency of LearnedSSD, we implemented our proposed techniques using PyTorch [72], the *scikit-learn* tool [5], and a production-level SSD simulator MQSim [70]. We perform experiments with a variety of storage traces. Our experimental results show that LearnedSSD delivers an SSD configuration that can achieve $1.28\text{--}34.61\times$ performance improvement for a target workload, without hurting the performance of non-target workloads, compared to the configurations specified by the released commodity SSDs. We also show that LearnedSSD can learn a new configuration in 37.3 seconds, and finalize an optimal configuration in 121 iterations on average for a workload, with a multi-core server processor. Overall, we make the following contributions:

- We present the first study of SSD hardware parameters and popular storage workloads with learning in mind, and demonstrate the feasibility of applying the learning-based approach for identifying optimal SSD specifications.
- We formulate the tuning problem of SSD device configurations using a learning-based approach, and develop an

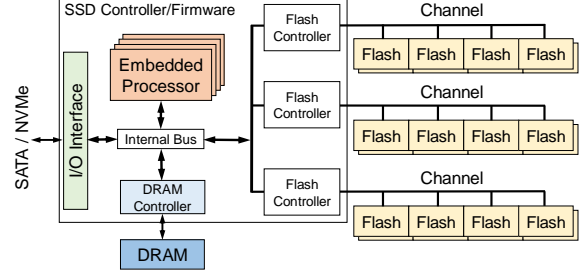


Figure 1. Internal architecture of flash-based SSDs.

automated learning framework that can efficiently recommend optimal SSD configurations for different workloads.

- We summarize a set of learning rules that can facilitate the hardware configurations and development of new SSDs, based on our study with LearnedSSD.
- We study the efficiency of LearnedSSD and show its benefits in comparison with released commodity SSD settings.

2 Background and Motivation

SSD has proven to be a revolutionary storage technology. It performs much faster than hard-disk drives (HDDs), while its price is reaching to that of HDDs. The rapidly shrinking process and manufacturing technologies have accelerated the SSD device development and enabled their widespread adoption in a variety of computing platforms, such as data centers [4, 22, 55, 68].

2.1 SSD Architecture

We present the internal system architecture of a typical SSD in Figure 1. An SSD consists of five major components: a set of flash memory packages, an SSD controller having embedded processors like ARM, off-chip DRAM (SSD DRAM), flash controllers, and the I/O interface that includes SATA and NVMe protocols [23, 31, 42]. The flash packages are organized in a hierarchical manner. Their organization not only determines the storage capacity but also affects the storage performance. Each SSD has multiple channels where each channel can receive and process read/write commands independently. Each channel is shared by multiple flash packages. Each package has multiple flash chips. Within each chip, there are multiple planes. Each plane includes multiple flash blocks, and each block consists of multiple flash pages. And the page size varies in different SSDs.

In order to manage the flash memory, the SSD controller usually implements the Flash Translation Layer (FTL) in its firmware. The FTL was developed for taking care of the intrinsic properties of SSDs. When a free flash page is written once, that page is no longer available for future writes until that page is erased. However, erase operation is performed only at a block granularity. To remove the expensive erase operations from the critical path, writes are issued to free pages that have been erased, which is also called out-of-place

update. The SSD controller will perform Garbage Collection (GC) later to clean the stale data. As each flash block has limited endurance, it is important for blocks to age uniformly (i.e., wear leveling). Modern SSD controllers employ out-of-place write, GC, and wear leveling to overcome these aforementioned shortcomings and maintain indirections for handling the address mapping in their FTL.

2.2 SSD Manufacturing and Parameter Tuning

According to the interviews with SSD product managers [21] and our discussions with SSD vendors, finalizing the SSD hardware specifications or parameters is on the critical path in the SSD design. These specifications are usually determined by the requirements from applications and customers. And they involve the components of the SSD controller as shown in Figure 1. Without these specifications, the SSD development cannot proceed to the manufacturing stage. With the confirmation from SSD vendors, there are more than a hundred tunable parameters in a typical SSD.

To finalize the SSD specifications, a straightforward approach is to test and profile application workloads with different hardware configurations. However, this is not scalable as we target different application workloads. Given an application workload, it is challenging for developers to test all the combinations of the device parameters. And likewise, give a new SSD specification, it requires significant manual effort to quantify the effectiveness of the selected specifications. In this work, we use the learning-based approach to automate the SSD hardware configurations.

2.3 Software-Defined Solid-State Drive

With the increasing demands on storage performance from applications, we have seen a trend that modern storage systems are embracing software-defined hardware techniques [4, 22]. This allows upper-level applications to achieve maximum performance benefits and resource efficiency with customized storage devices. For instance, the recent development of software-defined SSDs [19, 37] enables platform operators to customize the number of flash channels and chips in an SSD, with the cooperation with SSD vendors.

This is especially true for both public and private cloud platforms that require highly customized SSDs to support their cloud services, such as database-as-a-service [3], web services [2], web search [37], and batch data analytics [25]. For these applications, their workloads can be highly classified. For instance, we use our proposed learning-based workload characterization approach (see the detailed discussion in §3.3) to study the storage traces from a set of popular application workloads (see Table 1). Our experiments demonstrate that each workload type has its unique characteristics, and I/O traces from the same workload type have similarities in their data access pattern, as shown in Figure 2. This provides the evidence showing that it is feasible to customize storage devices for a specific category of applications.

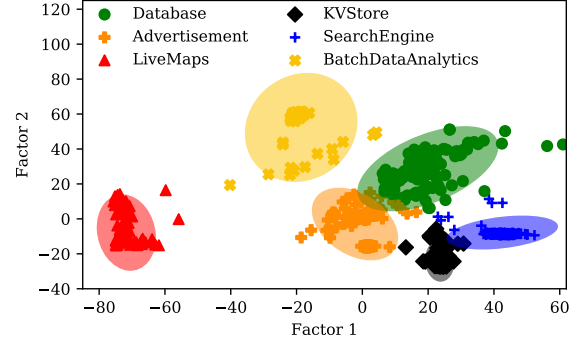


Figure 2. A clustering of popular storage workloads.

Unfortunately, our community lacks a framework that can efficiently transfer application demands and characteristics into the hardware configurations of SSDs.

2.4 Learning-Based Parameter Tuning

We have seen a disruptive advancement of machine learning (ML) techniques over the past decade. In general, we can categorize machine learning techniques into two types: supervised learning and unsupervised learning. As for supervised learning, it learns a set of rules with labeled datasets and then generalizes these rules to make predictions for new inputs. Typical example algorithms include decision trees [26], support vector machines [69], Bayesian networks [16], and artificial neural networks [14]. Unlike supervised learning, the unsupervised learning can identify unknown patterns based on unlabeled datasets, such as k-means clustering and principal component analysis (PCA) [27].

Recent studies [11, 15, 30, 43, 53, 58, 66, 73, 80] showed that the learning-based method is a promising approach to solve system optimization problems. However, none of them investigated their applications in SSD development, and it is unclear how can we utilize learning-based approaches to overcome the challenges of tuning hardware specifications of SSDs. In this work, we will use both supervised learning and unsupervised learning to develop LearnedSSD.

3 Design and Implementation

In this section, we first discuss the goals of LearnedSSD. And we will present the system architecture of LearnedSSD, and its core components respectively.

3.1 Design Goals

The high-level goal of LearnedSSD is to enable the automated tuning of hardware configurations of SSDs for a specific application workload with learning techniques. Specifically, LearnedSSD will achieve the following goals:

- It can generate an optimal SSD configuration for a target workload, while this hardware configuration has minimal negative impact on other workloads.

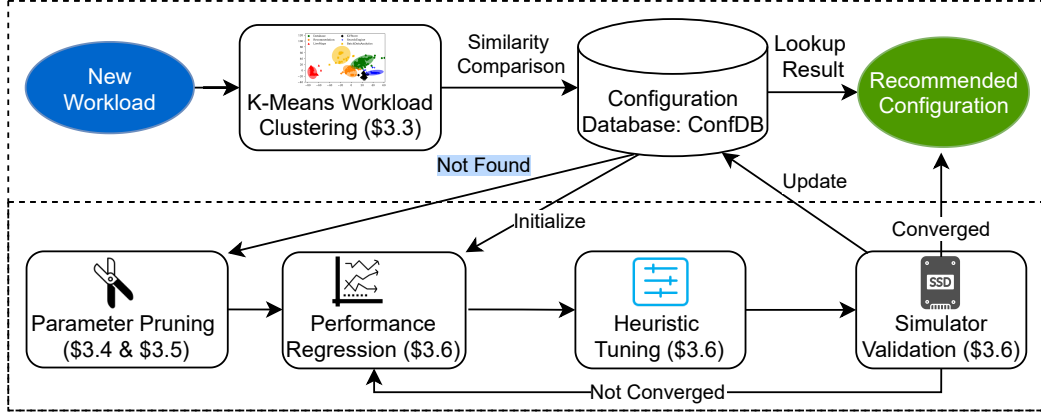


Figure 3. System overview of LearnedSSD: LearnedSSD first learns new workload features with a **learning-based workload clustering** (§3.3). If it is similar to workloads in existing clusters in the configuration database ConfDB, LearnedSSD will recommend an optimal configuration stored in the database. If not, LearnedSSD will first conduct **parameter pruning** (§3.4 and §3.5) to identify performance-critical parameters of SSDs. After that, LearnedSSD will conduct automated tuning which consists of three modules: **performance regression** with Gaussian process, **heuristic tuning**, and **performance validation** (§3.6).

