AI-Based Layout Optimization of HDDs in Full-Rack Heterogeneous Server and Storage Systems

Yi-Ju Liao [1], Jen-Yuan Chang [2], *Senior Member, IEEE*

[1, 2] Department of Power Mechanical Engineering, National Tsing Hua University, Hsinchu, Taiwan

Heterogeneous rack systems combining servers, storage arrays, and switches are common in modern data centers, yet hard disk drives (HDDs) within such configurations face performance degradation from complex structural and airflow-induced vibrations. This work proposes a slot-based, machine learning framework to predict HDD degradation risk and guide layout optimization. Each rack slot is characterized by a feature vector comprising modal frequency, local stiffness, airflow excitation, and vibrational coupling, estimated via representative simulations. Decision tree and neural network models are trained to predict IOPS degradation based on these descriptors. A reference configuration—switches at the top, servers in the middle, storage at the bottom—is used to demonstrate rack-level risk mapping. The framework supports explainability through SHAP analysis and enables layout recommendations that reduce slot-level exposure to high-risk conditions. The proposed method is modular, generalizable across rack designs, and suitable for deployment-oriented optimization of HDD placement strategies.

***Index Terms*—Data center reliability, hard disk drives, machine learning, vibration analysis**

I. Introduction

The increasing heterogeneity of data center rack systems—comprising co-located servers, storage arrays, and switch modules—poses new challenges to mechanical stability and long-term hard disk drive (HDD) reliability. While prior studies have investigated HDD failure under isolated conditions or uniform enclosures, real-world deployments often exhibit complex physical interactions due to structural discontinuities, varying airflow patterns, and device-specific mounting configurations [1]. These factors result in spatially nonuniform vibrational and thermal environments that can significantly degrade HDD input/output operations per second (IOPS) performance. However, current industry practices lack a predictive framework capable of modeling such position-dependent degradation and guiding layout optimization at the full-rack level. To address this gap, we propose a slot-based predictive framework that combines structural and thermal simulation with machine learning techniques to assess HDD degradation risk and support layout-aware planning. Each slot in the rack is abstracted into a physical feature vector describing its modal response, airflow exposure, local mounting conditions, and proximity-based coupling effects. These features are derived from representative finite element and airflow models of heterogeneous full-rack configurations. Supervised learning models—specifically decision trees and artificial neural networks—are then trained to predict IOPS degradation at the slot level. The resulting framework supports SHAP-based model interpretability, enabling identification of high-risk regions and informing placement strategies that align with rack design constraints. This paper introduces the proposed methodology, demonstrates its application on a reference configuration, and outlines future directions for integrating interpretable slot-level risk analysis into rack-level design workflows.

II. System Modeling

In the proposed framework, the heterogeneous rack comprising switches, servers, and storage modules—is abstracted into a set of discrete slots, each of which is represented by a feature vector encoding its physical and environmental characteristics [4]. This slot-level abstraction enables the modeling of complex interdependencies between device types, airflow dynamics, and structural responses within the rack.

Each slot is described by a set of features that include: (1) modal frequency, obtained from simplified structural simulations or vendor specifications; (2) local stiffness, which accounts for tray design and mounting conditions; (3) airflow velocity, derived from computational fluid dynamics (CFD) simulations or empirical fan profiles; (4) proximity-based coupling intensity, modeling vibrational energy transmission from adjacent slots; and (5) device type, a categorical indicator (e.g., HDD, server, switch, or empty). These features form a structured dataset with each row corresponding to a slot and each column to a physical parameter. This dataset serves as input to supervised learning models trained to predict slot-level IOPS degradation risks. A schematic of this modeling approach is shown in **Fig. 1**.

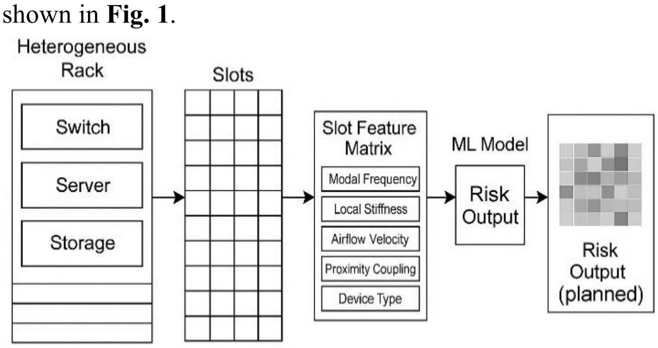

Fig. 1. Slot-based degradation modeling and explainable learning framework

III. MACINE LEARNING ARCHITECTURE

A supervised learning framework is employed to predict slot-specific hard disk drive (HDD) performance degradation using physical and environmental descriptors derived from rack-level modeling. Each rack configuration is transformed into a structured feature matrix, where each row corresponds to a logical slot and each column represents a characteristic such as modal frequency, airflow velocity, local stiffness, vibrational coupling intensity, or device type. For storage enclosures housing multiple HDDs, each drive is individually abstracted into a separate logical slot. Intra-enclosure effects such as shared airflow conditions or structural coupling are reflected through common feature values across these slots, enabling the model to preserve localized distinctions while maintaining global coherence. This abstraction permits flexible modeling of mixed server–storage systems and supports evaluation of layout variations without requiring device-specific knowledge.

