AI-based Proactive Storage Failure Management in Software-Defined Data Centres

Yongqing Zhu *
School of Science and Technology, Singapore University of Social Sciences, Singapore
yqzhu@suss.edu.sg

Kiam Cheng How
School of Science and Technology, Singapore University of Social Sciences, Singapore
caseyhowkc@suss.edu.sg

Horng Jyh Wu
School of Science and Technology, Singapore University of Social Sciences, Singapore
paulwuhj@suss.edu.sg

Qi Cao
School of Computing Science, University of Glasgow, UK/Singapore
qi.cao@glasgow.ac.uk

Abstract — Proactive failure management is essential to alleviate potential risks of service unavailability and downtime in Software-Defined Data Centres (SDDCs). Artificial Intelligence (AI) models enable proactive failure management by predicting and addressing potential failures before they actually occur. This paper proposes an AI-based Proactive Storage Failure Management (APSFM) solution for intelligent data centre management. The proposed solution includes a four-stage framework that employs AI models to predict failures efficiently. The study uses Random Forest and Artificial Neural Network models as examples to predict disk failures by employing Self-Monitoring, Analysis, and Reporting Technology (SMART) attributes. The experimental results have shown that both models can achieve high prediction performance.

Keywords: Storage Management; Artificial Intelligence; Failure Prediction; Software-Defined Data Centre.

1 INTRODUCTION

The global data centre traffic is growing tremendously, especially during the pandemic when many organizations and companies move their operations and services online. Software-Defined Data Centres (SDDCs) have been widely deployed to help organizations improve the infrastructure and operations. In the SDDC, hardware infrastructure is abstracted and virtualized as logical resources, which can be dynamically provisioned and orchestrated as integrated computing, storage, and networking services through software. SDDCs can improve the flexibility in resource provisioning and management, that increase the hardware resource utilization and efficiency [1].

It is crucial to maintain the hardware and software in SDDCs in good conditions in order to ensure high availability of data centre services. However, failures happen frequently in data centres, with storage failures being the primary cause of service disruptions. The research work [2] indicates that 78% of hardware replacement in data centres is due to hard disk failures. When failures occur, data centre engineers need to react by repairing or replacing the affected hardware devices through a remediation process. Such a reactive process will lead to downtime and performance degradation inevitably. To improve the service availability and reliability, remedial actions should be taken proactively when failures are predicted but yet to actually occur in SDDCs.

* Corresponding author.

Artificial Intelligence (AI) technologies have been extensively studied in data centres to achieve multifaceted benefits. Deep Reinforcement Learning technique is applied in [3] for data centre cooling optimization. Google uses Deepmind AI to reduce cooling costs in data centres [4]. Artificial Neural Network (ANN) model is used in [5] to predict energy demands and solve the energy dispatch problem in data centres. Machine learning models are reported in [6-9] for disk failure predictions to improve data centre service availability, including Naïve Bayes Classifier, Support-Vector Machine (SVM), etc. A cost-effective multi-failure resilient replication scheme (MRR) is introduced to handle both correlated and non-correlated machine failures [10]. MRR optimizes data availability by considering data popularities and aims to minimize consistency maintenance and storage expenses.

In this paper, an AI-based Proactive Storage Failure Management (APSFM) solution is proposed with the goal of improving service availability and reducing operational disruption in SDDCs. The AI model is used to predict storage failures based on data collected from storage devices. Whenever a failure is predicted to happen, remediation actions will be proactively performed before the actual failure occurs, therefore minimizing downtime. With the aid of AI, the proactive process will mitigate the problems of reactive alternatives such as service interruptions, data loss, performance deterioration, etc.

This paper is organized as follows. Section 2 elaborates the framework of the AI-based Proactive Storage Failure Management solution. Two AI models are presented in Section 3 for the disk failure prediction, followed by performance evaluations in Section 4. The whole paper is concluded in Section 5.

