EPJ Web of Conferences **295**, 09012 (2024) https://doi.org/10.1051/epjconf/202429509012
CHEP 2023

Digital Twins for Storage Systems and RAID Pools: Enhancing Data Management in High Energy Physics

Abdalaziz Al-Maeeni[1,*], *Aziz* Temirkhanov[1], *Artem* Ryzhikov[1], *Dmitry* Maevskiy[1], *Fedor* Ratnikov[1], and *Mikhail* Hushchyn [1]

[1]HSE University, Moscow, Russia

Abstract. In high-energy physics experiments, the acquisition of vast quantities of data poses formidable challenges in terms of efficient storage and management. We explore the application of the digital twin concept for SSD RAID pools, wherein digital replicas of physical systems are created to augment the efficiency of data storage within HEP experiments. By developing digital twins of storage systems, this research seeks to facilitate the continuous monitoring, comprehensive analysis, and strategic optimization of various workloads within the HEP domain. The key objectives of this study include the development of a digital twin for a data storage system and the formulation of generative models to assess the performance of the data storage system performance under specific configurations and data load parameters.

1 Introduction

Within data storage systems (DSS), where constituents like storage pools, controllers, and fast cache memory are present, the storage of data involves a combination of hard disk drives (HDDs) and solid-state drives (SSDs), organized through RAID configurations. These systems are essential for various applications, from business data management to scientific computing.

One significant challenge in optimizing DSS performance, particularly for SSDs, has prompted extensive research. Early studies, like the one by Huang et al. [1], recognized the fundamental differences between SSDs and HDDs. SSDs, known for their minimal latency, deliberate write operations, and costly block-level erasures, require unique modeling approaches. This led to the development of a black-box approach, which advocated for a simplified, information-independent way of modeling SSD performance. Surprisingly, despite its minimal reliance on prior knowledge of the device, this approach showed impressive accuracy when applied to SSDs, in contrast to its performance with HDD modeling.

In a related initiative, Kim et al. [2] developed SSDcheck, a model for assessing SSD performance. This innovative tool not only uncovers the complex processes within SSDs but is also proficient in forecasting the latency for subsequent interactions with standard SSDs. Such progress has enabled the practical implementation of SSD performance modeling across diverse applications. The application scenarios proposed take advantage of this model, along with additional functionalities, to realize a surge in productivity of as much as 130% when measured against the standard baseline.

*e-mail: al-maeeni@hse.ru

Furthermore, Li et al. [3] enhanced the black-box framework through the incorporation of regression trees. This strategy, particularly tailored for solid-state drives (SSDs), demonstrated a precise prognostication of the performance parameters. The mean relative error of the metrics is 20%, 13%, and 6% for latency, bandwidth, and throughput, respectively.

Moreover, machine learning techniques have gained popularity in the analysis of storage devices, as shown by Tarihi et al. [4]. They have trained SSD performance models capable of predicting response times to user requests using cost-effective and history-based features. These machine learning-based methods have simplified the creation of more accessible and adaptable SSD simulators for real-world use.

However, despite these valuable contributions, existing solutions have limitations. Most notably, these models require resource-intensive development because they simulate physical processes and firmware stacks within devices. In addition, they may not consider vendor-specific details due to proprietary constraints, leading to discrepancies between simulated and real-world performance.

In response to these challenges, this work introduces a novel data-driven approach to SSD modeling. By using accurate performance measurements, we address these limitations and aim to provide a more accurate and comprehensive framework for modeling and optimizing SSD performance, transcending the constraints of device-specific approaches.

A digital twin is a virtual representation or simulation of a physical object, system, or process. It uses real-time data and other information sources to create a dynamic digital replica or model that accurately mimics and predicts the behavior of the real-world counterpart. Digital twins are used in various industries, including manufacturing, automotive, aerospace, and infrastructure, to optimize processes, monitor performance, and discover potential issues before they occur in the physical system. By integrating data analytics, artificial intelligence, and IoT technologies, digital twins enable businesses to improve decision-making, enhance efficiency, and reduce operational costs.

