

Enterprise Disk Drive Scrubbing Based on Mondrian Conformal Predictors

Rahul Vishwakarma

RAHULDEO.VISHWAKARMA01@STUDENT.CSULLB.EDU

California State University Long Beach, 1250 Bellflower Blvd, Long Beach, CA 90840, United States

Jinha Hwang

JINHA.HWANG01@STUDENT.CSULB.EDU

California State University Long Beach, 1250 Bellflower Blvd, Long Beach, CA 90840, United States

Soundouss Messoudi

SOUNDOUSS.MESSOUDI@HDS.UTC.FR

HEUDIASYC - UMR CNRS 7253, Université de Technologie de Compiègne, 57 avenue de Landshut, 60203 Compiègne Cedex - France

Ava Hedayatipour

AVA.HEDAYATIPOUR@CSULB.EDU

California State University Long Beach, 1250 Bellflower Blvd, Long Beach, CA 90840, United States

Editors: Harris Papadopoulos, Khuong An Nguyen, Henrik Boström and Lars Carlsson

Abstract

Disk scrubbing is a process aimed at resolving read errors on disks by reading data from the disk. However, scrubbing the entire storage array at once can adversely impact system performance, particularly during periods of high input/output operations. Additionally, the continuous reading of data from disks when scrubbing can result in wear and tear, especially on larger capacity disks, due to the significant time and energy consumption involved. To address these issues, we propose a selective disk scrubbing method that enhances the overall reliability and power efficiency in data centers. Our method employs a Machine Learning model based on Mondrian Conformal prediction to identify specific disks for scrubbing, by proactively predicting the health status of each disk in the storage pool, forecasting n -days in advance, and using an open-source dataset. For disks predicted as non-healthy, we mark them for replacement without further action. For healthy drives, we create a set and quantify their relative health across the entire storage pool based on the predictor’s confidence. This enables us to prioritize selective scrubbing for drives with established scrubbing frequency based on the scrub cycle. The method we propose provides an efficient and dependable solution for managing enterprise disk drives. By scrubbing just 22.7% of the total storage disks, we can achieve optimized energy consumption and reduce the carbon footprint of the data center.

Keywords: Mondrian conformal prediction, calibration, disk scrubbing, storage array.

1. Introduction

A large-scale data center is a complex ecosystem primarily consisting of various types of storage devices, such as hard disk drives (HDD), solid-state drives (SSD), and hybrid storage devices, spread over multiple geographic locations. As the number of storage components grows in the storage ecosystem, it becomes increasingly difficult to manage its business continuity (Bajgorić et al., 2022) because of the underlying complexity and uncertainty associated with each individual component’s internal working mechanics. For instance, a data center can have a mixed workload of transactional databases, network-attached storage (NAS), and archival storage. Each workload has its unique characteristics and impact on the

reliability and remaining useful life of the storage system. Inadequate management of these storage components can result in data loss and eventually business disruption (Pinheiro et al., 2007). Monitoring the health of storage components is a common proactive approach to maintaining the reliability of storage systems. This is typically achieved through system logs and SMART logs of HDDs and SSDs (Zhang et al., 2023).

In the context of data centers, many device manufacturers and vendors (Ma et al., 2015; Vishwakarma and Perneti, 2021) utilize drive failure analysis as a metric to assess the overall reliability of the data center. Drive failure analysis typically categorizes failures into two types: complete failure and latent failure of the disk drive. Complete failure is relatively easy to detect and enables a transparent approach for replacing failed disks. However, latent failure is nearly undetectable and can lead to sudden failure and loss of critical data (Bairavasundaram et al., 2007; Schroeder et al., 2010).

Traditional system-level approaches, such as storage virtualization technology like Redundant Array of Independent Disks (RAID), tend to focus on passive fault tolerance rather than proactive approaches. Alternatively, many statistical (Hamerly et al., 2001; Singh and Vishwakarma, 2023) and machine learning (Pitakrat et al., 2013; Sun et al., 2019) approaches have been explored to enhance the reliability of the storage system. Machine learning approaches, while powerful, face challenges when it comes to making accurate predictions on unseen data, as production data is subject to change over time (Lu et al., 2018), requiring continuous model updates. One argument to consider is that even a small false positive rate (FPR), such as 0.1%, can be a concern when implementing the model in real-world data centers that may have millions of disk drives. Instead of using drive failure analysis for reliability enhancement, an alternative approach could be to leverage disk drive scrubbing (Iliadis et al., 2008, 2011) in a more fine-grained manner to identify specific disks that require further attention.

Disk scrubbing is a process of performing full media pack sweeps across allocated and unallocated disks to detect and rebuild latent medium errors (Ryu and Park, 2009), reducing the chances of bad block media detection during host I/O activity. However, running scrubbing tasks for the entire disk population in an array can significantly increase the load on the data storage system, potentially degrading its performance. Additionally, if the disk has a larger capacity (e.g., 12TB), it will take a considerable amount of time to complete the operation.

Our proposed method aims to optimize disk scrubbing by selectively targeting only the disks that require the operation. We use a learning framework based on Mondrian conformal prediction, which is agnostic to the specific machine learning algorithm, to identify specific disks for scrubbing. The method involves forecasting the health of a disk n -days ahead through binary classification. A set of healthy drives is created, and the health status of the entire storage pool is quantified based on the predictor’s confidence. Drives marked as unhealthy are then sent for treatment based on the administrator’s decision. The metrics obtained from Mondrian conformal prediction are used to prioritize selective scrubbing for the drives. This proposed method has two main advantages: first, it leverages the underlying drive failure analysis, and second, the quantified output from the forecast engine can be used as input for the disk scrubbing scheduler engine, optimizing the scrubbing process and enhancing reliability in the data center environment. A summary of our contributions is as below:

- Introducing a fine-grained approach for identifying disk drives to be scrubbed by complementing the existing failure analysis engine on the storage sub-system with an algorithm-agnostic framework based on Mondrian conformal prediction.
- Translating the output of the framework, i.e. the confidence of prediction for new data points, into a ranking mechanism that sorts the healthy drives in descending order for further treatment by the decision engine.
- Implementing a scrubbing frequency schedule for the entire storage array based on n-step ahead system load prediction using probabilistic weighted fuzzy time series, which is mapped to the scrubbing engine for optimized disk scrubbing.