- It can identify an optimal SSD configuration in a short time, without introducing much extra computation overheads.
- It can scale to support diverse target workloads as well as different device constraints.

In the following, we will discuss how LearnedSSD achieves these goals and overcomes the challenges in both systems building and configuration learning procedure.

3.2 System Overview

We develop an automated framework named LearnedSSD, which exploits both supervised and unsupervised learning techniques to drive the specifications for new SSD devices. Given a storage workload and design constraints, LearnedSSD will recommend an optimal SSD configuration that can deliver optimized storage performance shortly.

To be specific, we leverage linear regression techniques to expose the device specifications that have the strongest correlation with the storage efficiency. To present reasonable device specifications, we formulate different types of hardware parameters in the SSD, transfer them into the vectors of the learning models, and learn the optimal options with specified constraints, such as storage capacity and SSD interfaces (NVMe or SATA). To reduce the execution time of learning an optimal SSD configuration while ensuring the learning accuracy, we develop pruning algorithms to identify the most important hardware parameters in SSDs. We will also maintain a configuration database named ConfDB to store the learned workloads and the corresponding SSD specifications. Therefore, we can utilize the learned experiences for new workloads in LearnedSSD.

For a new workload, LearnedSSD will extract its features and compare them with the records in the **configuration**

database ConfDB, as shown in Figure 3. If LearnedSSD identifies similar workloads, it will either recommend a known optimal specification, or retune the SSD configurations with the recorded SSD specifications, such that we can utilize previous learned experiences. Otherwise, LearnedSSD will learn new SSD specifications and add the learned configurations into ConfDB for future references. As many storage workloads can be categorized into a general type (see Figure 2), LearnedSSD can also assist developers to identify the most critical parameters for a type of storage workloads. It is worth noting that LearnedSSD will ensure the recommended SSD specifications will not hurt the storage performance for other generic workloads.

3.3 Learning-based Workload Clustering

Unlike traditional ways of using the read/write ratio, and I/O patterns (e.g., sequential and random read/write) to categorize workloads, which cannot capture the whole picture of workload characteristics, we develop a learning-based clustering approach based on block I/O traces. We chose block I/O traces to learn the characteristics of storage workloads, because this approach does not have system dependencies and not require application semantics.

To develop the learning-based workload clustering, we **first partition each I/O trace into small windows**. According to our study of diverse workloads, we **use 3,000 trace entries in each window by default**. This is because fewer entries may lose the unique data access patterns of the trace, and more entries would generate less valid data points in a cluster, they both could hurt the accuracy of the workload clustering. The trace information used to conduct the workload characterization include I/O timestamp, I/O size, device number, block address, and operation types. We convert each window of

the I/O traces into a data point, and use Principal Component Analysis [7] to transfer the data points into **two dimensions**. After that, we use k-means to cluster these data points.

After we cluster all the data points of a new workload, we will calculate the distance between the center of the examined data points and the center of an existing cluster. If the distance is below a threshold, we make the conclusion that the majority of the examined data points fall into the existing workload cluster. In other words, the new workload belongs to this cluster. If LearnedSSD cannot identify a similar cluster, LearnedSSD will create a new cluster for the new workload. Our experimental results, as demonstrated in Figure 2, show that our learning-based workload clustering can successfully identify a cluster of storage workloads that belong to the same or similar **workload type**.

To further verify the effectiveness of the learning-based approach for workload clustering, we divide the same workload into the training and validation datasets, and evaluate the clustering results for each workload as listed in Table 1. We observe that 95% of their data points fall into the same workload cluster on average. This provides the evidence showing that our proposed learning-based approach is sufficient to identify an appropriate cluster for new workloads. It also drives the LearnedSSD design by offering the insight that it is feasible to develop an SSD which can deliver optimized performance for a category of application workloads.

3.4 Transfer SSD Specifications into ML Parameters

We now discuss how we can transfer the tuning problem of SSD specifications for a workload category into a ML problem, such that we can automate the tuning procedure with high accuracy. To address this problem, we first need to formulate the SSD specifications into ML parameters. However, this is not easy, since we have to ensure the parameter formulation does not lose the meaningful semantic of a specific SSD hardware parameter, and the ML parameters should be able to represent the characteristics of different hardware specifications and their correlations.

To model the SSD specifications via ML parameters, we formulate them into three major parts in the ML model: (1) the performance metrics used as the optimization targets for SSDs; (2) SSD hardware configurations that can be vectorized as parameters in a ML model; and (3) the configuration constraints (e.g., the SSD capacity) that bound the optimization space of the ML model. We describe each of them as follows.

Performance metrics used in the ML model. As for storage performance, LearnedSSD focuses on **the storage latency and throughput**. To quantify whether a SSD configuration delivers optimized performance or not, we use reference performance as the baseline (e.g., the latency and throughput obtained from a commercial SSD’s configurations), and the relative performance improvements as the evaluation metrics. We set the performance optimization goal as follows:

$$\begin{aligned} Goal(conf) = & (1 - \alpha) \times \log\left(\frac{Latency_{conf}}{Latency_{ref}}\right) \\ & - \alpha \times \log\left(\frac{Throughput_{conf}}{Throughput_{ref}}\right) \end{aligned} \quad (1)$$

where α is a tunable coefficient factor for balancing the latency and throughput at a proper scale. We set $\alpha = 0.9$ by default in LearnedSSD, based on our study of different coefficients. In our evaluation (see §4.5), we will examine the impact of the α on the learning efficiency. As discussed, the $Latency_{ref}$ and $Throughput_{ref}$ will be given as the reference performance in Formula 1 in the ML model.

SSD hardware specifications as ML parameters. To represent SSD hardware specifications in the ML models, we transfer them into four types of parameters and use different ways to set their values. They include *continuous*, *discrete*, *boolean*, and *categorical* parameters.

- **Continuous parameter:** typical examples of continuous parameter include over-provisioning ratio for GC, and the number of flash channels. To set the value of this type of parameters as we run the ML model, we identify a range of possible values it could take in advance, and divide the range uniformly into N small pieces. Therefore, LearnedSSD can take N endpoints as the possible values. For each continuous parameter, we set the range to cover all common values in commodity SSDs for ensuring the learned SSD specifications are practical.
- **Discrete parameter:** typical examples of discrete parameter include the SSD DRAM capacity, I/O queue depth, and page size. We select all their possible values and store them in a list. Therefore, we can use the list index in the vector of the ML model. Their possible values also cover all the common values. Each discrete parameter follows different rules as LearnedSSD sets the value at runtime. For instance, the SSD DRAM capacity will take the power of 2 as we increase it; and there are only five PCIe bandwidth settings, according to the PCIe protocol.
- **Boolean parameter:** we use the boolean parameters in the ML model to indicate whether a function or feature (e.g., statistic wear leveling, and greedy GC) will be enabled in the SSD or not. With the 0-1 boolean parameter, 0/1 means function enabled/disabled respectively.
- **Categorical parameter:** As for the categorical parameter, we convert it to the dummy variable [28]. For example, there are 16 possible values for the plane allocation scheme, we create a list with the length of 16. When LearnedSSD selects one scheme, it will set the value of the corresponding index of the list to 1, and others to 0.

Configuration constraints. LearnedSSD allows users to specify the configuration constraints for its configuration tuning. Typical examples include the SSD storage capacity

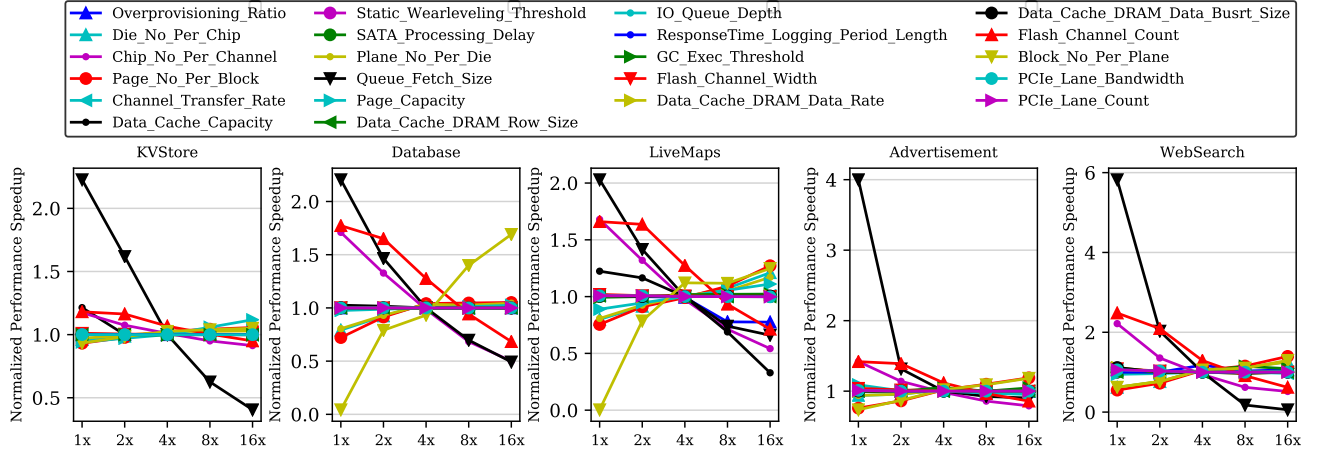


Figure 4. A study of coarse-grained parameter pruning for different storage workloads.

and the interface (e.g., NVMe or SATA) supported by the SSD for interacting with the host machine. When LearnedSSD sets different values for its ML parameters, it will simply abandon those configurations that violate the specified constraints. And LearnedSSD will always check the constraints when it learns a new configuration. Note that LearnedSSD mainly works on the tunable parameters in SSD specifications. For the lower-level circuit-relevant specifications that are strictly limited by the hardware, LearnedSSD does not cover them in its tuning model.