We utilize the concept of a digital twin to model storage systems, since in essence, it represents a model and provides a link between the actual storage pool and the model. In this research, we mainly focus on modeling without emphasizing the link. As a digital twin, our model can be updated online using new measurements from the system.

Our research delves into simulating key elements of data storage infrastructures, with an emphasis on Solid-State Drives (SSD) storage pools. The central objective is to develop predictive models adept at forecasting the performance of these components when subjected to diverse configurations and data volume scenarios. We quantify this performance by measuring critical metrics, namely the Input and Output Operations Per Second (IOPS) and the corresponding average latencies. These metrics serve as benchmarks to gauge the efficiency and responsiveness of SSD storage pools in various operational contexts.

The parameters defining the data load in our research are detailed in Table 1. We consider two primary load types: random and sequential, each encompassing a mix of read and write operations. These load types mirror the practical complexities of data storage in high-energy physics experiments. Additionally, we delve into the configuration parameters of SSD storage pools, including the total number of disks and the specific RAID scheme utilized. The RAID scheme is further characterized by the number of data and parity blocks it employs.

 https://doi.org/10.1051/epjconf/202429509012

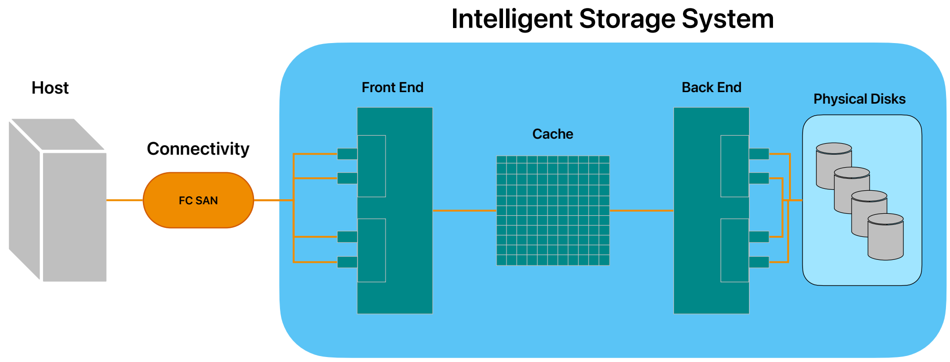

Figure 1: Development of a digital twin of a data storage system (DSS) to predict its performance for a given configuration and data load parameters.

2 Data

The datasets were gathered for the SSD pool under both random and sequential data loads. We harnessed Perf [5], a performance analysis tool, to create a diverse set of 512 distinct data loads. In our data collection process, we considered the pool configuration parameters and data load parameters. Each data load persisted for a duration of 120 seconds, within which we conducted separate measurements of Input-Output Operations Per Second (IOPS) and average latency for both read and write operations on a per second basis.

To ensure the comprehensive coverage of various pool configurations, we employed the Sobol sequence technique; see figure 2, resulting in a total of 512 unique data loads. Each of these loads also maintained a 120-second duration, during which we recorded IOPS and average latency for read and write operations on a per-second basis. This dataset serves as a foundation for our subsequent modeling and analysis efforts.

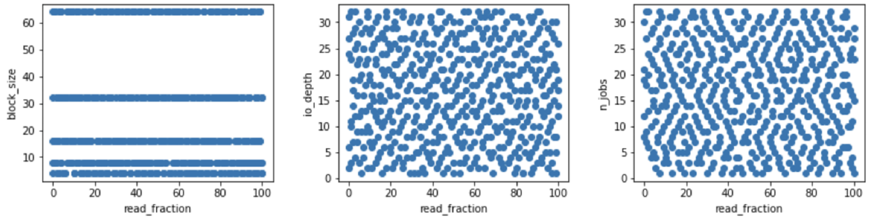

Figure 2: Efficient data collection using Sobol sampling of the data load parameters for SSD disks. Note the even coverage of the parameters space