The method contributes to a proactive approach that provides value for business continuity, such as resource and power savings in data centers during data scrubbing by selectively spinning disks down. This means that only the required disks are scrubbed on a priority basis, while the healthy disks can be scrubbed less frequently or not at all. This optimized approach helps improve operational efficiency and reduces unnecessary disk wear, resulting in potential cost savings and enhanced reliability for the data storage system.

This paper is structured as follows: [section 2](#) provides the motivation for our work and outline the design considerations. In [section 3](#), we present a review of related work in the field. We provide a definition and algorithmic overview of conformal prediction in [section 4](#). In [section 5](#), we deliver a synopsis of our proposed solution, followed by the presentation of experimental results in [section 6](#). The usability and interpretation of disk drive health metrics are discussed in [section 7](#), and we conclude with [section 8](#), summarizing our findings.

2. Motivation and design goals

In data centers, a significant number of unhealthy drives go undetected due to latent failure attributes, resulting in fail-stop scenarios. One common approach to mitigate such scenarios is disk scrubbing, which consists of verifying disk data through a background scanning process to identify bad sectors. However, this process can consume energy and cause performance degradation depending on the trigger schedule. This scenario raises concerns in the industry, especially as disk capacities increase. We notice a missing link in addressing 'which disk to scrub', 'when to scrub', based on frequency of scrub cycle while minimizing storage array performance impact and also maximizing the reliability. In this paper, we consider the following objectives and design approaches to tackle this challenge :

- **Which disk to scrub?** Depending on the specific scrubbing process, it can temporarily degrade the performance of the drive. To ensure that the drive remains fast and responsive, minimizing the frequency of scrubbing is crucial. Instead of performing scrubbing for all disks in the storage array, our approach focuses on selectively scrubbing only the disks that require it, thereby reducing the overall time required to complete the process.
- **When to scrub?** We can optimize the disk drive scrubbing schedule by considering factors such as the workload of the system, the importance of the data on the drive, and the availability of resources. This approach ensures that scrubbing is performed at the most appropriate times, minimizing the impact on the overall system performance.

3. Related Work

Storage device reliability has long been a critical concern in the industry, and existing solutions often rely on failure analysis of storage systems. However, traditional methods like accelerated life tests (Cho et al., 2015) have not proven to be reliable indicators of actual failure rates in production environments. Recent machine learning-based approaches, such as multivariate time-series (Yu, 2019) and time-series classification (Ircio et al., 2022), have focused on improving model accuracy, but often lack deep integration of domain knowledge. Moreover, the multi-modal approach by (Lu et al., 2020) using performance metrics (disk-level and server-level) and disk spatial location only focuses on fail-stop scenarios, which may not be helpful in detecting latent failures. A most recent study (Lu et al., 2023) has addressed this issue by investigating grey failures (fail-slow drives) using a regression model to pinpoint and analyze fail-slow failures at the granularity of individual drives.

Another important factor of disk scrubbing is the implementation cost and power consumption. (Mi et al., 2008) and (Jiang et al., 2019) address performance degradation due to scrubbing and propose assigning a lower priority to the background process during idle time, i.e. when the disk drive is not actively engaged in processing data or performing any other tasks. (Liu et al., 2010) and (Oprea and Juels, 2010) propose a method to mitigate power consumption and determine when to scrub in systems with inexpensive data but require designing another method to identify less critical data. Drive space management in case of replacing the failed disk is discussed in (Pâris et al., 2010), along with reducing the need for frequent scrubbing. A multilevel scrubbing is proposed in (Zhang et al., 2020) using a Long Short-Term Memory (LSTM) model to detect latent sector errors in a binary classification setup. However, using machine learning-based models may treat healthy and relatively less healthy disks the same, leading to unnecessary scrubbing of healthy disks.

To the best of our knowledge, our work is the first to adopt Mondrian conformal prediction for assigning a health score to each individual disk drive and using the metrics to design a scrubbing cycle aligned with the system idle time.

4. Conformal prediction

Conformal prediction (Shafer and Vovk, 2008) is a powerful framework in machine learning that allows for prediction with uncertainty. Unlike traditional point prediction methods used in classification tasks, conformal prediction provides a set prediction. This means that instead of outputting a single predicted label, conformal prediction provides a range of possible labels that are likely to be correct, along with confidence and credibility measures in the correctness of these predictions. This is particularly useful in cases where the prediction task may be uncertain or when the model is dealing with previously unseen data, especially when the application is of high risk (Luo et al., 2022).

Conformal prediction is not limited to classification tasks alone, but it is also valid for regression tasks by providing prediction intervals instead of a single-point prediction. These prediction intervals represent a range of possible values for the target variable, along with a measure of confidence in the correctness of these intervals. This allows for more nuanced and interpretable predictions in regression tasks, where the goal is to estimate a continuous value rather than a discrete class label.

One of the key advantages of conformal prediction is its agnostic nature, which means it can be used with any machine learning algorithm. This flexibility allows for the integration of conformal prediction into various machine learning pipelines without being constrained by the choice of a specific algorithm. This makes conformal prediction a versatile tool that can be applied to various tasks and domains (Messoudi et al., 2020).

A great resource for learning more about conformal prediction is the book "Algorithmic Learning in a Random World" (Vovk et al., 2022), which provides a detailed overview of the theory and applications of conformal prediction. Additionally, a repository of conformal prediction implementations and resources is maintained on GitHub¹(Manokhin, 2022), providing practical tools and examples for applying conformal prediction in various machine learning settings. This makes it easier for researchers and practitioners to implement and experiment with conformal prediction in their own work.

The overall approach of conformal prediction in the inductive setting is presented in [algorithm 1](#).

Algorithm 1: Inductive conformal prediction (ICP)

Input: Divide the training sequence Z_{tr} into two disjoint subsets; the proper training set Z_t and the calibration set $Z_c = \{(x_1, y_1), \dots, (x_q, y_q)\}$

Task: Predict $y_{n+1} \in \mathbf{Y}$ for any new object $x_{n+1} \in \mathbf{X}$.