3.5 Learning-based Parameter Pruning

After we transfer the SSD specifications into ML parameters, we can start to train the model. However, modern SSDs usually have hundreds of hardware specifications or parameters. Although ML models today can handle a large set of parameters, it is still desirable to develop efficient and lightweight models for reducing both training and learning time as well as saving computation cycles. For example, we develop a model with 64 SSD parameters, it takes 30.7 hours to converge the model on a modern multi-core server (see the experimental setup in §4.1). Moreover, we find that not all SSD parameters are strongly correlated to the storage performance, and it is not needed to include these insensitive parameters in the learning model. To this end, we propose a parameter pruning approach to identify the impactful parameters that could affect the storage performance of SSDs.

However, we have to overcome two major challenges within the parameter pruning. First, we need an accurate measurement method to examine the importance of a parameter. This is challenging as SSD parameters usually have dependencies. As we tune a single parameter while keeping the values of other parameters fixed, this may violate the configuration constraints. For example, increasing the number of flash channels could violate the constraint of the SSD capacity. On the other hand, as we tune a single

parameter while updating the values of other parameters accordingly for meeting the configuration constraints, we cannot accurately determine which parameters affect the storage performance significantly. Second, removing some of the SSD parameters may hurt the overall accuracy of the learning model. And it is challenging to quantify how each SSD parameter could affect the learning accuracy. To address these challenges, we conduct the parameter pruning procedure with two stages.

Coarse-grained parameter pruning. We first adopt a coarse-grained pruning method that adjusts the values of continuous and discrete numerical parameters with large stride length. But we ensure the values of these parameters are configured in a reasonable range, and still satisfy the configuration constraints. At this stage, we eliminate the parameters that do not have much impact on the storage performance, no matter how we change their values. As shown in Figure 4, we increase the values of the 23 numerical parameters of SSDs from their baseline setting to 16 \times , and measure the storage performance with different workloads. We observe that some parameters do not affect the storage performance significantly (those flat lines in Figure 4), we call them as *insensitive parameters* in this paper. We also find that these insensitive parameters will be different for different storage workload types, therefore, LearnedSSD will conduct the coarse-grained parameter pruning for each workload type and identify the corresponding insensitive parameters. In general, we identify 9 insensitive parameters, such as *Page_Metadata_Size* and *SATA_Processing_Delay*, according to our study (see Figure 4). Note that these insensitive parameters would be updated for a new workload type.

Fine-grained parameter pruning. After eliminating the insensitive parameters with the coarse-grained pruning, we continue the parameter pruning with a fine-grained approach.

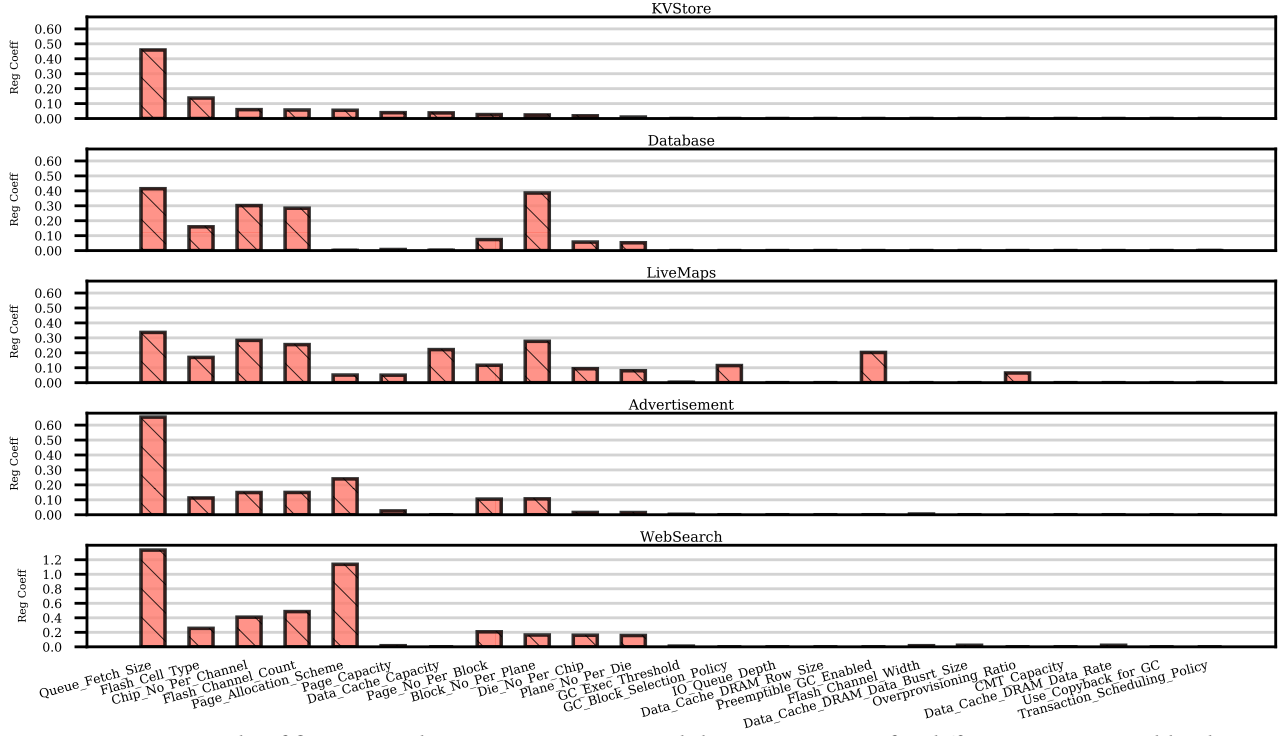


Figure 5. A study of fine-grained parameter pruning with linear regression for different storage workloads.

It employs the [linear regression technique LASSO](#) [6] to identify the linear correlations between the SSD parameters and performance. Following the discussion in §3.4, we set a regression space by maintaining the SSD capacity constraint, as we vary the values of SSD parameters. Since the SSD capacity is mainly determined by the parameters related to the chip layout, such as flash page size, the number of flash channels, and the flash block size, we first set the values for these parameters. After that, we vary the values of other parameters, and measure the regression coefficient for each SSD parameter. A higher regression coefficient score of a parameter means it has a closer correlation with the SSD performance. Based on the reported coefficient scores, we abandon the parameter whose score is below a threshold (0.01 by default in LearnedSSD). Therefore, we can focus on the parameter tuning for the important ones. As shown in Figure 5, we remove the insensitive parameters identified by the coarse-grained parameter tuning, and also include the boolean and categorical parameters in the fine-grained parameter tuning. Given the threshold of 0.01 for the regression coefficient score, we can eliminate more insensitive parameters for different workloads.

Observations. The learning-based parameter pruning of LearnedSSD can not only help us to eliminate the insensitive parameters, but also offer interesting insights that would benefit SSD development. Specifically, we observe that: (1) Different workloads have different parameter sensitivity. For

example, the performance of latency-critical workloads like Advertisement and WebSearch is not sensitive to the page size and over-provisioning ratio of SSDs, as they are read intensive. In contrast, the I/O-intensive workloads such as key-value stores and LiveMaps are sensitive to the flash page size. (2) Not all parameters have linear correlation with SSD performance, which generates difficulties for manual tuning, and further motivates us to utilize ML techniques to pinpoint the optimal SSD specifications.

3.6 Automated Tuning of SSD Configurations

After the SSD parameter pruning, we now develop the ML model to learn the optimal SSD specifications for various workloads. We present the system workflow of LearnedSSD in Figure 6. Given a workload, LearnedSSD will first use the configurations stored in the ConfDB as the initial configuration set, and leverage both Gaussian process regression (GPR) and discrete stochastic gradient descent (SGD) algorithms to learn different configurations. For each learned configuration, LearnedSSD will use a cycle-accurate SSD simulator to validate its performance until the model converges (i.e., the optimal SSD configuration is identified). In the following, we discuss each step of the workflow in details.

Identify the initial configuration set for a new workload. For a new workload, LearnedSSD will use the learning-based workload clustering as discussed in §3.3 to cluster the workload, and look up the learned configurations for the

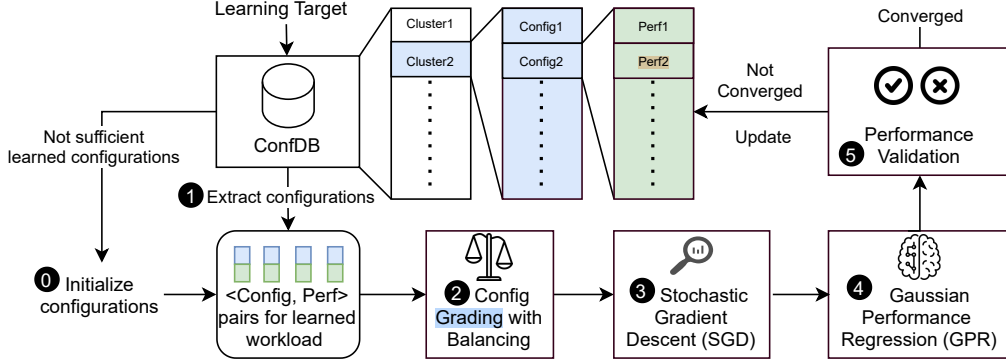


Figure 6. The automated learning workflow of LearnedSSD.

corresponding workload cluster in ConfDB (①). LearnedSSD will use these configurations and their delivered performance to initialize the ML model. However, if there are insufficient configurations in ConfDB (e.g., the ConfDB is empty), LearnedSSD will use a configuration from existing commodity SSDs and its measured performance to initialize the model (①).