2 AI-BASED PROACTIVE STORAGE FAILURE MANAGEMENT

A proactive storage management solution can effectively monitor data storage status, predict potential storage failures, and automatically trigger remediation actions to prevent those failures. Recent research has focused on the application of AI technologies to data centres, with the target to improve performance and reduce operation costs intelligently. In this paper, we propose the APSFM solution for achieving the smart data centre management. Figure 1 illustrates the framework of the proposed solution, which consists of four stages: data collection, model training, failure prediction, and proactive remediation.

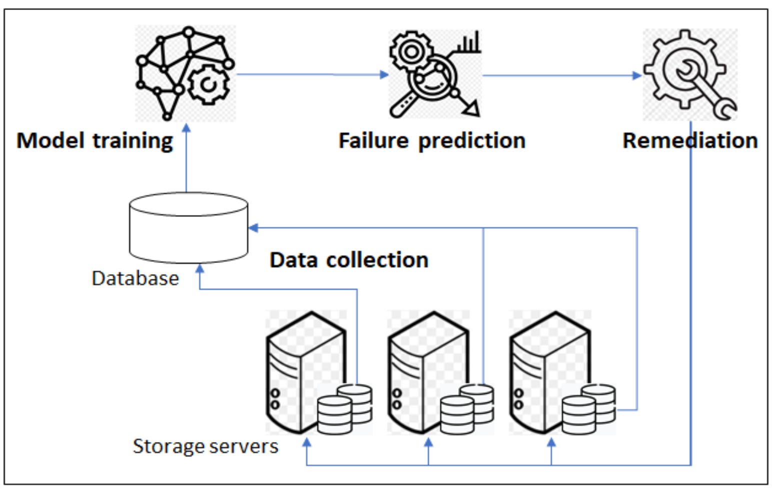

Figure 1. Framework of the proposed AI-based Proactive Storage Failure Management APSFM

During the initial stage of data collection, various data can be collected from different sources within the data centre to provide a comprehensive view of storage status. Some relevant data sources include: Self-Monitoring, Analysis, and Reporting Technology (SMART) data [11] from disk drives; Input/Output (I/O) data from file systems (such as read/write throughput and speed); sensor data from the data centre (such as humidity and temperature); etc. All of the data serve as good indicators for the current status of the storage devices. They are collected periodically (e.g., every 2 hours) from various sources and streamed to be stored in a central database.

In the second stage of model training, the data collected in the previous stage are extracted from the database to train an AI model for failure prediction. Machine learning algorithms have been widely studied and employed to predict failures proactively. The prediction problem can be formulated as a binary classification problem using a machine learning approach. The historical data is fed into the machine learning model to train and build the final reference model. Some machine learning algorithms that are suitable for the failure prediction include Naïve Bayes Classifier, SVM, Random Forest, ANN, etc.

The reference model, after being trained and developed, is then utilized to predict failures based on new dataset in the next stage. The trained model classifies the relevant data as positive or negative, thus to indicate whether failures are going to occur in the storage devices. The prediction results are stored in a separate database, which will be used to trigger remediation process in the following stage. If a storage device is functioning for now but being classified as a failed one, a ticket is issued or warning message is sent out to alarm the data centre engineers. The device will then be further examined for rectification or replacement, and a potential failure will be avoided accordingly.

In the last stage of the solution, proactive remediation process is triggered when failures are predicted to happen in the storage devices, e.g., hard disks. The remediation includes the actions to rectify or replace the devices going to fail. For hard disks, the spare disks will be used to replace the disks being predicted to fail. Since the actions take place before the actual failures happen, the remediation process is considered "proactive" approach. Comparatively, the traditional remediation is a type of "reactive" approach, where data centre engineers respond to failures that have already occurred. Figure 2 illustrates the differences between "reactive" approach and "proactive" approach. The reactive approach will inevitably result in performance degradation and service unavailability. For example, when hard disk failures happen in a Redundant Array of Inexpensive Disks (RAID) system, the reactive remediation will incur a rebuilding process to recover data in the failed disk from the rest of disks in RAID [12]. The rebuilding process will consume additional storage and computing resources, that slows down the storage services provided by the RAID system. Instead, a proactive remediation approach will mitigate risks with higher availability and less downtime comparing to the reactive one.