Parameter	Random	Sequential
Block size	4, 8, 16, 32, 64 KB	128, 256, 512, 1024 KB
Read fraction	0 - 100%	0%
Number of jobs	1 - 32	1 - 20
Queue depth	1 - 32	1 - 32
RAID (K+M)	1+1, 2+1, 2+2, 4+1, 4+2, 8+2	
Number of disks	K+2M, 24, +3 values in between	

Table 1: Data load parameters and their value ranges for the storage SSD pool data set. For sequential and random data loads. We generated 512 different data loads using the Sobol sequence, each load ran for 120 seconds, during which we measured IOPS and average latency for read and write operations.

3 Methods

In this investigation, we apply the CatBoost regression technique as a parametric generative model to discern the intricate connections between the Input and Output Operations Per Second (IOPS) and the delays experienced across varied data volumes. This assessment utilizes the principle of Little's law to shed light on these correlations.

Little's law is a pivotal theorem that interconnects key performance indicators within a system. The law is mathematically represented as:

$$Q \times J = \mathrm{IOPS}_{\mathrm{read}} \times \mathrm{Latency}_{\mathrm{read}} + \mathrm{IOPS}_{\mathrm{write}} \times \mathrm{Latency}_{\mathrm{write}}$$

Where Q signifies the queue's depth, and J is the quantity of jobs. Crucially, Little's law proposes a specific relationship: the logarithm of IOPS is directly linked to the negative logarithm of Latency. This concept forms the bedrock of our methodical approach.

Recognizing the inherently unpredictable nature of IOPS and delays in practical systems is crucial. To address this randomness, our methodology incorporates a probabilistic standpoint. We estimate the distributions of the logarithmic measures of IOPS and latencies by utilizing conditional two-dimensional Gaussian distributions:

$$\hat{z}_i = \log(\hat{y}_i), \quad \hat{z}_i \sim \mathcal{N}(\hat{\mu}(x_i), \hat{\Sigma}(x_i))$$

In this formula, $\hat{y}_i$ denotes a vector containing the estimated figures for IOPS and latency, while $\hat{\mu}(x_i)$ and $\hat{\Sigma}(x_i)$ represent the estimated average and the covariance matrix conditioned on the input vector x_i. This vector x_i reflects the data volume and configuration settings. The CatBoost regression model's role is to prognosticate $\hat{\mu}(x_i)$ and $\hat{\Sigma}(x_i)$.

Next, we determine the mean vectors μ_j and covariance matrices Σ_j corresponding to each specific data load under scrutiny. To guarantee the predicted covariance matrices Σ_j are positive semi-definite, Cholesky decomposition is utilized [6]. This method enables us to articulate the inverse of Σ_j as the multiplication of a lower triangular matrix L_j and its transpose L_j^T.

We calibrate our CatBoost regression model using a MultiRMSE loss function, which is articulated as:

$$L^2 = \frac{1}{2m} \sum_{j=1}^{2m} (\hat{\mu}(x_j) - \mu_j)^2 + \frac{1}{2m} \sum_{j=1}^{2m} (\hat{L}(x_j) - L_j)^2$$

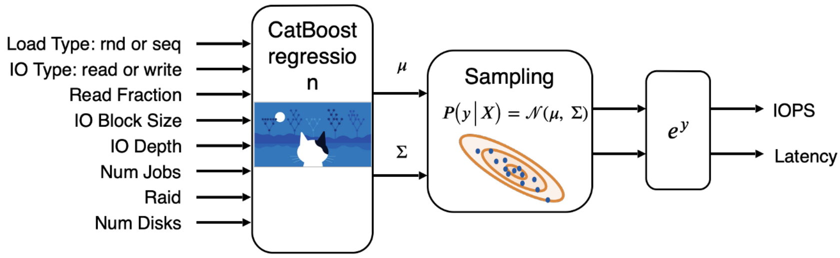

Figure 3: CatBoost model for performance predictions of the storage pools and cache for the given values of data load and configuration parameters.