Quantify the configurations with a unified grading mechanism. LearnedSSD will start the configuration tuning procedure based on the initial configurations. In order to check the effectiveness of these configurations, LearnedSSD develops a **grading mechanism** (②) with the goal of unifying different performance metrics (see §3.4). To achieve the maximal optimizations for both **data access latency and throughput**, LearnedSSD uses the Formula 1 as the goal. To ensure the learned configuration for a target workload does not hurt the performance of other workloads, LearnedSSD introduces a new factor β named *penalty balance* in its grading. Therefore, we define the performance grade for a workload as follows:

$$\begin{aligned} \text{Grade}(\text{conf}) &= (1 - \beta) \times \text{Goal}(\text{conf}) \\ &+ (\beta) \times \frac{\sum_{\text{conf}' \in \text{non-target}} \text{Goal}(\text{conf}')}{\text{NumClusters} - 1} \end{aligned} \quad (2)$$

where $\text{Goal}(\text{conf})$ is defined in Formula 1, and $\beta = 0.9$ by default based on our study (see Figure 10 in our evaluation).

Search optimal configurations with the stochastic gradient descent technique. With the initial SSD configurations and their grades, LearnedSSD will use SGD to search an optimal configuration (③). Specifically, LearnedSSD first identifies the top three best configurations (i.e., the configurations whose grades rank at the top) from the learned configurations, and randomly selects one as the search root. In the gradient descent process, LearnedSSD expands the search space from the root by checking all the adjacent configurations under configuration constraints (e.g., SSD capacity) in each searching iteration. Given the capacity constraint,

LearnedSSD tunes one or two relevant parameters at one time, and then keep those parameters that satisfy the constraint.

For other parameters, LearnedSSD will adjust their values back and forth in the search space. Once we finalize the parameters for one configuration, LearnedSSD will use the GPR model to identify the configuration with the best predicted performance grade (④). If its performance grade is better than the search root, LearnedSSD will **set this configuration as the new search root** and continue the next search iteration.

The main challenge with the SGD procedure (③) is to balance the learning accuracy and **exploitation overhead**. Since there is no guarantee that the initial configuration set will cover the entire search space, LearnedSSD has to gradually expand its search space to ensure it can identify the optimal ones. However, this may cause a search space explosion. To address this issue, we introduce a heuristic exploit factor, which is the minimum Manhattan distance [77] between the configuration being exploited and the existing learned configurations. We also set a threshold for the number of search iterations (20 iterations by default in LearnedSSD) in the configuration exploration.

Predict the grades of explored configurations. As discussed briefly in previous descriptions, LearnedSSD uses GPR [59] to **predict the grades for new configurations** (④). This is for three major reasons. First, GPR can provide nearly the same performance as the deep neural networks, especially in the modeling of searching optimal configurations and making recommendations. Second, it offers excellent trade-offs between the explorations of new knowledge and learned knowledge [44, 67]. Third, GPR provides confidence intervals with low computation overhead by default [9].

In LearnedSSD, we build a new GPR model by specifying its mean function and covariance function. The mean function is configured as trainable, as the mean of the performance metrics is unknown before the learning in LearnedSSD. We use the covariance function to represent the correlation

between two adjacent points in the model, and adopt both radial basis function (RBF) kernel [75] and rational quadratic kernel [76] as the regression covariance. We also add a white kernel [47] for random noise simulation.

Validate the explored configurations. The learning procedure of a new configuration will terminate by checking two conditions: (1) no configuration is better than the current root configuration in the search space; or (2) the search exceeds the threshold for the number of iterations. After that, LearnedSSD will use a cycle-accurate SSD simulator to validate the efficiency of the learned configurations (5). LearnedSSD will run all the available workloads in the ConfDB with the SSD simulator, and report the grade for the tested configuration. Before the validation, LearnedSSD will warm up the SSD simulator by running diverse workload traces randomly. In the validation, LearnedSSD maintains a set of optimal configurations whose grades rank at the top of all the learned configurations. After a certain number of search iterations, if the overall grade of this configuration set is not significantly updated, the learning procedure will be converged. Otherwise, LearnedSSD will update the ConfDB with the new learned configuration and start another search iteration until the learning procedure is converged.

3.7 Implementation Details

We implement the LearnedSSD framework with Python programming language. LearnedSSD supports the storage traces collected with *blktrace* which is available on a majority of computing systems. It uses Principal Component Analysis and k-means algorithms in the learning-based workload clustering. LearnedSSD utilizes the Sklearn library [5] to develop the statistic learning model that supports both SGD and GPR algorithms. LearnedSSD adopts the MQSim [70] as the back-end SSD simulator to validate the learned configurations. Note that LearnedSSD is also compatible with other SSD simulators, therefore, SSD vendors can replace the open-sourced MQSim simulator with their own simulators. LearnedSSD implements the ConfDB with the key-value store LevelDB, in which the key is the workload cluster ID, and the value includes the corresponding SSD configurations and their performance obtained from the SSD simulator. The value is organized in JSON format. LearnedSSD provides a simple interface *set_cons* (*capacity*, *interface*, *flash_type*) to enable end users to specify their configuration constraints SSD capacity, interface (i.e., NVMe or SATA), and the flash type (i.e., SLC, MLC, and TLC). We will open source LearnedSSD to benefit future study.

3.8 Discussion and Future Work

In this work, LearnedSSD mainly works under the configurations constraints that include storage capacity, interfaces, and flash types. However, its learning techniques and workflow

Table 1. Application workloads used in our evaluation.

Workload Category	Description
KVStore	The YCSB benchmarks are executed against LevelDB.
Database	TPCC and TPCE are executed against MySQL.
LiveMaps	The storage trace collected from a map service.
Advertisement	The storage trace collected from advertisement servers.
BatchDataAnalytics	The storage trace collected from MapReduce workloads.
WebSearch	The storage trace collected from web search services.
CloudStorage	The storage trace collected from a cloud storage service.

are also suitable for identifying the optimal SSD configurations with other constraints, such as the economic cost and energy efficiency of the SSD. Unfortunately, a majority of SSD vendors are not willing to open source the specifications of the cost and energy consumption of each hardware component of the SSD. Therefore, we do not study these configuration constraints in this work. We wish to explore these dimensions as the future work.

4 Evaluation

Our evaluation shows that: (1) LearnedSSD can learn optimal SSD configurations for a given workload, and the learned configurations can deliver improved storage performance, compared with commodity SSD configurations (§4.2); (2) LearnedSSD can instantly learn an optimal configuration with low performance overheads (§4.3); (3) LearnedSSD works efficiently under different configuration constraints (§4.4); and (4) LearnedSSD itself is also tunable for satisfying various performance requirements from end users (§4.5).

4.1 Experimental Setup

In our evaluation, we use 7 different workload categories as shown in Table 1. These workloads cover various workload types that include key-value stores, databases, map services, advertisement recommendations, batch data analytics, web search services, and cloud storage [40]. Each workload type includes multiple storage traces. All the storage traces are either collected from university servers or enterprise servers.

We run the LearnedSSD framework on a server, which is configured with 48 Intel Xeon CPU (E5-2687W v4) processors running at 3.0GHz, 96GB DRAM, and 4TB SSD. Since LearnedSSD uses the statistic learning models, it does not require GPUs in its learning procedure. We use the configurations of Intel 590 SSD, Samsung 850 PRO SSD and Z-SSD as the baselines, and compare the learned configurations with them to evaluate the efficiency of LearnedSSD.

4.2 Efficiency of Learned Configurations

We first evaluate the efficiency of the learned configurations with LearnedSSD. We use the Intel 590 SSD as the reference. We set the configuration constraints as [*SSD capacity* = 1TB, *interface* = NVMe, *flash type* = MLC]. With its configuration in the SSD simulator, we run all the workloads in Table 1 to measure their performances. After that, we use the reference configuration and the measured performances to initialize ConfDB. And then, we feed the storage traces from different

Table 2. Performance of the learned configurations for NVMe MLC SSDs (normalized to Intel 590 SSD).

BLC/TGT	Advertisement	KVStore	Database	WebSearch	BatchDataAnalytics	CloudStorage	LiveMaps
Advertisement	3.99/1.00	1.67/1.00	1.59/0.99	1.50/1.00	1.61/1.00	1.62/1.00	1.30/1.00
KVStore	2.13/1.00	31.11/1.00	30.96/1.00	6.06/1.00	30.67/1.00	30.84/1.00	22.41/1.00
Database	3.20/1.50	6.64/1.50	6.89/1.50	3.94/1.50	6.84/1.50	6.86/1.50	4.11/1.49
WebSearch	1.00/1.00	1.25/1.00	1.26/1.00	1.28/1.00	1.22/1.00	1.27/1.00	0.98/1.00
BatchDataAnalytics	1.23/1.22	9.48/5.25	9.69/5.25	1.98/1.17	9.81/5.25	9.49/5.25	9.81/3.50
CloudStorage	1.35/1.19	28.11/4.16	27.72/4.16	1.58/1.14	28.82/4.16	29.30/4.16	11.11/3.57
LiveMaps	0.67/1.00	2.94/1.00	3.07/1.00	4.95/1.00	2.99/1.00	3.21/1.00	34.61/1.00
Average	1.94/1.13	11.6/2.13	11.60/2.12	3.04/1.12	11.70/2.13	11.80/2.13	12.04/1.79

Table 3. Learned configurations for different workloads. AD: Advertisement, KV: KVStore, DB: Database, WS: WebSearch, BD: BatchDataAnalytics, CS: CloudStorage, LM: LiveMaps.