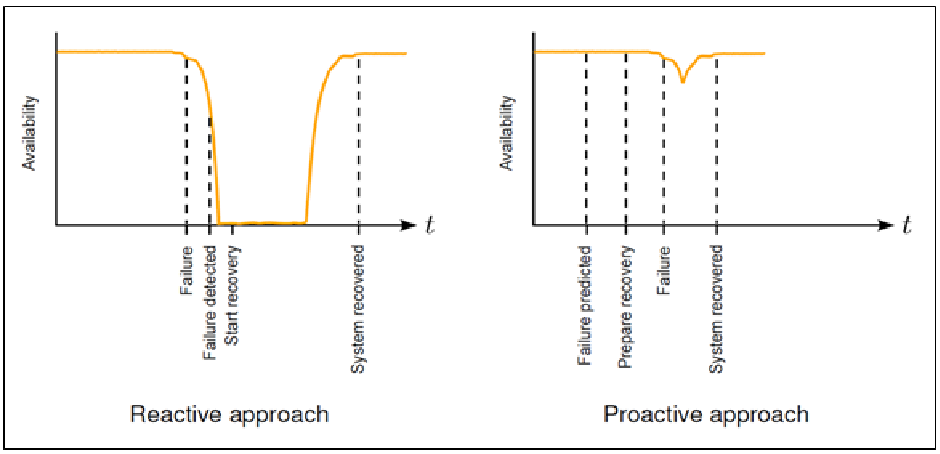

Figure 2. Proactive approach vs Reactive approach

3 PROACTIVE FAILURE PREDICTION

Model training and failure prediction are very important in the APSFM solution, as the proactive remediation depends on the accuracy of the prediction results. When a failure is predicted to happen, the remediation process initiates to rectify or replace the device even though it is still operational. Therefore, it is crucial to have a reference model that can accurately predict failures, as false positives will result in unnecessary hardware device replacements and waste resources.

Recent research has explored various AI models for intelligent data centre management, which can be utilized in the stages of model training and failure prediction in APSFM. Given that more than three-quarters of hardware replacement in data centres are caused by hard disk failures [2], we are particularly interested in applying AI models to predict disk failures. Several machine learning models [6-9] have been proposed for disk failure prediction using SMART data, but the prediction results of these approaches are far from satisfactory. More recently, deep learning techniques have been applied to disk failure predictions as well [13-14]. However, a recent study [15] suggested that complex deep neural networks are not necessary for disk failure prediction when SMART attributes are the only data source. Therefore, this study uses two machine learning models, Random Forest and ANN, as examples for predicting hard disk drives in data centres.

3.1 Random Forest

Random Forest [17] is an ensemble learning method that can be used to solve both classification and regression problems. It is composed of multiple Decision Trees, which are popular supervised learning algorithms. Figure 3 shows a binary Decision Tree that includes three types of nodes: root node, decision node, and leaf node. In a classification task, the algorithm firstly splits training dataset into two branches based on certain feature, leading to a root node. The segregation of data continues based on different features, and thus forms different levels of decision nodes. The process stops when a leaf node is arrived and data cannot be further split, meaning either maximum tree depth or minimum node records are satisfied. After the Decision Tree model is built, it is then used to classify testing dataset for predictive analysis.