Algorithm:

1. Split the original data set \mathbf{Z} into a proper training set \mathbf{Z}^{tr} with $|\mathbf{Z}^{tr}| = m$ and a calibration set \mathbf{Z}^{cal} with $|\mathbf{Z}^{cal}| = n - m = q$.
 2. Train using $h : \mathbf{X} \rightarrow \mathbf{Y}$ on \mathbf{Z}^{tr} and obtain the nonconformity measure $f(z)$. The non-conformity measure for h is defined as $f(z) = 1 - \hat{P}_h[y | x]$.
 3. Apply the non-conformity measure $f(z)$ to each example z_i of \mathbf{Z}^{cal} to get the nonconformity scores $\alpha_1, \dots, \alpha_q$.
 4. Choose a significance level $\epsilon \in (0, 1)$ to get a prediction set with a confidence level of $1 - \epsilon$
 5. For a new example x_{n+1} , compute a non-conformity score for each class $C_k \in \mathbf{Y}$: $\alpha_{n+1}^{C_k} = f((x_{n+1}, y = C_k))$
 6. For each class $C_k \in \mathbf{Y}$, compute the p -value : $p_{n+1}^{C_k} = \frac{|\{i \in \{1, \dots, q\} : \alpha_{n+1}^{C_k} \leq \alpha_i\}|}{q}$
 7. Build the prediction set: $\Gamma^\epsilon = \{C_k \in \mathbf{Y} : p_{n+1}^{C_k} > \epsilon\}$
-

4.1. Mondrian conformal prediction (MCP)

Mondrian conformal prediction (MCP) is a variant of the conformal prediction framework that provides a guarantee on a subset of the dataset, or on specific categories of the dataset. This variant is originally established for a classification problem by creating class-conditional

1. <https://github.com/valeman/awesome-conformal-prediction>

or attribute-conditional categories (Vovk et al., 2003). However, a Mondrian version exists for regressors (Boström and Johansson, 2020).

In imbalanced datasets, the minority class (i.e., the class with fewer instances) may lead to biased predictions and inaccurate confidence measures. Mondrian conformal prediction is a powerful tool to handle this issue by maintaining the same error rate for both the majority and minority classes, ensuring that the predictions are not biased toward the majority class.

When calculating the non-conformity scores in MCP, we only consider the scores related to the examples that share the same class as the object x_{n+1} , which we are testing hypothetically. Consequently, the p -value is substituted in step 6 of algorithm 1 and calculated as:

$$p_{n+1}^{C_k} = \frac{\left| \left\{ i \in 1, \dots, q : y_i = C_k, \alpha_{n+1}^{C_k} \leq \alpha_i \right\} \right|}{\left| \{ i \in 1, \dots, q : y_i = C_k \} \right|}.$$

Mondrian conformal prediction has been implemented for various domains in many real-life use cases for academia as well as industry. For instance, (Messoudi et al., 2021) apply MCP for tenant debt prediction in real estate, (Alvarsson et al., 2021) use it for modeling ABC transporters in drug discovery, and (Vishwakarma and Liu, 2021) leverage conformal predictors for detecting persistent storage failure analysis in the enterprise storage domain.

4.2. Evaluation metrics

To evaluate the performance of conformal prediction models, several metrics can be employed. In our study, we will focus only on two :

- **Confidence:** reflects the certainty of the model that a prediction is a singleton, or a unique outcome. Confidence is based on the concept of p-values, which are used to assess the probability of obtaining an outcome as extreme as the one observed, assuming that the null hypothesis is true. A higher confidence value suggests that the model is more confident about the accuracy of its prediction and that the predicted label is likely to be correct. Conversely, a lower confidence value implies that there may be alternative labels that are equally likely. This metric is defined as:

$$\text{Confidence}(x) = \sup\{1 - \epsilon : |\Gamma_\epsilon(x)| \leq 1\}.$$

- **Credibility** quantifies the likelihood that a sample comes from the training set, as determined by the minimal significance level that would result in an empty prediction region. In other words, credibility is expressed as the largest p-value, which serves as the lower bound for the value of the significance level ϵ that would result in an empty prediction. A higher credibility value indicates a higher likelihood that the sample is consistent with the training set, while a lower credibility value indicates a higher likelihood of the sample being inconsistent with the training set. This metric is defined as:

$$\text{Credibility}(x) = \inf\{\epsilon : |\Gamma_\epsilon(x)| = 0\}.$$

5. Mondrian conformal prediction for Disk Scrubbing: our approach

In contrast to the conventional studies mentioned above, we propose a novel approach for disk drive scrubbing based on Mondrian conformal prediction to quantitatively assess the health status of disk drives and use it as a metric for selecting drives for scrubbing. Figure 1 shows a high-level overview of the proposed method.