Parameters	Intel 590	AD	KV/DB/CS/BD	WS	LM
IOQueueDepth	8192	4096	8192	4096	8192
QueueFetchSize	3072	3072	2048	4096	3072
DataCacheCapacity	800MB	512MB	2GB	800MB	2GB
CMTCapacity	2MB	2MB	2MB	4MB	8MB
PageAllocationScheme	CWDP	PWCD	CDPW	PDCW	PWCD
OverprovisioningRatio	0.22	0.22	0.22	0.20	0.20
FlashChannelCount	12	12	16	16	20
ChipNoPerChannel	5	6	4	5	7
DieNoPerChip	8	8	8	4	2
PlaneNoPerDie	1	1	1	1	1
BlockNoPerPlane	512	512	512	1024	1024
PageNoPerBlock	512	512	512	512	512
PageSize	4096	4096	16384	4096	8192

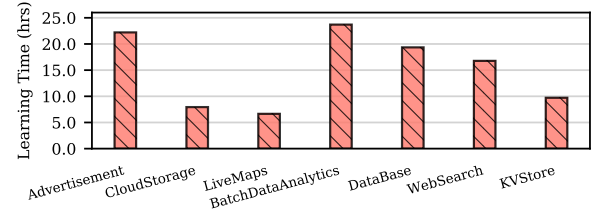
workload types into LearnedSSD to learn new configurations. Once the learning converges, LearnedSSD will report the best SSD configuration for each workload.

We show the performance of learned configurations in Table 2. In comparison with Intel 590 SSD, the learned SSD configurations can reduce the storage latency by 1.28–34.61× for the target workload, while decreasing the storage latency by 8.38× on average for non-target workloads. The learned configurations can also improve the storage throughput by up to 5.25× for bandwidth-intensive applications such as the database and cloud storage workloads, without hurting the throughput of other workloads.

To further understand the learned configurations, we list the critical parameters of the learned configurations in Table 3, in comparison with the reference configuration of Intel 590 SSD. As we can see, for different target workloads, LearnedSSD will learn different values for these parameters, although some workload types share the same configurations. To achieve improved SSD performance, LearnedSSD increases the number of channels for most of the workloads, while adjusting the flash chip layout (e.g., the number of chips per channel) accordingly to satisfy the SSD capacity constraint. LearnedSSD will also adjust the page allocation scheme [71] to optimize the data layout for different workloads. Similarly, LearnedSSD tunes the SSD DRAM capacity and I/O queue depth according to workload characteristics. Our work LearnedSSD proves that it is feasible to utilize ML techniques to automate the tuning of SSD configurations.

4.3 Learning Time of LearnedSSD

We now examine the learning time of LearnedSSD. We report the numbers for different target workloads in Figure 7.

**Figure 7.** Learning time of LearnedSSD for different workloads.**Table 4.** Overhead sources of LearnedSSD.

Component	Execution Time (secs)
Extract workload features per 100K I/O requests	0.84
Workload Similarity Comparison	4.65
Workload Clustering	0.57
ConfDB Database Lookup	0.02
New configuration learning per iteration	37.3
Simulator Validation	691.1

LearnedSSD can learn an optimal configuration in 6.65–23.70 hours. And it will incur 121 search iterations on average to pinpoint the optimal configuration. To further understand the overhead source of LearnedSSD, we profile the execution time of its critical components on the multi-core server as described in §4.1, and show the results in Table 4. Our profiling results demonstrate that LearnedSSD can finish each search iteration within only 37.3 seconds. And the major performance overhead of LearnedSSD comes from the simulator validation, as we need to warm up the simulator before each validation. However, LearnedSSD only needs to validate the best configuration (selected based on the predicted grade with GPR) in each search iteration.

4.4 Sensitivity to Configuration Constraints

We now evaluate how LearnedSSD performs as we change the configuration constraints that include the flash types and device interface. To evaluate the sensitivity to flash types, we use Samsung Z-SSD, which is a NVMe SLC SSD, as the reference configuration. To evaluate the sensitivity to device interface, we use Samsung 850 PRO, which is a SATA MLC SSD, as the reference configuration.

We present the performance of the learned configurations for different workloads in Table 5 and Table 6 respectively. Table 5 shows that the configurations learned by LearnedSSD can reduce the storage latency by 3.61–19.28×, and improve the storage throughput by up to 8.00× for NVMe SLC SSDs for the target workload, compared to the Samsung Z-SSD.

Table 5. Performance of learned configurations for NVMe SLC SSDs (normalized to Samsung Z-SSD).

BLC/TGT	Advertisement	KVStore	Database	WebSearch	BatchDataAnalytics	CloudStorage	LiveMaps
Advertisement	3.99/1.00	3.35/1.00	1.21/1.00	3.25/1.00	3.64/1.00	3.92/1.00	3.71/1.00
KVStore	2.13/1.00	11.49/1.00	2.57/1.00	0.78/1.00	9.65/1.00	2.17/1.00	9.53/1.00
Database	3.20/1.50	1.31/1.49	3.61/1.50	0.91/1.50	3.33/1.49	3.26/1.50	2.00/1.49
WebSearch	1.00/1.00	1.01/1.00	0.63/1.00	3.78/1.00	0.88/1.00	0.99/1.00	1.01/1.00
BatchDataAnalytics	5.18/4.01	1.50/2.00	8.72/4.01	11.34/4.01	9.72/4.00	5.04/4.01	1.59/2.00
CloudStorage	14.17/8.00	1.55/1.60	28.04/1.60	14.24 / 1.60	21.99/1.60	14.25/8.00	2.43/2.00
LiveMaps	0.67/1.00	0.44/1.00	6.98/1.00	7.5/1.00	1.05/1.00	0.66/1.00	19.28/1.00
Average	4.03/2.30	2.95/1.30	7.39/1.59	5.97/1.59	7.18/1.59	4.33/2.50	5.65/1.36

Table 6. Performance of learned configurations for SATA MLC SSDs (normalized to Samsung 850 PRO).

BLC/TGT	Advertisement	KVStore	Database	WebSearch	BatchDataAnalytics	CloudStorage	LiveMaps
Advertisement	1.41/1.00	1.05/1.00	1.40/1.00	1.07/1.00	1.39/1.00	1.06/1.00	1.02/1.00
KVStore	1.00/1.00	1.01/1.00	0.99/1.00	1.00/1.00	1.00/1.00	0.97/1.00	0.98/1.00
Database	1.00/1.00	0.99/1.00	1.01/1.00	0.95/1.00	1.00/1.00	1.04/1.00	1.03/1.00
WebSearch	1.00/1.00	1.01/1.00	1.02/1.00	1.04/1.00	0.99/1.00	1.02/1.00	0.98/1.00
BatchDataAnalytics	1.03/1.00	0.97/1.00	1.00/1.00	1.04/1.00	1.07/1.00	1.03/1.00	1.05/1.00
CloudStorage	1.56/1.31	2.42/2.63	1.64/1.31	2.41/2.63	1.22/1.11	2.38/2.63	2.34/2.63
««« HEAD LiveMaps	1.67/1.00	8.30/1.00	1.75/1.00	8.19/1.00	1.01/1.00	8.32/1.00	8.41/1.00
Average	1.24/1.04	2.25/1.23	1.26/1.04	2.24/1.23	1.09/1.02	2.16/1.23	2.26/1.23

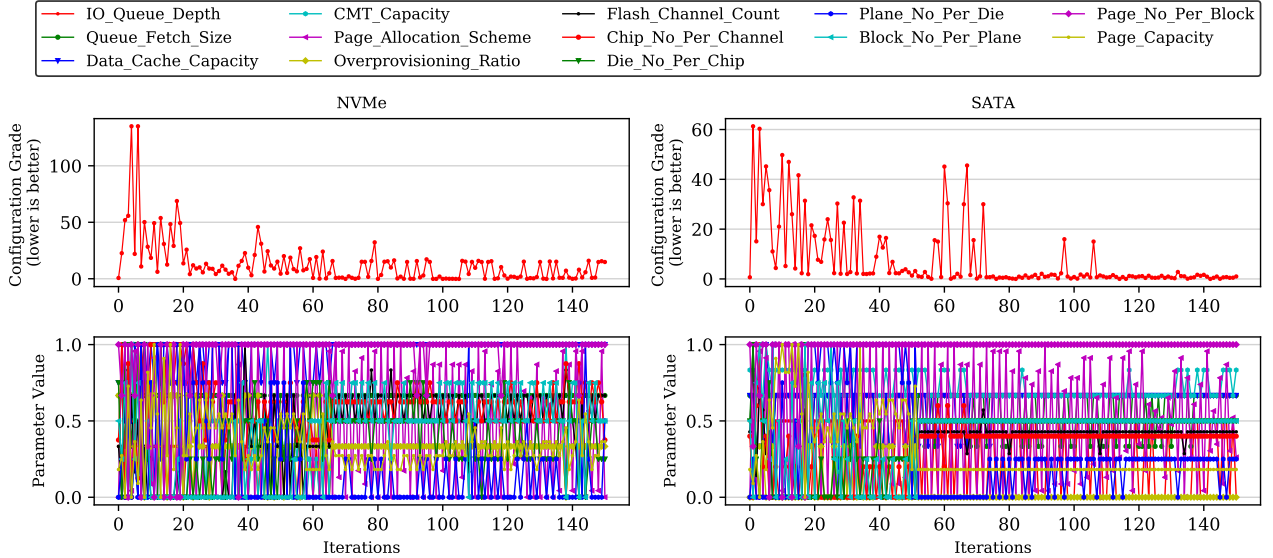
**Figure 8.** The learning procedure of LearnedSSD for NVMe and SATA SSDs, as we target batch data analytics workload.