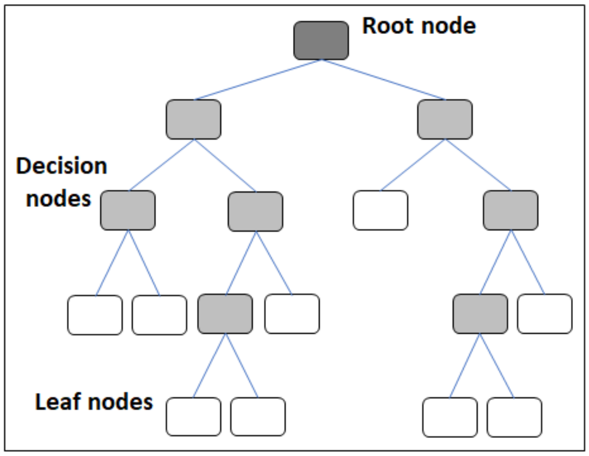

Figure 3. Decision Tree classification model

Though popular in various applications, a single Decision Tree is likely to bring problems like bias and overfitting. Random Forest addresses these problems by averaging or combining results of different Decision Trees. In Random Forest algorithm, training dataset is randomly selected into multiple training samples with different features, each of which is used to train a Decision Tree independently. After all trees are constructed, testing dataset is fed into different trees to get prediction results. The algorithm then applies a voting mechanism to select the result with the most votes as the final prediction result. Using the ensemble method, Random Forest can reduce the risk of overfitting and improve prediction accuracy.

3.2 Artificial Neural Network

ANN, also known as Neural Networks, operates on the same concept as biological neural networks [18]. An ANN comprises a group of neurons (nodes) organized into different layers: an input layer, one or more hidden layers, and an output layer. Neurons are connected via weighted connections. Each neuron takes in data with numerical values from the previous layer, along with the associated weights, and applies a nonlinear function for the computations. The output is then sent to the neurons in the subsequent layer for further calculations. After cycles of learning, the ANN model can be trained to improve performance by adjusting the weights of connections.

Multi-Layer Perceptron (MLP) is a fully-connected feedforward ANN that achieves failure predictions by classifying devices into good and failed groups. In MLP, each node (except for those in the input layer) is fully connected to all nodes in the preceding layer. As illustrated in Figure 4, data flow in from the input layer and are forwarded through all hidden layers directly to the output layer. Each node in the hidden and output layers performs calculations using a nonlinear activation function. As a supervised learning model, the output layer results are compared with the expected results, and errors are computed during the learning phase. MLP uses a backpropagation method to adjust the connection weights and ensure the error to be reduced after each learning cycle.

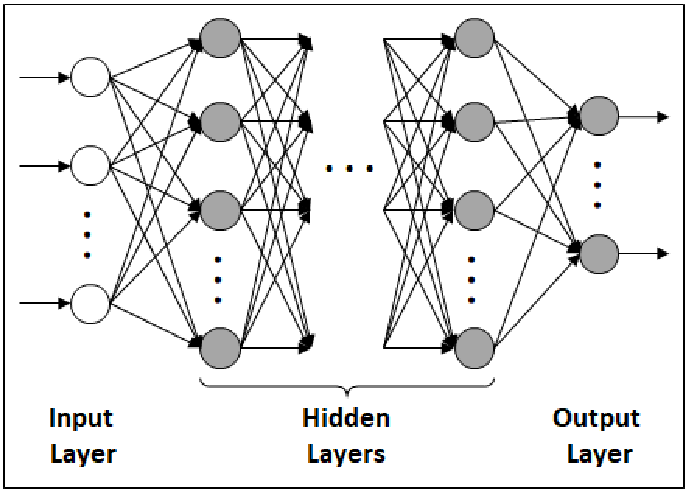

Figure 4. A fully-connected MLP model

4 PERFORMANCE EVALUATIONS

We have evaluated the two machine learning models, Random Forest and ANN (specifically MLP), for model training and failure prediction in the proposed APSFM solution. PyTorch [20], an open source machine learning tool, was used to build the two models for the performance evaluations. The machine learning models classified the hard disks into the failed disk and good disk categories.

4.1 SMART Dateset

During the hard disk manufacturing process, SMART [11] is implemented to monitor disk health and detect failures. In this work, SMART dataset collected by Baidu Inc. [21] is used to evaluate these two machine learning models. A total set of 4,006,453 records were collected from 23,395 enterprise disks in the data centres, including 22,962 disks with good conditions and 433 disks going to be failed within predefined time. SMART data records were collected every hour for each disk. Records of good disks were kept in the system for one week. While data records of the failed disks were kept for 20 days before the failures happened eventually.