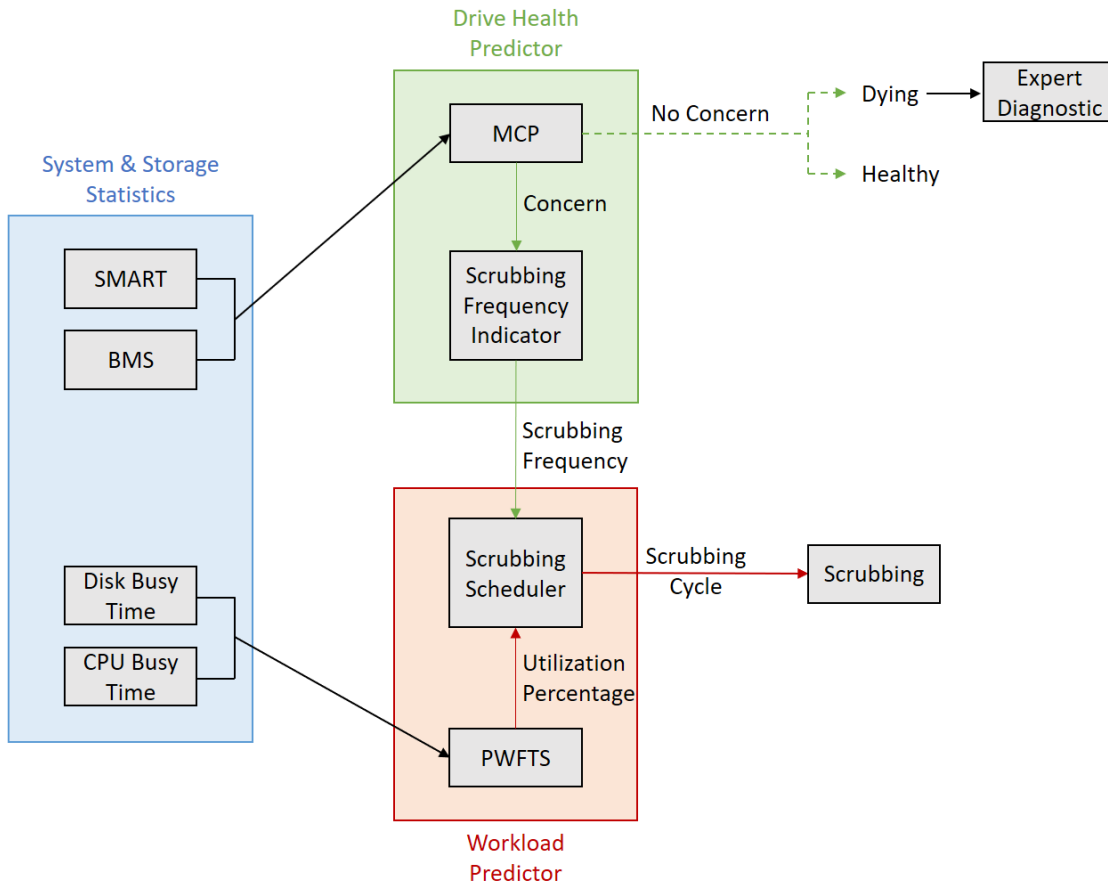


Figure 1: Overall approach of Mondrian conformal disk drive scrubbing.

The proposed architecture consists of three subsystems. The first subsystem is responsible for collecting storage and system statistics, which includes retrieving disk drive data from the storage array, as well as capturing CPU and disk busy statuses. The second subsystem, referred to as the drive health predictor engine, predicts the health status of the drives. It uses MCP to output a set of "No concern" drive disks, i.e. unhealthy/dying drives that can be flagged for manual diagnostics by experts (not discussed in this paper) or completely healthy drives that do not need any scrubbing, as well as a set of "Concern" disks with assigned health scores based on the predictor's confidence, which then are turned into scrubbing frequencies with the scrubbing frequency indicator. The underlying non-conformity score used is margin error function. The third subsystem is the workload predictor engine, which first predicts the resources' utilization percentage by taking into

account SAR logs², and then combine this result with the scrubbing frequencies in order to schedule when and how frequently disk drive scrubbing is performed. Finally, the scrubbing operation is triggered on the storage array based on the scrubbing cycle. In the following subsections, each component of the overall architecture is described in detail.

5.1. System and Storage statistics

The main components of this subsystem are:

- **SMART:** stands for Self-Monitoring, Analysis, and Reporting Technology, and refers to a set of predefined parameters provided by device manufacturers that offer insights into various aspects of a storage device’s performance, including temperature, error rates, reallocated sectors, and more. Each attribute has a threshold value assigned by the manufacturer, indicating the acceptable limit for that parameter. When a parameter exceeds its threshold value, it may indicate a potential issue with the storage device. We use SMART parameters as input features for the drive health predictor engine.
- **BMS:** stands for Background Media Scanning, and is a passive process that differs from disk scrubbing, which actively scans the disk for errors during idle periods without reading or writing data. BMS involves scanning the disk for errors in the background without interrupting normal operations. In our proposed architecture, we also extract this BMS feature, which is a numerical value for the number of times it encounters errors while performing a scan on the same drive, and feed it to the drive health predictor engine.
- **Disk and CPU busy time:** The performance of a drive is heavily dependent on its critical processes, such as data access and write speed. The numeric values range between 1 to 100 in terms of percentage and change over time with a sampling period of 1 hour. These system statistics are extracted from the SAR logs (standard logs for system utilization) and converted into time series data, which can then be used by the workload predictor engine.

5.2. Which disk to scrub: Drive health predictor

In a normal data center setting, all disk drives are classified as either healthy or unhealthy. Unhealthy disks are supposed to be dying or imminently failing, thus they are not marked for scrubbing, while healthy disks are marked for scrubbing.

In our approach, we propose to assign a relative ‘degree of health’ score to each disk. Drives that are marked as of *No concern* are either dying/imminently failing or completely healthy, while those marked as of *Concern* have different degrees of health other than failing or healthy. The conformal prediction framework then classifies the “No-concern” and “Concern” drives, and only selects the disks which are in the set of “Concern” drives

2. The System Activity Report is a command that provides information about different aspects of system performance. For example, data on CPU usage, memory and paging, interrupts, device workload, network activity, and swap space utilization

for further ranking. These are the drives which are concerning to us and is used as input for the scrubbing scheduler.

Our focus, as shown in Figure 2, is on identifying disks in the system that are currently of concern or may become concerning soon, and only selecting those disks for scrubbing. This approach reduces the number of disks meant for scrubbing, since even completely healthy drives are not scrubbed, making the process more efficient and targeted. By doing so, we optimize time, power, and energy consumption and reduce the carbon footprint of data centers.



Figure 2: Quantifying the health of disk drives: The disks which are healthy and non-healthy are not selected for scrubbing, while the disks of concern are marked for scrubbing.

When dealing with disk drives in a usual data center environment, failures are rare over a period of time, resulting in a highly imbalanced dataset with a small number of failed disks and the majority of disks being healthy. To handle this imbalanced data, we adopt a Mondrian Conformal Prediction approach, in order to get the prediction labels "0": failed and "1": healthy, along with their confidence score that serves as a health score in our case. This means that our MCP algorithm selects disks with a confidence score depending on the threshold chosen by the administrator.

For instance, if the administrator sets a threshold of 1%, this will lead to excluding disks with health scores above 99% as healthy or failing (depending on the label) and only selecting disks with a health score lower than 99% for scrubbing. Furthermore, the selected drives can be mapped to distinct scrubbing frequencies. Thus, drives with poor health scores may require more frequent scrubbing (every week), while those with good health scores will need less frequent scrubbing (every 3 months). For the same threshold of 1%, the administrator can then map the disk health with a scrubbing frequency, as in Table 1.

Disk health	Scrubbing frequency	Health score
Best	LOW (CYCLE - A)	[95%, 99[
Medium	MEDIUM (CYCLE - B)	[80%, 95[
Poor	HIGH (CYCLE - C)	< 80%

Table 1: Mapping of the disk health with the scrubbing frequency based on health score.

5.3. When to scrub: Workload predictor

After identifying the disks to be scrubbed using the drive health predictor engine, the next step is to determine the optimal time to perform scrubbing using the workload predictor. This component needs to consider the availability of system resources, i.e. disk and CPU utilization information in the system and storage statistics subsystem.