This loss function helps to refine the predictions by minimizing the discrepancy between the predicted and actual values of both the mean and the transformation of the covariance matrix. Here, m represents the number of samples in our study, and $\hat{\Sigma}^{-1}(x_i) = \hat{L}(x_j)\hat{L}(x_j)^T$. Our model comprises 5000 decision trees, and we estimate optimal hyperparameter values through a grid search process tailored to each sample in our dataset. We optimize the following hyperparameters: the depth of the trees, with values ranging from 2 to 8, and the learning rate, within the parameters of 0.01, 0.05, and 0.1.

3.1 Quality Metric

To assess the quality of our predictive models, we employ the Mean/Median Absolute Percentage Error (MEAPE) metric, which quantifies the models' ability to predict the mean values of IOPS and latency for each data load, providing a valuable measure of predictive accuracy and reliability.

MEAPE is defined as:

$$\text{MEAPE} = \left|\frac{\hat{\mu} - \mu}{\mu}\right| \times 100\%$$

Where, $\hat{\mu}$ represents the predicted mean value of IOPS or latency, and μ is the true mean value, which is calculated as follows:

$$\mu = \begin{cases} \frac{1}{k}\sum_{i=1}^{k} y_i & \text{for mean calculation} \\ \text{median}(y) & \text{for median calculation} \end{cases}$$

MEAPE assesses the relative absolute error between the predicted mean values of our model and the true mean values of IOPS and latency for various data load scenarios. This metric offers insight into the accuracy of our predictive models, providing a basis for evaluating their performance in capturing the statistical characteristics of the data. By incorporating MEAPE into our analysis, we enhance our understanding of the model's predictive capabilities and their ability to provide reliable estimates of system performance.

4 Results

We evaluated the quality of our models. Our modeling approach involved fitting all previously described models to the same training samples, generating predictions on identical test samples, and subsequently assessing their quality using the MEAPE metric.

The table below presents the MEAPE values for the CatBoost model when applied to the SSD pool under random data loads. Figures 4, 5, and 6 provide an illustrative example of these predictions for a specific data load from the test sample. Our model achieves mean prediction errors of 6.7% and 8.3% for IOPS and latency, respectively.

		Mean MEAPE %	**Median MEAP %**
SSD Pool, seq	IOPS	5.6	5.3
	Latency	12.2	6.9
SSD Pool, rand	IOPS	6.7	4.3
	Latency	8.3	4.1

Table 2: We outline the MEAPE values for the SSD pool under sequential data loads, yielding mean prediction errors of 5.6% and 12.2% for IOPS and latency, respectively. Figures 4 and 5 offer a visual representation of these predictions for a specific data load from the test sample.

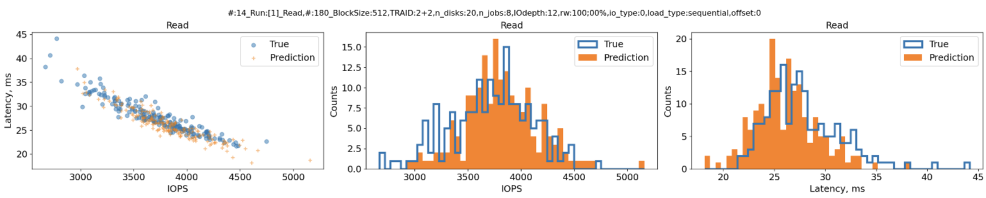

Figure 4: Block size: 512, RAID: 2+2, 15 disks, 6 jobs, IO depth: 1, Read fraction: 1%, IO type: Read, Load type: Sequential.