Each data record in the dataset contains 14 columns, with the first column being the index of the disk ranging from 1 to 23,395. The second column is the class label of the disk, with value -1 for the failed disks and +1 for the good disks. The remaining 12 columns are selected SMART attributes from the numerous attributes defined by hard disk manufacturers as follows.

- Raw Read Error Rate;
- Spin Up Time;
- Reallocated Sectors Count;
- Seek Error Rate;
- Power On Hours;
- Reported Uncorrectable Errors;
- High Fly Writes;
- Temperature Celsius;
- Hardware ECC Recovered;
- Current Pending Sector Count;
- Raw value of Reallocated Sectors Count;
- Raw value of Current Pending Sector Count.

Following the method in [16], the values of all SMART attributes are normalized to be within the range of [-1, 1]. These 12 SMART attributes are used as multiple features to input to the machine learning models.

4.2 Evaluation Results of Random Forest

The SMART dataset is used to train the Random Forest model. Each SMART attribute is considered a feature, based on which the training dataset are split recursively at the root nodes or decision nodes. The number of trees in the forest is set to 50 and the number of features sampled for each tree is automatically decided. The training and testing data account for 67% and 33%, respectively. For each of the 50 trees in the Random Forest, training dataset is sampled randomly for training of the tree. After being constructed, these trees are used to predict disk failures independently based on the testing dataset. The final prediction result will be further selected using a voting mechanism.

The Random Forest algorithm was implemented using PyTorch. Figure 5 shows one of the Decision Trees generated by the Random Forest algorithm based on the SMART dataset. It is observed in Figure 6 that the Random Forest can achieve good prediction results with high accuracy as 99.96%. Additionally, a Decision Tree was implemented and trained using the entire SMART dataset. The prediction accuracy obtained by the Decision Tree is very good as well, though slightly lower than that achieved by Random Forest algorithm. The evaluation results suggest that both Random Forest and Decision Tree models could be suitable for predicting disk failures using the SMART attributes.

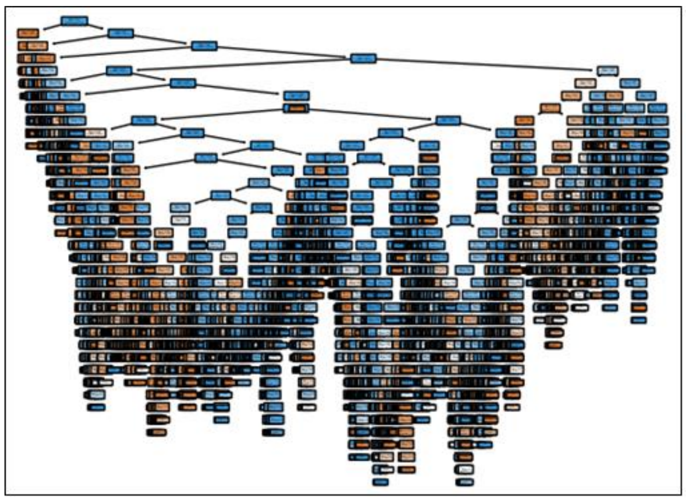

Figure 5. A Decision Tree constructed from the SMART dataset

```
Confusion Matrix:
 [[  51177     308]
 [    231 1270414]]

Classification Report:
               precision    recall  f1-score   support

        -1.0       1.00      0.99      0.99     51485
         1.0       1.00      1.00      1.00   1270645

    accuracy                           1.00   1322130
   macro avg       1.00      1.00      1.00   1322130
weighted avg       1.00      1.00      1.00   1322130

Accuracy:: 0.9995923245066672
```

Figure 6. Performance evaluation of the Random Forest model

4.3 Evaluation Results of MLP Model

Our previous work showed that, the disk failure prediction problem did not require a complex MLP model with large number of layers and neutrons [19]. In this work, we employed an MLP model including four layers with PyTorch. The input layer includes

12 nodes, taking in 12 SMART attributes values. Two hidden layers are implemented in the model, with 12 neurons and 6 neurons, respectively. The output layer consists of one node only, indicating the predicted status of hard disks.