The workload predictor employs a Probabilistically Weighted Fuzzy Time Series algorithm (PWFTS), as detailed in (Orang et al., 2020). This algorithm forecasts n-days ahead system utilization, by predicting the system utilization percentage for the next 12 hours, with 1-hour intervals. Then, this information is combined with one of the three possible scrubbing cycles (A, B, or C as in Table 1) obtained from the drive health predictor. Finally, the scrubbing is triggered. During the 1-hour interval, if the scrubbing is complete, then we stop, if not, the administrator is notified. The high-level flowchart for the system workload predictor is outlined in Figure 3.

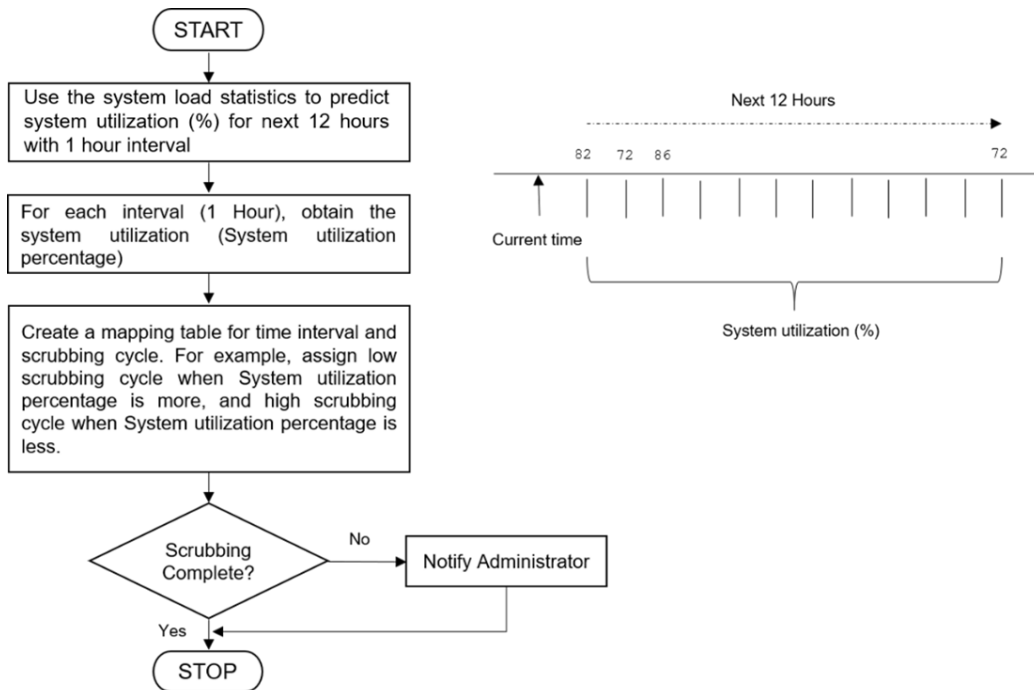


Figure 3: Flowchart of the workload predictor using the PWFTS algorithm.

In Figure 4, we showcase the n-days ahead forecasting of the system utilization percentage. It is evident from the figure that the system exhibits a lower load on day 0 and a higher load on day 2. Consequently, scheduling the scrubbing operations at day 0, when the system is under a lower load, would be more favorable. This approach optimizes the utilization of system resources, ensuring efficient scrubbing of the disks, and leading to lower processing time, lower energy consumption, and a reduced carbon footprint of the data center.

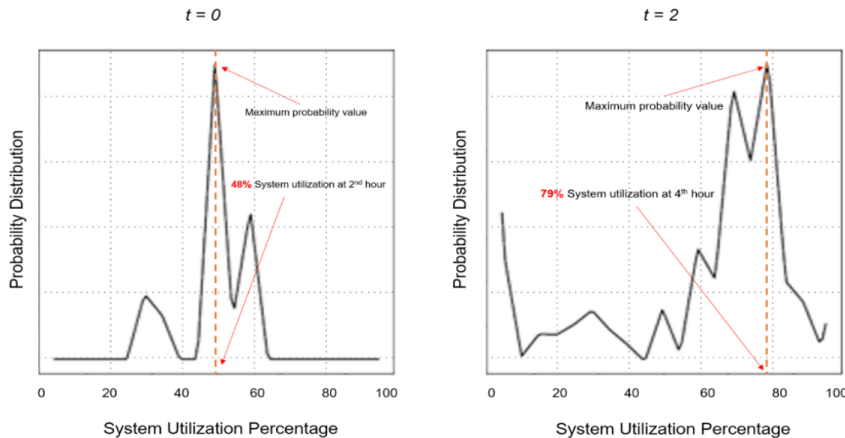


Figure 4: Probability distribution of system utilization percentage for n-days ahead forecasting.

6. Experimental setting

In this section, we detail the dataset used for our study and the conducted experiments as well as their results.

6.1. Open-source Baidu dataset

This dataset (DrTycoon, 2023) consists of samples collected from Seagate ST31000524NS enterprise-level HDDs, with a total of 23395 units and 13 features describing SMART attributes as shown in Table 2. The labeling of each disk was based on its operational status, categorized as either functional or failed. A significant proportion of disks, totaling 22962, were classified as functional, while a smaller subset of 433 was marked as failed, resulting in an imbalanced dataset. The SMART attribute values were recorded at an hourly interval for each disk, generating 168 samples per week for operational disks which gives 1,048,573 actual rows in the dataset corresponding to 23,395 disks (sampling frequency of 1 hour over a period of 2 years). The number of rows represents only the sample of operational disks that are provided in the dataset. However, the failed disks had varying numbers of samples, up to 20 days prior to failure.

Column No	Feature	Description
1	Index of the disk serial number	Ranging from 1 to 23395
2	Value of SMART ID #1	Raw Read Error Rate
3	Value of SMART ID #3	Spin Up Time
4	Value of SMART ID #5	Reallocated Sectors Count
5	Value of SMART ID #7	Seek Error Rate
6	Value of SMART ID #9	Power On Hours
7	Value of SMART ID #187	Reported Uncorrectable Errors
8	Value of SMART ID #189	High Fly Writes
9	Value of SMART ID #194	Temperature Celsius
10	Value of SMART ID #195	Hardware ECC Recovered
11	Value of SMART ID #197	Current Pending Sector Count
12	Raw Value of SMART ID #5	Reallocated Sectors Count
13	Raw Value of SMART ID #197	Current Pending Sector Count
14	Class label of the disk	0 for failed and 1 for functional

Table 2: Features’ description for the Open-source Baidu dataset.