Table 6 demonstrates that our learned configurations can deliver up to 8.41 \times latency reduction and 2.63 \times throughput improvement for SATA SSDs for the target workload, in comparison with Samsung 850 PRO.

In order to understand how LearnedSSD tunes the device parameters, we record its learning procedure and configuration grades at runtime. As shown in Figure 8, we present the profiling results of learning the optimal configurations for the target workload *BatchDataAnalytics* for NVMe and SATA SSDs, respectively. The top subfigures in Figure 8 demonstrate how the configuration grade will be updated after each search iteration. As discussed in §3.6, the learning procedure will converge when the grade of the configurations becomes stable. We show the learning procedure of the critical SSD parameters in the below subfigures in Figure 8. We observe that, with different configuration constraints

(NVMe vs. SATA), (1) the learning procedure will be different; (2) for each parameter, its correlation with the SSD performance is also different, making it impossible for developers to manually tune them; (3) not all parameters are equal, some parameters are insensitive to storage performance. LearnedSSD framework can help developers identify such parameters for different workloads under different configuration constraints, which could improve the productivity of SSD development.

4.5 Performance Impact of the Balance Coefficient

As discussed in §3.4 and §3.6, LearnedSSD uses the coefficient factor α (Formula 1) to balance the storage latency and throughput in the learning procedure, and defines the coefficient factor β (Formula 2) to balance the penalty (weight) between the target workload and non-target workloads. Both of them are tunable in LearnedSSD, which allows end users

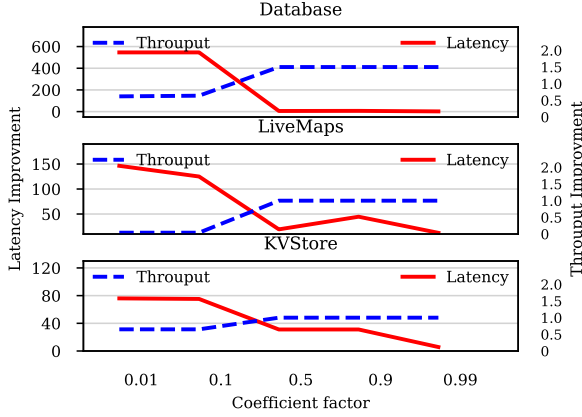


Figure 9. Performance impact of the coefficient factor for balancing the latency and throughput for a target workload.

to adjust them per their needs. In this part, we evaluate their impact on storage performance. We vary their values from 0.01 to 0.99, and measure the performance of the learned configurations for the three representative workloads database, key-value store, and LiveMaps. As we examine each value of α and β , we reset the ML model and initialize the ConfDB. We show the experimental results in Figure 9 and Figure 10.

With the coefficient factor α , our goal is to achieve the maximum improvement for both latency and throughput. In Figure 9, as we increase the value of α from 0.01 to 0.3, the latency of the target workload is dramatically improved, however, its throughput is lower than the reference configuration. As we further increase its value to 0.9, we can achieve both improved latency and throughput for all the three target workloads. Thus, LearnedSSD sets $\alpha = 0.9$ by default.

With the coefficient factor β , our goal is to achieve the maximum performance improvement for both the target workload and non-target workloads. As LearnedSSD learns new configurations, it is usually easy to achieve the improved performance for the target workload. However, this may decrease the performance for non-target workloads, which could impede the widespread adoption of the learned configurations. As we vary the value of β , we observe that there is such a sweet spot ($\beta = 0.9$) that can delivers maximal performance improvement for the target workload, while having minimal negative impact on the non-target workloads.

5 Related Work

SSD Performance Optimization. SSDs has been widely used in modern storage systems to meet the I/O performance and storage capacity requirements of data-intensive applications, such as databases, cloud storage, web search, and big-data analytics [35, 46, 50, 63, 64, 81]. Although these applications have high demands on I/O performance and their workload have unique data access patterns [17, 34, 82], they normally employ generic SSD devices [13, 57, 65], which

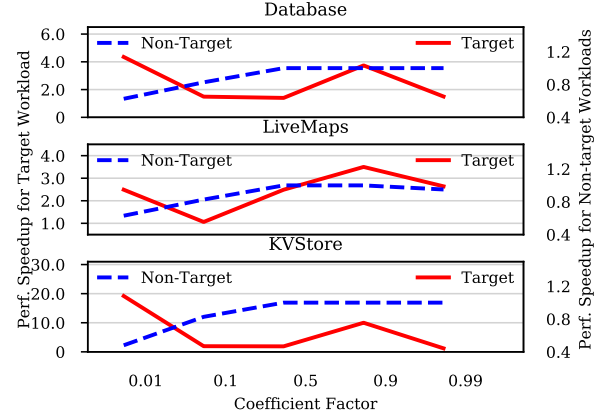


Figure 10. Performance impact of the coefficient factor for balancing the performance between the target workload and non-target workloads.

causes suboptimal performance and resource efficiency. In this paper, we develop LearnedSSD to facilitate the development of customized SSD devices for applications with improved performance. Recently, researchers proposed the software-defined flash and open-channel SSDs to enable applications to develop their own storage stack for improved storage utilization and performance isolation [36, 48, 56]. They show that there is an increasing demand on software-defined storage. However, there is a longstanding gap between the application demands and device specifications. We develop LearnedSSD with the goal of bridging this gap.

Machine Learning for Systems. Most recently, researchers have started to leverage machine learning techniques to solve system optimization problems, such as the task scheduling [58, 74, 83], cluster resource management [12, 18, 24, 52, 79], performance optimizations [33, 45, 51, 84], data management [10, 43, 49, 73], and others [53, 78]. However, few studies conduct a systematic investigation of applying the learning techniques to develop SSD devices. To the best of our knowledge, LearnedSSD is the first work that utilize the learning techniques to enable the automated tuning of SSD specifications. We believe it will not only benefit SSD vendors and manufacturers but also platform operators such as those for cloud services and data centers.

SSD Device Development. Along with the architecture innovation, the industry community has developed mature manufacturing techniques and fabrication process to produce new storage devices, such as Z-SSD [61], Optane SSD [39], ZNS SSDs [60, 85]. As the industrial revolution has moved into the fourth/fifth generation (Industry 4.0/5.0) powered by the artificial intelligence [29, 38], storage devices should also become highly customizable for applications. Unfortunately, we are lacking an effective framework that can transfer application demands into storage device development. In this work, we focus on building a learning-based framework to address a critical challenge with the SSD development – how

to efficiently identify the optimal SSD specifications for meeting the needs from target applications under constraints.

6 Conclusion

We build a learning-based framework named LearnedSSD for enabling the automated tuning of SSD specifications. Given a storage workload, LearnedSSD can efficiently learn an optimal SSD configuration that delivers the maximum performance improvement even under different configuration constraints. LearnedSSD can significantly reduce the manual efforts in the SSD device development. Our experiments show that our learned SSD configurations can significantly improve the storage performance for a target workload, without hurting the performance of non-target workloads.