The activation functions used in the MLP model are ReLu for the hidden layers, and Sigmoid for the output layer. During the model training, the errors are calculated using the Binary Cross-Entropy with logits loss function BCEWithLogitsLoss. The connection weights are adjusted at each learning to minimize the loss. The dataset was split following two different training-testing ratios: one is 80:20 with 80% for training data and 20% for testing data; and another is 67:33 with 67% for training data and 33% for testing data. The training and testing sets include both good disks and failed disks. The data was trained in batches of 100 records, and all data was trained for 20 times.

The evaluation results of the MLP model are shown in Figure 7. Both settings of training-testing ratios are able to achieve high performance in terms of accuracy, precision, recall, and F-score. Both accuracy rates are above 95%, while both recalls are as high as above 99%. The setting with 80:20 ratio can obtain slightly higher accuracy and recall than the one with 67:33 ratio, while the latter one gains better performance in precision and F-Score than the former one.

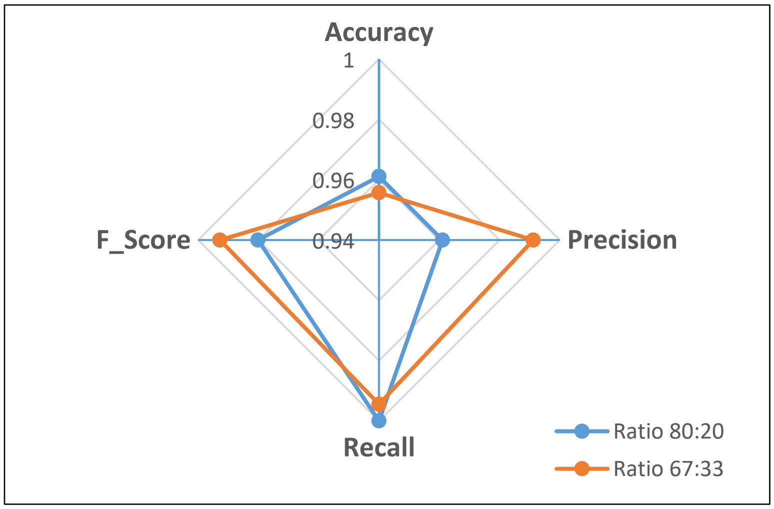

Figure 7. Performance evaluation of the MLP model

5 CONCLUSIONS

This paper proposes an intelligent storage failure management solution APSFM for software-defined data centres. The proposed solution comprising four stages of processes, namely data collection, model training, failure prediction, and proactive remediation. It brings intelligence to the data centre management that can potentially improve service availability and reduce downtime. Two machine learning models, Random Forest and Artificial Neural Network, are implemented for the model training and failure predictions. The evaluation results indicate that both models can effectively predict disk failures using the SMART attributes with high accuracy. Overall, the APSFM solution provides a promising direction for intelligent data centre management in the era of digital transformation.

Although simple, Random Forest and Decision Tree have demonstrated good results of disk failure prediction using SMART dataset. In the future, we will further study how these models would perform with multiple data sources including SMART attributes, system performance and environmental data. Additionally, we will explore to employ different machine learning models for storage failure predictions.

REFERENCES

[1] P. Kerpan, “Software defined data centres will change the way enterprises migrate and deploy to cloud-based applications”, CohesiveFT, Report.

[2] K. V. Vishwanath and N. Nagappan, “Characterizing cloud computing hardware reliability”, in Proceedings of the 1st ACM symposium on Cloud computing (SoCC), Indianapolis, IN, June 2010, pp. 193–204.

[3] Yuanlong Li, Yonggang Wen, Dacheng Tao, Kyle Guan “Transforming Cooling Optimization for Green Data Center via Deep Reinforcement Learning”, IEEE Transactions on Cybernetics, Volume: 50, Issue: 5, May 2020, pp. 2002-2013.