6.2. Experimental results

For our experiments, we employed the Python programming language and used the MAPIE³ library ([map](#)) for implementing Mondrian Conformal Prediction. The underlying algorithm employed in our experiments was the k Nearest Neighbors (kNN) algorithm.

The main goal of conducting the experimental evaluation is to showcase the significant reduction in the number of disk drives to be scrubbed that can be achieved by using the drive health predictor engine, i.e. exploiting the Mondrian conformal predictor.

[Table 3](#) shows a comparison between the confusion matrix for the drive disk classification problem using the underlying algorithm alone kNN and adding Mondrian Conformal Prediction, where label "0" indicates a disk failure and label "1" indicates a functional one. We can notice that, adding MCP, the number of disks correctly classified as failing has increased from 51314 to 51669, i.e., a difference of 355. This shows MCP helps to identify more disks of the minority class, but with a drawback that is a decrease in the number of disks correctly classified as healthy which has reduced from 296689 to 268616, i.e., a difference of 28073.

This issue can be solved by considering the confidence scores and their respective health status, as shown in [Figure 5](#). There are nearly 126,224 drives with a health score greater than 99.95% for the disks labeled as healthy (left), out of total 349,525 disks, but when considering the relative health score, we categorize the 79,396 disk drives with a health score less than 99.9% as less healthy. Consequently, as shown in [Table 4](#), we only select these 79,396 disk drives for scrubbing and skip the remaining 270,129. This approach significantly reduces the number of disks to be scrubbed to only 22.7%, resulting in lower power and energy consumption, which is noteworthy.

3. <https://github.com/adamzenith/MAPIE/tree/Mondrian>

		kNN		MCP		
		Predicted	0	1	0	1
Actual	0	51314	547	51669	537	
	1	975	296689	28703	268616	

Table 3: Comparison of confusion matrix results for disk drive classification using kNN and MCP.

Heath score	0.998 - 0.9985	0.9985 - 0.999	0.999 - 0.9995	0.9995 and 1
Disk count	16468	62928	63087	126244

Table 4: The number of relatively healthy drives based on the health score intervals

7. Discussion

The proposed method for disk identification for scrubbing offers a dual benefit. Firstly, it can be utilized to assess the reliability of the storage system. Secondly, it employs a disk ranking mechanism to assign relative health scores to individual disks. The choice of classification algorithm depends on factors such as dataset size and available compute resources. However, the decision can be guided by the expertise of the system administrator.

In addition, we discuss how the use of the Mondrian conformal predictor can aid in identifying latent failures of disks, which could be a potential area for future research. Furthermore, we identify three key aspects for designing optimal scheduling and cover performance metrics, including effective coverage and size of the average prediction set.

Lastly, we provide a hypothetical evaluation of energy and power savings resulting from selective scrubbing. This showcases the potential benefits of the proposed method in terms of reduced power and energy consumption, highlighting its effectiveness in optimizing disk scrubbing operations.

7.1. Optimal scheduling aspect

With respect to disk scrubbing frequency scheduling, we can design three aspects of scheduling: time window, frequency, and space allocation. Each of them is described below:

- Time window focuses on scheduling the time window for scrubbing based on the workload pattern. Scrubbing is done when the system is predicted to be idle.
- Frequency involves scheduling the frequency of scrubbing based on the health status of the drive. For drives with the best health, scrubbing is done less frequently. For drives with medium health, scrubbing is done more frequently.
- Space deals with scheduling space allocation based on the spatial and temporal locality of sector errors. Instant scrubbing is performed on problematic chunks to ensure efficient disk scrubbing.

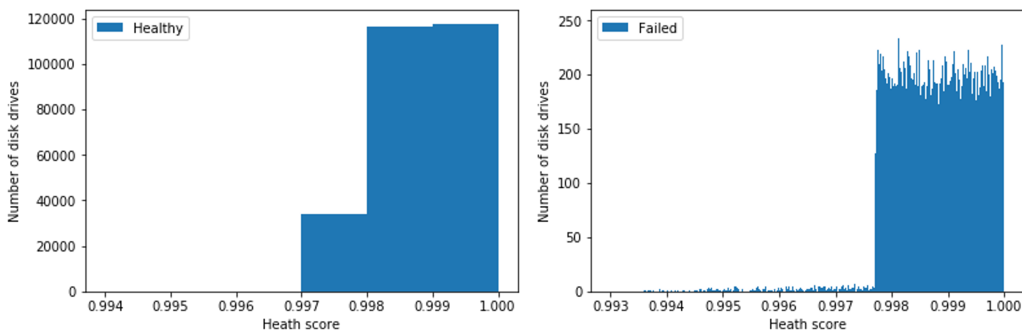


Figure 5: Histogram of the health scores corresponding to healthy predictions (left) and faulting predictions (right).

7.2. Performance metrics

We captured the effective coverage (i.e., for any chosen confidence level, prediction intervals will fail to include the correct label) and prediction set size for the open source dataset in . The plot in Figure 6 demonstrates that there is a positive correlation between the confidence level and the coverage. The split-conformal method results in a higher mean coverage than the cross-validation method, indicating that the calibration set selection has a considerable influence on the effective coverage. Furthermore, the right side of the figure displays the average size of the prediction set, which increases as the confidence level increases. Similarly, the split-conformal method yields a consistently higher mean prediction set size than the cross-validation method. The metrics can be used to evaluate how well the Mondrian conformal predictor is performing.

