

# Tier-Scrubbing: An Adaptive and Tiered Disk Scrubbing Scheme

Ji Zhang<sup>§†</sup>, Ke Zhou<sup>§</sup>, Ping Huang<sup>§</sup>, Sebastian Schelter<sup>†</sup>, Bin Cheng<sup>‡</sup>, Yongguang Ji<sup>‡</sup>  
<sup>§</sup>Huazhong University of Science and Technology, <sup>†</sup>New York University, <sup>‡</sup>Tencent Inc.

## 1 Introduction

Sector errors are a common type of error in disks. A sector error that occurs during I/O operations might cause inaccessibility of an application [1]. Many disk scrubbing schemes have been proposed to solve this problem [2, 3]. However, existing approaches have limitations. First, schemes use machine learning (ML) to predict latent sector errors (LSEs) based on S.M.A.R.T data [4], but only leverage a single snapshot of training data to make a prediction, and thereby ignore sequential dependencies between different statuses of a drive over time. Second, they accelerate the scrubbing at a fixed rate based on the results of a binary classification model, which may result in unnecessary increases in scrubbing cost. Third, they naively accelerate the scrubbing of the full disk which has LSEs based on the predictive results, but neglect partial high-risk areas (the areas that have a higher probability of encountering LSEs). Lastly, they do not employ strategies to scrub these high-risk areas in advance based on I/O accesses patterns, in order to further increase the efficiency of scrubbing. We address these discussed disadvantages of the existing methods by designing our Tier-Scrubbing (*TS*) scheme to achieve a lower Mean Time To Detection (MTTD) accompanied by a decrease in the scrubbing cost.

## 2 Proposed Approach

Figure 1 provides an overview of our proposed scrubbing scheme *TS*. It mainly combines the following three parts:

- (1) **ASRC at Disk Level.** We propose an Adaptive Scrubbing Rate Controller (ASRC) that contains a Long Short-Term Memory based model [5], capable of learning long-term dependencies to predict the sector risk degree (e.g., in range of 0-7) rather than just a binary classification and leverage the results to determine an adaptive scrubbing rate at disk level.
- (2) **Locate High-risk Areas at Sector Level.** Our experimental results show that the probability of the occurrence of sector errors in a single disk is not evenly distributed, but has peaks around certain localities. The probability that a new sector error occurs around a previously observed sector errors of a disk is much higher than in other areas. Therefore, we focus on these high-risk areas of LSE disks so as to scrub these areas with a higher priority at the sector level.
- (3) **Piggyback Scrubbing at the Level of I/O Operations.** We

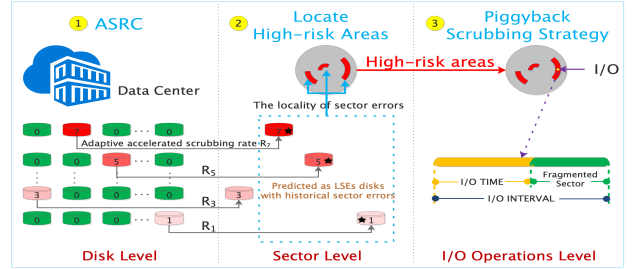
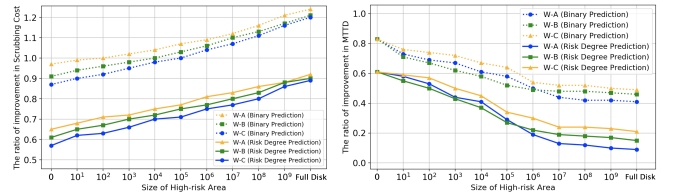


Figure 1: The overall scrubbing scheme of proposed *TS*.

propose a piggyback scrubbing strategy to scrub high-risk areas at the level of I/O operations. When the application I/O operations access these areas, we execute a piggyback read operation which immediately scrubs the fragmented sectors untouched by the I/O operation. In this case, the disk head just needs to seek within a tiny area after conducting I/O operations and ultimately reduce the cost of sequential scrubbing because we reduce the frequency of head movement for the fragmented sectors.



(a)  $C_r$  (The lower, the better) (b)  $M_r$  (The lower, the better)  
 Figure 2: Scheme *TS* achieves lower MTTD and scrubbing cost.

## 3 Preliminary Results

We measure the ratio of improvement in MTTD ( $M_r$ ) and scrubbing cost ( $C_r$ ) compared to the state-of-the-art scheme *SU* [3]. Figures 2(a) & 2(b) show  $M_r$  and  $C_r$  under different sizes of high-risk areas using three real world workloads. *TS* achieves lower MTTD and scrubbing cost in all cases. Moreover, the results of the solid lines achieve better performance than the dotted ones, which demonstrates that the predictor we designed in ASRC for the sector risk degree is more efficient than just predicting whether the disk is an LSE disk or not. Limiting the high-risk area size to  $10^7$  sectors can simultaneously decrease the MTTD by about 80% and the scrubbing cost by about 20%, compared to the scheme *SU*.

## References

- [1] Yong Xu, Kaixin Sui, Randolph Yao, Hongyu Zhang, Qingwei Lin, Yingnong Dang, Peng Li, Keceng Jiang, Wenchi Zhang, Jian-Guang Lou, Murali Chintalapati, and Dongmei Zhang. Improving service availability of cloud systems by predicting disk error. In *2018 USENIX Annual Technical Conference (USENIX ATC 18)*, pages 481–494, 2018.
- [2] Farzaneh Mahdisoltani et al. Improving storage system reliability with proactive error prediction. In *Proceedings of the USENIX Annual Technical Conference*, pages 391–402, Santa Clara, CA, July 2017. USENIX Association.
- [3] Tianming Jiang, Ping Huang, and Ke Zhou. Scrub unleveling: Achieving high data reliability at low scrubbing cost. In *Design, Automation & Test in Europe Conference & Exhibition, DATE 2019, Florence, Italy, March 25-29*, pages 1403–1408, 2019.
- [4] Bruce Allen. Monitoring hard disks with smart. *Linux J.*, 2004(117):9, January 2004.
- [5] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, November 1997.