The predictive framework employs decision trees (DTs) for interpretability and artificial neural networks (ANNs) for capturing nonlinear interactions across slot-level features. Training data are generated through a combination of representative simulations and hypothetical rack configurations, ensuring generalizability across diverse deployment scenarios. The model output is a degradation risk score associated with each HDD slot, corresponding to predicted IOPS loss under specific vibrational and thermal conditions. To support interpretability, a SHAP-based feature attribution module is planned, enabling identification of dominant physical drivers of degradation across the slot population. While this component is under development, its intended integration highlights the extensibility of the framework for deployment-aware and layout-sensitive design optimization [2] [3].

Each slot's predicted degradation risk score is modeled as a function of its physical and environmental characteristics. The risk score R_i for slot i is computed as a weighted sum of feature contributions:

$$R_i = \beta_1 f_{modal,i} + \beta_2 v_{air,i} + \beta_3 k_{stiff,i} + \beta_4 c_{couple,i} + \beta_5 d_{type,i} \quad (1)$$

Where $f_{modal,i}$ is the first-mode frequency, $v_{air,i}$ is the local airflow velocity, $k_{stiff,i}$ is the slot stiffness, $c_{couple,i}$ is the coupling intensity from adjacent slots, and $d_{type,i}$ denotes the device type (e.g., HDD, server, or switch). The coefficients β_j are learned by the regression model. Equation (1) is used both as a conceptual formulation and as the basis for SHAP-based feature attribution in model interpretation.

TABLE I
SLOT-LEVEL FEATURE DESCRIPTIONS

Feature Name	Unit	Description
Modal Frequency	Hz	First mode frequency of the slot structure
Airflow Velocity	m/s	Estimated airflow affecting the slot
Local Stiffness	N/m	Tray stiffness or mounting-induced compliance
Coupling Intensity	–	Vibrational influence from adjacent slots
Device Type	Categorical	Type of device installed (Storage, Server, Switch, Empty)

IV. Preliminary Evaluation and Deployment Plan

The proposed framework is designed to support early-stage evaluation through simulated slot configurations and hypothetical rack layouts. In the initial phase, training data will be generated using synthesized modal and thermal descriptors representative of typical server–storage deployments. A limited matrix of 48–96 slots will be constructed to emulate vibration-sensitive regions, and corresponding IOPS degradation scores will be estimated using simplified mapping functions based on modal alignment thresholds and airflow-induced excitation levels.

Model training will begin with decision trees to establish interpretable rule sets, followed by artificial neural networks to capture complex dependencies. Initial SHAP analysis will be applied to assess feature contributions across slots, and risk maps will be visualized to support slot-level performance assessment. This workflow will enable deployment-focused layout evaluations without reliance on proprietary device data. Results from this preliminary phase will be showcased at the poster session, with ongoing work focused on expanding dataset diversity and integrating optimization modules.

V. Conclusion

This work introduces a slot-based machine learning framework for predicting HDD performance degradation across heterogeneous rack systems. The proposed modeling approach enables configuration-level flexibility, modular scalability, and integration with explainability techniques. Future work will focus on extending the framework to support real-time layout optimization and system-wide deployment scenarios.

References

[1] J. H. L. Hansen, W. Song, and T. Zhang, "Vibration-induced performance degradation of hard disk drives in data centers," *IEEE Trans. Magn.*, vol. 53, no. 11, pp. 1–7, Nov. 2017.

[2] A. Shrikumar, P. Greenside, and A. Kundaje, "Learning important features through propagating activation differences," *Proc. 34th Int. Conf. Mach. Learn.*, vol. 70, pp. 3145–3153, 2017.

[3] S. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, vol. 30, pp. 4765–4774, 2017.

[4] T. Kim and B. Karp, "The impact of airflow and rack layout on hard drive reliability in high-density servers," *Proc. ASME InterPACK*, pp. 1–8, 2019.

[5] D. Silver, J. Barto, and R. S. Sutton, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA: MIT Press, 2018, pp. 145–150.