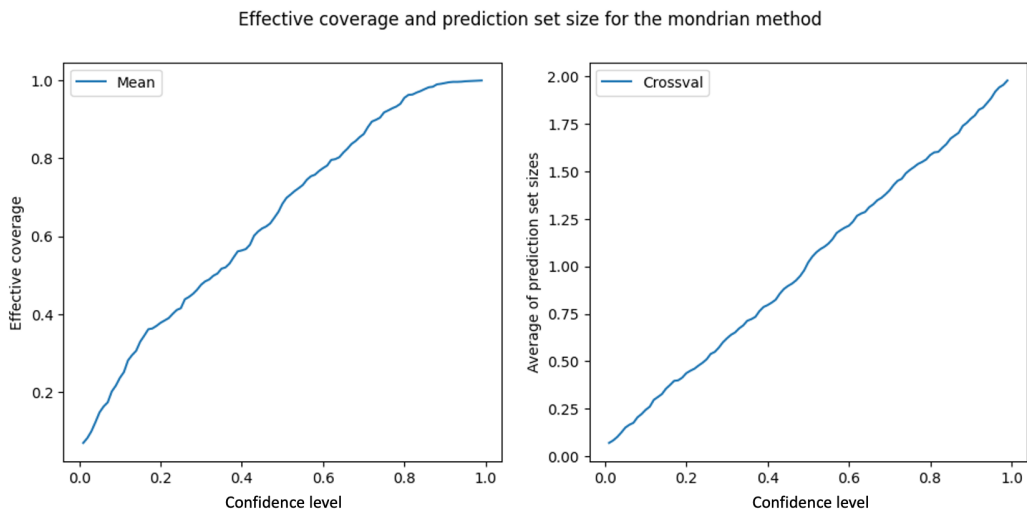


Figure 6: Effective coverage and average prediction set size plotted against confidence level

7.3. Power saving from selective scrubbing

Scrubbing is a resource-intensive operation that can impact the performance of the system during its execution. The time taken to complete a scrubbing operation depends on various factors, such as the size of the HDD being scrubbed. For instance, scrubbing a 1TB HDD may take a few to several hours, while scrubbing an 8TB HDD could take significantly longer, potentially a day or more. Assuming an average power consumption of 7 watts during a 6-hour scrubbing operation for a single HDD, the total energy consumed would be 42 watt-hours (Wh). It’s important to note that power consumption during scrubbing can vary for different disks in a data center, depending on factors like disk size, manufacturer, and storage operations. Taking an average value for power usage comparison, if selective scrubbing is performed on 28,000 disks instead of scrubbing all 120,000 disks in a data center based on results from the Baidu open-source dataset, significant power and energy savings can be achieved for the entire data center.

8. Conclusion

The complexity and uncertainty of individual storage components in large-scale data centers pose challenges to business continuity. While proactive approaches like monitoring and failure analysis have been implemented, machine learning approaches may have false positive concerns in real-world applications with numerous disk drives. In this paper, we propose a fine-grained approach to disk scrubbing using a learning framework based on Mondrian conformal prediction, evaluated on the Baidu open-source dataset.

Our method provides a modest yet effective contribution from a methodological perspective. It tackles the issue of aggressive scrubbing of the entire storage array by utilizing Mondrian conformal predictors to assign health scores to each drive and selectively targeting disks with lower scores for scrubbing. This approach generates a prioritized list for the scheduler engine, leveraging drive failure analysis and quantifying disk health across the entire storage pool. As a result, only 22.7% of the drives need to be scrubbed, leading to power savings and improved reliability.

Future work could involve incorporating Venn-Abers predictors, which offer calibrated probabilities for predictions and could further enhance the accuracy and effectiveness of our approach (Vovk and Petej, 2012). By incorporating such predictors, we could potentially refine and optimize our method for even better performance in identifying and addressing potential disk failures in large-scale data centers.

References

- MAPIE - Model Agnostic Prediction Interval Estimator ; MAPIE 0.6.3 documentation — mapie.readthedocs.io. <https://mapie.readthedocs.io/en/latest/>. [Accessed 08-Apr-2023].
- Jonathan Alvarsson, Staffan Arvidsson McShane, Ulf Norinder, and Ola Spjuth. Predicting with confidence: using conformal prediction in drug discovery. *Journal of Pharmaceutical Sciences*, 110(1):42–49, 2021.

- Lakshmi N Bairavasundaram, Garth R Goodson, Shankar Pasupathy, and Jiri Schindler. An analysis of latent sector errors in disk drives. In *Proceedings of the 2007 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, pages 289–300, 2007.
- Nijaz Bajgorić, Lejla Turulja, and Amra Alagić. Downtime and business continuity. In *Always-On Business: Aligning Enterprise Strategies and IT in the Digital Age*, pages 29–50. Springer, 2022.
- Henrik Boström and Ulf Johansson. Mondrian conformal regressors. In *Conformal and Probabilistic Prediction and Applications*, pages 114–133. PMLR, 2020.
- Euy-Hyun Cho, Jeong-Kyu Park, and Jong-Gyu Chae. The accelerated life test of hard disk in the environment of pacs. *Journal of Digital Contents Society*, 16(1):63–70, 2015.
- DrTycoon. Hdds dataset (baidu inc.), Jan 2023. URL <https://www.kaggle.com/datasets/drtycoon/hdds-dataset-baidu-inc>.
- Greg Hamerly, Charles Elkan, et al. Bayesian approaches to failure prediction for disk drives. In *ICML*, volume 1, pages 202–209. Citeseer, 2001.
- Ilias Iliadis, Robert Haas, Xiao-Yu Hu, and Evangelos Eleftheriou. Disk scrubbing versus intra-disk redundancy for high-reliability raid storage systems. *ACM SIGMETRICS Performance Evaluation Review*, 36(1):241–252, 2008.
- Ilias Iliadis, Robert Haas, Xiao-Yu Hu, and Evangelos Eleftheriou. Disk scrubbing versus intradisk redundancy for raid storage systems. *ACM transactions on storage (TOS)*, 7(2):1–42, 2011.
- Josu Ircio, Aizea Lojo, Jose A Lozano, and Usue Mori. A multivariate time series streaming classifier for predicting hard drive failures [application notes]. *IEEE Computational Intelligence Magazine*, 17(1):102–114, 2022.
- Tianming Jiang, Ping Huang, and Ke Zhou. Scrub unleveling: Achieving high data reliability at low scrubbing cost. In *2019 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pages 1403–1408, 2019. doi: 10.23919/DATE.2019.8715169.
- Junping Liu, Ke Zhou, Zhikun Wang, Liping Pang, and Dan Feng. Modeling the impact of disk scrubbing on storage system. *J. Comput.*, 5(11):1629–1637, 2010.
- Jie Lu, Anjin Liu, Fan Dong, Feng Gu, Joao Gama, and Guangquan Zhang. Learning under concept drift: A review. *IEEE transactions on knowledge and data engineering*, 31(12):2346–2363, 2018.
- Ruiming Lu, Erci Xu, Yiming Zhang, Fengyi Zhu, Zhaosheng Zhu, Mengtian Wang, Zongpeng Zhu, Guangtao Xue, Jiwu Shu, Minglu Li, et al. Perseus: A {Fail-Slow} detection framework for cloud storage systems. In *21st USENIX Conference on File and Storage Technologies (FAST 23)*, pages 49–64, 2023.