References

- [1] How ssds are made: Phases of solid-state drive development, 2012. <https://www.gamernexus.net/guides/956-how-ssds-are-made>.
- [2] Aws web services, 2021. <https://aws.amazon.com/>.
- [3] Database-as-a-service, 2021. <https://www.trustradius.com/database-as-a-service-dbaas>.
- [4] The open compute project, 2021. <https://www.opencompute.org/>.
- [5] scikit-learn: Machine learning in python, 2021. <https://scikit-learn.org/stable/index.html>.
- [6] A Complete Understanding of LASSO Regression. <https://www.mygreatlearning.com/blog/understanding-of-lasso-regression/>.
- [7] A Step-by-Step Explanation of Principal Component Analysis (PCA). <https://builtin.com/data-science/step-step-explanation-principal-component-analysis>.
- [8] Nitin Agrawal, Vijayan Prabhakaran, Ted Wobber, John D. Davis, Mark Manasse, and Rina Panigrahy. Design tradeoffs for ssd performance. In *Proceedings of the 2008 USENIX Annual Technical Conference (ATC'08)*, Boston, Massachusetts, 2008.
- [9] Dana Van Aken, Dongsheng Yang, Sebastien Brillard, Ari Fiorino, Bohan Zhang, Christian Billian, and Andrew Pavlo. An Inquiry into Machine Learning-based Automatic Configuration Tuning Services on Real-World Database Management Systems. *Proceedings of the VLDB Endowment*, 14(7):1241–1253, 2021.
- [10] Dana Van Aken, Dongsheng Yang, Sebastien Brillard, Ari Fiorino, Bohan Zhang, Christian Billian, and Andrew Pavlo. An Inquiry into Machine Learning-based Automatic Configuration Tuning Services on Real-World Database Management Systems. *Proceedings of the VLDB Endowment*, 14(7):1241–1253, 2021.
- [11] Ibrahim Umit Akgun, Ali Selman Aydin, Aadil Shaikh, Lukas Velikov, and Erez Zadok. A machine learning framework to improve storage system performance. In *Proceedings of the 13th ACM Workshop on Hot Topics in Storage and File Systems (HotStorage'21)*, July 2021.
- [12] Pradeep Ambati, Inigo Goiri, Felipe Frujeri, Alper Gun, Ke Wang, Brian Dolan, Brian Corell, Sekhar Pasupuleti, Thomas Moscibroda, Sameh Elnikety, Marcus Fontoura, and Ricardo Bianchini. Providing slos for resource-harvesting vms in cloud platforms. In *Proceedings of the 14th USENIX Symposium on Operating Systems Design and Implementation (OSDI'20)*, November 2020.
- [13] Ali Anwar, Yue Cheng, Hai Huang, Jingoo Han, Hyogi Sim, Dongyoon Lee, Fred Douglass, and Ali R. Butt. Bspokv: Application tailored scale-out key-value stores. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC'18)*, 2018.
- [14] Artificial Neural Network. https://en.wikipedia.org/wiki/Artificial_neural_network.
- [15] Noman Bashir, Nan Deng, Krzysztof Rzadca, David Irwin, Sree Kodak, and Rohit Jnagal. Take it to the limit: Peak prediction-driven resource overcommitment in datacenters. In *Proceedings of the Sixteenth European Conference on Computer Systems (EuroSys'21)*, Virtual Event, April 2021.
- [16] Bayesian Network. https://en.wikipedia.org/wiki/Bayesian_network.
- [17] Jean Luca Bez, Francieli Zanon Boito, Ramon Nou, Alberto Miranda, Toni Cortes, and Philippe O. A. Navaux. Detecting i/o access patterns of hpc workloads at runtime. In *Proceedings of the 31st International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD'19)*, 2019.
- [18] Ricardo Bianchini, Marcus Fontoura, Eli Cortez, Anand Bonde, Alexandre Muzio, Ana-Maria Constantin, Thomas Moscibroda, Gabriel Magalhaes, Girish Bablani, and Mark Russinovich. Toward ml-centric cloud platforms. *Communication of ACM*, 63(2), January 2020.
- [19] Matias Björling, Javier Gonzalez, and Philippe Bonnet. LightNVM: The Linux Open-Channel SSD Subsystem. In *Proceedings of the 15th USENIX Conference on File and Storage Technologies (FAST'17)*, Santa Clara, CA, 2017.
- [20] Steve Burke. How ssds are made: Phases of solid-state drive development. <https://www.gamernexus.net/guides/956-how-ssds-are-made>, 2012.
- [21] Steve Burke. How ssds are made: Phases of solid-state drive development, 2012. <https://www.gamernexus.net/guides/956-how-ssds-are-made>.
- [22] Laura Caulfield. Project denali to define flexible ssds for cloud-scale applications, 2018. <https://azure.microsoft.com/en-us/blog/project-denali-to-define-flexible-ssds-for-cloud-scale-applications/>.
- [23] Feng Chen, Rubao Lee, and Xiaodong Zhang. Essential roles of exploiting internal parallelism of flash memory based solid state drives in high-speed data processing. In *Proceedings of the IEEE 17th International Symposium on High Performance Computer Architecture (HPCA'11)*. IEEE, 2011.
- [24] Eli Cortez, Anand Bonde, Alexandre Muzio, Mark Russinovich, Marcus Fontoura, and Ricardo Bianchini. Resource central: Understanding and predicting workloads for improved resource management in large cloud platforms. In *Proceedings of the 26th Symposium on Operating Systems Principles (SOSP'17)*, Shanghai, China, 2017.
- [25] Jeffrey Dean and Sanjay Ghemawat. Mapreduce: simplified data processing on large clusters. *Communications of the ACM*, 51(1):107–113, 2008.
- [26] Decision Tree. https://en.wikipedia.org/wiki/Decision_tree.
- [27] Chris Ding and Xiaofeng He. K-means clustering via principal component analysis. In *Proceedings of the Twenty-First International Conference on Machine Learning (ICML'04)*, page 29, Banff, Alberta, Canada, 2004.
- [28] Dummy Variable (Statistics). [https://en.wikipedia.org/wiki/Dummy_variable_\(statistics\)](https://en.wikipedia.org/wiki/Dummy_variable_(statistics)).
- [29] Fourth Industrial Revolution. https://en.wikipedia.org/wiki/Fourth_Industrial_Revolution.
- [30] Yu Gan, Mingyu Liang, Sundar Dev, David Lo, and Christina Delimitrou. Sage: Practical and scalable ml-driven performance debugging in microservices. In *Proceedings of the 26th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS'21)*, Virtual, USA, 2021.
- [31] Aayush Gupta, Youngjae Kim, and Bhuvan Ugaonkar. DFTL: A Flash Translation Layer Employing Demand-based Selective Caching of Page-level Address Mappings. In *Proceedings of the 14th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS'09)*, Washington, DC, USA, 2009.
- [32] Mingzhe Hao, Huaicheng Li, Michael Hao Tong, Chrisma Pakha, Riza O. Suminto, Cesar A. Stuardo, Andrew A. Chien, and Haryadi S. Gunawi. Mittos: Supporting millisecond tail tolerance with fast rejecting slow-aware os interface. In *Proceedings of the 26th Symposium on Operating*