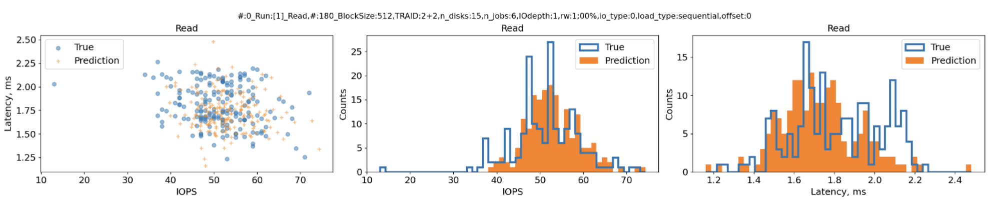

Figure 5: Block size: 256, RAID: 1+1, 8 disks, 15 jobs, IO depth: 12, Read fraction: 0%, IO type: Write, Load type: Sequential.

Figures 4, 5 and 6 show predictions under different load configurations, we observe the relation mentioned above $IOPS \propto \dfrac{1}{Latency}$ with different degrees of proportionality between conditions.

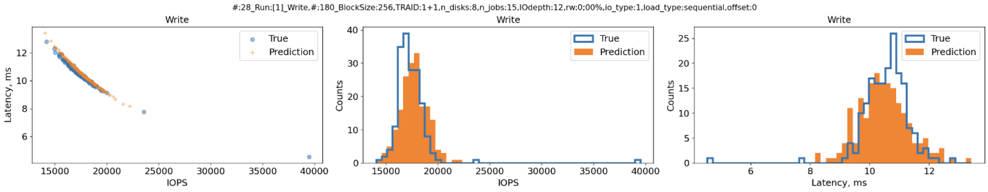

Figure 6: Block size: 16, RAID: 4+1, 20 disks, 2 jobs, IO depth: 25, Read fraction: 77%, IO type: Read, Load type: Random.

5 Conclusions

The digital twins approach used in this study presents a promising avenue for enhancing SSD design and optimizing performance. Leveraging both parametric models, this approach enables the accurate prediction of mean IOPS and latency values, contributing to the overall improvement of data storage systems. Furthermore, our approach allows for identifying inefficiencies within SSD disk storage systems, thereby facilitating cost reduction through enhanced resource management and configuration adjustments. It also plays a role in ensuring the reliability and scalability of storage system disks by enabling informed decision-making processes. Little's law, prediction reliability, helps identify non-realistic forecasts. The dataset collected in this study, encompassing IOPS and latency metrics for SSD drives, constitutes a treasure trove for future research. These data underpin a wide array of potential explorations, such as the development of data-centric performance models, the creation of conditional generative models, the evaluation of uncertainty, and the assessment of model dependability. These investigations are pivotal for advancing our comprehension of SSD performance dynamics and enhancing the predictive accuracy of storage system behaviors under varying conditions.

6 Acknowledgments

Research leading to these results has received funding from the Basic Research Program at the National Research University Higher School of Economics.

This research was supported in part through computational resources of HPC facilities at HSE University [7]

References

[1] H.H. Huang, S. Li, A. Szalay, A. Terzis, *Performance Modeling and Analysis of Flash-Based Storage Devices*, in *2011 IEEE 27th Symposium on Mass Storage Systems and Technologies (MSST)* (2011), pp. 1–11, ISSN 2160-1968

[2] J. Kim, K. Choi, W. Lee, J. Kim, **17**, 14:1 (2021)

[3] S. Li, H.H. Huang, *Black-Box Performance Modeling for Solid-State Drives*, in *2010 IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems* (IEEE, 2010), pp. 391–393, ISBN 978-1-4244-8181-1, `http://ieeexplore.ieee.org/document/5581555/`

[4] M. Tarihi, S. Azadvar, A. Tavakkol, H. Asadi, H. Sarbazi-Azad, **10**, 1821 (2022)

[5] *Perf: Linux profiling with performance counters*, `https://perf.wiki.kernel.org/index.php/Main_Page`, accessed: 2022-01-12

[6] G.H. Golub, C.F. Van Loan, *Matrix Computations (3rd Ed.)* (Johns Hopkins University Press, USA, 1996), ISBN 0801854148

[7] P.S. Kostenetskiy, R.A. Chulkevich, V.I. Kozyrev, Journal of Physics: Conference Series **1740**, 012050 (2021)