- Sidi Lu, Bing Luo, Tirthak Patel, Yongtao Yao, Devesh Tiwari, and Weisong Shi. Making disk failure predictions smarter! In *FAST*, pages 151–167, 2020.
- Rachel Luo, Shengjia Zhao, Jonathan Kuck, Boris Ivanovic, Silvio Savarese, Edward Schmerling, and Marco Pavone. Sample-efficient safety assurances using conformal prediction. In *Algorithmic Foundations of Robotics XV: Proceedings of the Fifteenth Workshop on the Algorithmic Foundations of Robotics*, pages 149–169. Springer, 2022.
- Ao Ma, Fred Douglass, Guanlin Lu, Darren Sawyer, Surendar Chandra, and Windsor Hsu. RAIDShield: Characterizing, monitoring, and proactively protecting against disk failures. In *13th USENIX Conference on File and Storage Technologies (FAST 15)*, pages 241–256, Santa Clara, CA, February 2015. USENIX Association. ISBN 978-1-931971-201. URL <https://www.usenix.org/conference/fast15/technical-sessions/presentation/ma>.
- Valery Manokhin. Awesome conformal prediction, April 2022. URL <https://doi.org/10.5281/zenodo.6467205>.
- Soundouss Messoudi, Sylvain Rousseau, and Sébastien Destercke. Deep conformal prediction for robust models. In *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 528–540. Springer, 2020.
- Soundouss Messoudi, Sébastien Destercke, and Sylvain Rousseau. Class-wise confidence for debt prediction in real estate management: discussion and lessons learned from an application. In *Conformal and Probabilistic Prediction and Applications*, pages 211–228. PMLR, 2021.
- Ningfang Mi, Alma Riska, Evgenia Smirni, and Erik Riedel. Enhancing data availability in disk drives through background activities. In *2008 IEEE International Conference on Dependable Systems and Networks With FTCS and DCC (DSN)*, pages 492–501. IEEE, 2008.
- Alina Oprea and Ari Juels. A clean-slate look at disk scrubbing. In *FAST*, pages 57–70, 2010.
- Omid Orang, Rodrigo Silva, Petrônio Cândido de Lima e Silva, and Frederico Gadelha Guimarães. Solar energy forecasting with fuzzy time series using high-order fuzzy cognitive maps. In *2020 IEEE international conference on fuzzy systems (FUZZ-IEEE)*, pages 1–8. IEEE, 2020.
- Jehan-François Pâris, Thomas Schwarz, Ahmed Amer, and Darrell DE Long. Improving disk array reliability through expedited scrubbing. In *2010 IEEE Fifth International Conference on Networking, Architecture, and Storage*, pages 119–125. IEEE, 2010.
- Eduardo Pinheiro, Wolf-Dietrich Weber, and Luiz André Barroso. Failure trends in a large disk drive population. 2007.
- Teerat Pitakrat, André van Hoorn, and Lars Grunske. A comparison of machine learning algorithms for proactive hard disk drive failure detection. In *Proceedings of the 4th International ACM Sigsoft Symposium on Architecting Critical Systems, ISARCS '13*, page 1–10,

- New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450321235. doi: 10.1145/2465470.2465473. URL <https://doi.org/10.1145/2465470.2465473>.
- Junkil Ryu and Chanik Park. Effects of data scrubbing on reliability in storage systems. *IEICE TRANSACTIONS on Information and Systems*, 92(9):1639–1649, 2009.
- Bianca Schroeder, Sotirios Damouras, and Phillipa Gill. Understanding latent sector errors and how to protect against them. *ACM Transactions on storage (TOS)*, 6(3):1–23, 2010.
- Glenn Shafer and Vladimir Vovk. A tutorial on conformal prediction. *Journal of Machine Learning Research*, 9(3), 2008.
- Jitendra Singh and Rahul Deo Vishwakarma. System and method for survival forecasting of disk drives using semi-parametric transfer learning, January 24 2023. US Patent 11,561,701.
- Xiaoyi Sun, Krishnendu Chakrabarty, Ruirui Huang, Yiquan Chen, Bing Zhao, Hai Cao, Yinhe Han, Xiaoyao Liang, and Li Jiang. System-level hardware failure prediction using deep learning. In *Proceedings of the 56th Annual Design Automation Conference 2019*, pages 1–6, 2019.
- Rahul Deo Vishwakarma and Bing Liu. System and method for persistent storage failure prediction, April 22 2021. US Patent App. 16/656,875.
- Rahul Deo Vishwakarma and Jayanth Kumar Reddy Perneti. Method and system for reliably forecasting storage disk failure, February 4 2021. US Patent App. 16/529,499.
- Vladimir Vovk and Ivan Petej. Venn-abers predictors. *arXiv preprint arXiv:1211.0025*, 2012.
- Vladimir Vovk, David Lindsay, Ilia Nouretdinov, and Alex Gammerman. Mondrian confidence machine. *Technical Report*, 2003.
- Vladimir Vovk, Alexander Gammerman, and Glenn Shafer. *Algorithmic learning in a random world*. Springer International Publishing, Cham, Switzerland, 2 edition, December 2022.
- Jie Yu. Hard disk drive failure prediction challenges in machine learning for multi-variate time series. In *Proceedings of the 2019 3rd International Conference on Advances in Image Processing*, pages 144–148, 2019.
- Ji Zhang, Yuanzhang Wang, Yangtao Wang, Ke Zhou, Schelter Sebastian, Ping Huang, Bin Cheng, and Yongguang Ji. Tier-scrubbing: An adaptive and tiered disk scrubbing scheme with improved mtd and reduced cost. In *2020 57th ACM/IEEE Design Automation Conference (DAC)*, pages 1–6, 2020. doi: 10.1109/DAC18072.2020.9218551.
- Yuqi Zhang, Wenwen Hao, Ben Niu, Kangkang Liu, Shuyang Wang, Na Liu, Xing He, Yongwong Gwon, and Chankyu Koh. Multi-view feature-based {SSD} failure prediction: What, when, and why. In *21st USENIX Conference on File and Storage Technologies (FAST 23)*, pages 409–424, 2023.