- Systems Principles (SOSP'17)*, Shanghai, China, 2017.
- [33] Mingzhe Hao, Levent Toksoz, Nanqin Li, Edward Edberg Halim, Henry Hoffmann, and Haryadi S. Gunawi. LinnOS: Predictability on Unpredictable Flash Storage with a Light Neural Network. In *Proceedings of the 14th USENIX Symposium on Operating Systems Design and Implementation (OSDI'20)*, November 2020.
 - [34] Jun He, John Bent, Aaron Torres, Gary Grider, Garth Gibson, Carlos Maltzahn, and Xian-He Sun. I/o acceleration with pattern detection. In *Proceedings of the 22nd International Symposium on High-Performance Parallel and Distributed Computing (HPDC'15)*, page 25–36, New York, New York, USA, 2013.
 - [35] Yang Hu, Hang Liu, and H. Howie Huang. Tricore: Parallel triangle counting on gpus. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis (SC'18)*, 2018.
 - [36] Jian Huang, Anirudh Badam, Laura Caulfield, Suman Nath, Sudipta Sengupta, Bikash Sharma, and Moinuddin K. Qureshi. Flashblox: Achieving both performance isolation and uniform lifetime for virtualized ssds. In *Proceedings of the 15th USENIX Conference on File and Storage Technologies (FAST'17)*, Santa Clara, CA, February 2017.
 - [37] Jian Huang, Anirudh Badam, Laura Caulfield, Suman Nath, Sudipta Sengupta, Bikash Sharma, and Moinuddin K. Qureshi. FlashBlox: Achieving Both Performance Isolation and Uniform Lifetime for Virtualized SSDs. In *Proceedings of the 15th Usenix Conference on File and Storage Technologies (FAST'17)*, Santa clara, CA, 2017.
 - [38] Industry 5.0: Top 3 Things You Need to Know. <https://www.mastercontrol.com/gxp-lifeline/3-things-you-need-to-know-about-industry-5.0/>.
 - [39] Intel Corporation. Intel® Optane™ SSD DC P4801X Series. 2018.
 - [40] Swaroop Kavalanekar, Bruce Worthington, Qi Zhang, and Vishal Sharda. Characterization of storage workload traces from production windows servers. In *Proceedings of the 2008 IEEE International Symposium on Workload Characterization*, 2008.
 - [41] Kingston. Design-in solid state drives for system designers and builders. <https://www.kingston.com/unitedstates/us/embedded/design-in-ssd>, 2021.
 - [42] Gunjae Koo, Kiran Kumar Matam, Te I, H. V. Krishna Giri Narra, Jing Li, Hung-Wei Tseng, Steven Swanson, and Murali Annavaram. Summarizer: Trading communication with computing near storage. In *Proceedings of the 50th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO'17)*, Cambridge, Massachusetts, 2017.
 - [43] Tim Kraska, Umar Farooq Minhas, Thomas Neumann, Olga Papaemmanouli, Jignesh M. Patel, Chris Re, and Michael Stonebraker. ML-In-Databases: Assessment and Prognosis. *IEEE Computer Society Technical Committee on Data Engineering*, 2021.
 - [44] Andreas Krause and Cheng Soon Ong. Contextual gaussian process bandit optimization. In *Proceedings of the 24th International Conference on Neural Information Processing Systems (NIPS'11)*, Granada, Spain, 2011.
 - [45] Daniar H. Kurniawan, Levent Toksoz, Anirudh Badam, Tim Emami, Sandeep Madireddy, Robert B. Ross, Henry Hoffmann, and Haryadi S. Gunawi. Ionet: Towards an open machine learning training ground for i/o performance prediction. *Technical Report*, 2021.
 - [46] Aapo Kyrola, Guy Blelloch, and Carlos Guestrin. Graphchi: Large-scale graph computation on just a pc. In *Proc. 10th USENIX OSDI*, Hollywood, CA, October 2012.
 - [47] Scikit Learn. White kernel. https://scikit-learn.org/stable/modules/generated/sklearn.gaussian_process.kernels.WhiteKernel.html, 2021.
 - [48] Sungjin Lee, Ming Liu, Sangwoo Jun, Shuotao Xu, Jihong Kim, and Arvind. Application-managed flash. In *Proceedings of the 14th USENIX Conference on File and Storage Technologies (FAST'16)*, Santa Clara, CA, February 2016.
 - [49] Guoliang Li, Xuanhe Zhou, Shifu Li, and Bo Gao. Qtune: A query-aware database tuning system with deep reinforcement learning. *Proceedings of the VLDB Endow.*, 12(12), August 2019.
 - [50] Hang Liu and H. Howie Huang. Simd-x: Programming and processing of graph algorithms on gpus. In *Proceedings of the 2019 USENIX Annual Technical Conference (USENIX ATC'19)*, Renton, WA, 2019.
 - [51] Martin Maas, David G. Andersen, Michael Isard, Mohammad Mahdi Javanmard, Kathryn S. McKinley, and Colin Raffel. Learning-based memory allocation for c++ server workloads. In *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS'20)*, Lausanne, Switzerland, 2020.
 - [52] Hongzi Mao, Malte Schwarzkopf, Shaileshh Bojja Venkatakrishnan, Zili Meng, and Mohammad Alizadeh. Learning scheduling algorithms for data processing clusters. In *Proceedings of the ACM Special Interest Group on Data Communication (SIGCOMM'19)*, Beijing, China, 2019.
 - [53] Azalia Mirhoseini, Anna Goldie, Mustafa Yazgan, Joe Jiang, Ebrahim M. Songhori, Shen Wang, Young-Joon Lee, Eric Johnson, Omkar Pathak, Sungmin Bae, Azade Nazi, Jiwoo Pak, Andy Tong, Kavaya Srinivasa, William Hang, Emre Tuncer, Anand Babu, Quoc V. Le, James Laudon, Richard C. Ho, Roger Carpenter, and Jeff Dean. Chip placement with deep reinforcement learning. *CoRR*, abs/2004.10746, 2020.
 - [54] Mihir Nanavati, Jake Wires, and Andrew Warfield. Decibel: Isolation and sharing in disaggregated rack-scale storage. In *14th USENIX Symposium on Networked Systems Design and Implementation (NSDI'17)*, Boston, MA, March 2017.
 - [55] Jian Ouyang, Shiding Lin, Song Jiang, Zhenyu Hou, Yong Wang, and Yuanzheng Wang. SDF: Software-defined Flash for Web-scale Internet Storage Systems. In *Proceedings of the 19th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS'14)*, Salt Lake City, UT, 2014.
 - [56] Jian Ouyang, Shiding Lin, Song Jiang, Yong Wang, Wei Qi, Jason Cong, and Yuanzheng Wang. SDF: Software-Defined Flash for Web-Scale Internet Storage Systems. In *Proceedings of 19th International Conference on Architectural Support for Programming Language and Operating Systems (ASPLOS'14)*, Salt Lake City, UT, March 2014.
 - [57] Arnab K. Paul, Brian Wang, Nathan Rutman, Cory Spitz, and Ali R. Butt. Efficient Metadata Indexing for HPC Storage Systems. In *2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGrid'20)*, Virtual Event, 2020.
 - [58] Haoran Qiu, Subho S. Banerjee, Saurabh Jha, Zbigniew T. Kalbarczyk, and Ravishankar K. Iyer. FIRM: An intelligent fine-grained resource management framework for slo-oriented microservices. In *Proceedings of the 14th USENIX Symposium on Operating Systems Design and Implementation (OSDI'20)*, November 2020.
 - [59] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006.
 - [60] Samsung Introduces Its First ZNS SSD with Maximized User Capacity and Enhanced Lifespan. <https://news.samsung.com/global/samsung-introduces-its-first-zns-ssd-with-maximized-user-capacity-and-enhanced-lifespan>.
 - [61] Samsung Semiconductors. Ultra-Low Latency with Samsung Z-NAND SSD. *Technical Report*, 2018.
 - [62] Sudharsan Seshadri, Mark Gahagan, Sundaram Bhaskaran, Trevor Bunker, Arup De, Yanqin Jin, Yang Liu, and Steven Swanson. Willow: A User-programmable SSD. In *Proceedings of the 11th USENIX Conference on Operating Systems Design and Implementation (OSDI'14)*, Broomfield, CO, 2014.
 - [63] Mo Sha, Yuchen Li, Bingsheng He, and Kian-Lee Tan. Accelerating Dynamic Graph Analytics on GPUs. *Proceedings of VLDB Endowment*, 11(1), September 2017.
 - [64] Xuanhua Shi, Zhigao Zheng, Yongluan Zhou, Hai Jin, Ligang He, Bo Liu, and Qiang-Sheng Hua. Graph Processing on GPUs: A Survey. *ACM Computing Survey*, 50(6), January 2018.
 - [65] Hyogi Sim, Geoffroy Vallée, Youngjae Kim, Sudharshan S. Vazhkudai, Devesh Tiwari, and Ali R. Butt. An analysis workflow-aware storage system for multi-core active flash arrays. *IEEE Transactions on Parallel*

- and *Distributed Systems*, 30(2), 2019.
- [66] Gagan Somashekar and Anshul Gandhi. Towards optimal configuration of microservices. In *Proceedings of the 1st Workshop on Machine Learning and Systems (EuroMLSys'21)*, Online, United Kingdom, 2021.
 - [67] Niranjan Srinivas, Andreas Krause, Sham Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the 27th International Conference on International Conference on Machine Learning (ICML'10)*, Haifa, Israel, 2010.
 - [68] StorPool Distributed Storage. Choosing the right ssds for your cloud: Datacenter ssds vs. consumer ssds, 2019. <https://storpool.com/blog/choosing-the-right-ssds-for-your-cloud-datacenter-ssds-vs-consumer-ssds>.
 - [69] Support-Vector Machine. https://en.wikipedia.org/wiki/Support-vector_machine.
 - [70] Arash Tavakkol, Juan Gómez-Luna, Mohammad Sadrosadati, Saugata Ghose, and Onur Mutlu. Mqsim: A framework for enabling realistic studies of modern multi-queue SSD devices. In *Proceedings of the 16th USENIX Conference on File and Storage Technologies (FAST'18)*, Oakland, CA, February 2018.
 - [71] Arash Tavakkol, Pooyan Mehrvarzy, Mohammad Arjomand, and Hamid Sarbazi-Azad. Performance evaluation of dynamic page allocation strategies in ssds. *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, June 2016.
 - [72] PyTorch Team. Pytorch: From research to production, 2021. <https://pytorch.org/>.
 - [73] Dana Van Aken, Andrew Pavlo, Geoffrey J. Gordon, and Bohan Zhang. Automatic database management system tuning through large-scale machine learning. In *Proceedings of the 2017 ACM International Conference on Management of Data (SIGMOD'17)*, Chicago, Illinois, USA, 2017.
 - [74] Yawen Wang, Kapil Arya, Marios Kogias, Manohar Vanga, Aditya Bhandari, Neeraja J. Yadwadkar, Siddhartha Sen, Sameh Elnikety, Christos Kozyrakis, and Ricardo Bianchini. Smartharvest: Harvesting idle cpus safely and efficiently in the cloud. In *Proceedings of the Sixteenth European Conference on Computer Systems (EuroSys'21)*, 2021.
 - [75] Wikipedia. Radial basic function kernel. https://en.wikipedia.org/wiki/Radial_basis_function_kernel, 2021.
 - [76] Wikipedia. Rational quadratic covariance function. https://en.wikipedia.org/wiki/Rational_quadratic_covariance_function, 2021.
 - [77] Wikipedia. Taxicab geometry: Manhattan distance. https://en.wikipedia.org/wiki/Taxicab_geometry, 2021.
 - [78] Amir Yazdanbakhsh, Christof Angermüller, Berkin Akin, Yanqi Zhou, Albin Jones, Milad Hashemi, Kevin Swersky, Satrajit Chatterjee, Ravi Narayanaswami, and James Laudon. Apollo: Transferable architecture exploration. *CoRR*, abs/2102.01723, 2021.
 - [79] Di Zhang, Dong Dai, Youbiao He, Forrest Sheng Bao, and Bing Xie. RLScheduler: An Automated HPC Batch Job Scheduler Using Reinforcement Learning. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC'20)*, Virtual Event, November 2020.
 - [80] Yanqi Zhang, Weizhe Hua, Zhuangzhou Zhou, G. Edward Suh, and Christina Delimitrou. Sinan: ML-based and qos-aware resource management for cloud microservices. In *Proceedings of the 26th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS'21)*, Virtual, USA, 2021.
 - [81] Da Zheng, Disa Mhembere, Randal Burns, Joshua Vogelstein, Carey E. Priebe, and Alexander S. Szalay. FlashGraph: Processing Billion-Node Graphs on an Array of Commodity SSDs. In *Proceedings of 13th USENIX Conference on File and Storage Technologies (FAST'15)*, Santa Clara, CA, 2015.
 - [82] Qing Zheng, Kai Ren, Garth Gibson, Bradley W. Settlemyer, and Gary Grider. Deltafs: Exascale file systems scale better without dedicated servers. In *Proceedings of the 10th Parallel Data Storage Workshop (PDSW'15)*, page 1–6, Austin, Texas, 2015.
 - [83] Zhiheng Zhong, Minxian Xu, Maria Alejandra Rodriguez, Chengzhong Xu, and Rajkumar Buyya. Machine learning-based orchestration of containers: A taxonomy and future directions. *CoRR*, abs/2106.12739, 2021.
 - [84] Giulio Zhou and Martin Maas. Learning on distributed traces for data center storage systems. In *Proceedings of the Machine Learning and Systems (MLSys'21)*, Austin, TX, March 2021.
 - [85] ZNS SSDs Just Got Real - Ultrastar DC ZN540 Now Sampling. <https://blog.westerndigital.com/zns-ssd-ultrastar-dc-zn540-sampling/>.