[4] DeepMind AI Reduces Google Data Centre Cooling Bill by 40%. [Online] Available at: https://deepmind.com/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-40/

[5] Yanan Liu, Xiaoxia Wei, Jinyu Xiao, Zhijie Liu, Yang Xu, Yun Tian, “Energy consumption and emission mitigation prediction based on data center traffic and PUE for global data centers”, Global Energy Interconnection, Volume 3, Issue 3, June 2020, Pages 272-282.

[6] Gordon F Hughes, Joseph F Murray, Kenneth Kreutz-Delgado, and Charles Elkan, “Improved disk-drive failure warnings”, IEEE transactions on reliability, 51(3):350–357, 2002.

[7] J. F. Murray, G. F. Hughes, and K. Kreutz-Delgado, “Hard drive failure prediction using non-parametric statistical methods,” in Proceedings of the International Conference on Artificial Neural Networks, June 2003.

[8] J. F. Murray, G. F. Hughes, and K. Kreutz-Delgado, “Machine learning methods for predicting failures in hard drives: A multiple-instance application,” Journal of Machine Learning Research, vol. 6, pp. 783–816, May 2005.

[9] G. Hamerly and C. Elkan, “Bayesian approaches to failure prediction for disk drives,” in Proceedings of the 18th International Conference on Machine Learning, June 2001, pp. 202–209.

[10] J. Liu et al., "A Low-Cost Multi-Failure Resilient Replication Scheme for High-Data Availability in Cloud Storage," in IEEE/ACM Transactions on Networking, vol. 29, no. 4, pp. 1436-1451, Aug. 2021, doi: 10.1109/TNET.2020.3027814.

[11] B. Allen, “Monitoring hard disks with SMART”, Linux Journal, No. 117, Jan 2004.

[12] P. M. Chen, E. K. Lee, G. A. Gibson, R. H. Katz, and D. A. Patterson, “RAID: high-performance, reliable secondary storage”, ACM Computing Surveys, Vol 26, No. 2, June1994, pp. 145–185.

[13] Preethi Anantharaman, Mu Qiao, and Divyesh Jadav, “Large scale predictive analytics for hard disk remaining useful life estimation”, In Proceedings of the 2018 IEEE International Congress on Big Data (BigData Congress), pages 251–254. IEEE, 2018.

[14] Fernando Dione dos Santos Lima, Gabriel Maia Rocha Amaral, Lucas Goncalves de Moura Leite, João Paulo Pordeus Gomes, and Javam de Castro Machado, “Predicting failures in hard drives with LSTM networks”, In Proceedings of the 2017 Brazilian Conference on Intelligent Systems (BRACIS), pages 222–227. IEEE, 2017.

[15] Sidi Lu, Bing Luo, Tirthak Patel, Yongtao Yao, Devesh Tiwari, Weisong Shi, “Making Disk Failure Predictions SMARTer!”, in Proceedings of the 18th USENIX Conference on File and Storage Technologies (FAST20), 2020, pp. 151–167.

[16] B. Zhu, G. Wang, X. Liu, D. Hu, S. Lin, and J. Ma, “Proactive drive failure prediction for large scale storage systems,” in 29th IEEE Conference on Massive Storage Systems and Technologies (MSST), Long Beach, CA, 2013, pp. 1–5.

[17] Leo Breiman, “Random forest”, Machine Learning, volume 45, 2001, pp. 5–32.

[18] S. J. Russell and P. Norvig, “Artificial Intelligence: A Modern Approach (2nd Edition)”, Prentice Hall, 2003.

[19] Yongqing Zhu, Paul Horng Jyh Wu, Fang Liu, Renuga Kanagavelu, “Disk Failure Prediction for Software-Defined Data Centre (SDDC)”, the 7th IEEE International Conference on Cloud and Big Data Computing (CBDCom 2021), Oct 2021.

[20] Pytorch. [Online] Available at: https://pytorch.org/

[21] SMART dataset. [Online] Available at: http://pan.baidu.com/share/link?shareid=189977&uk=4278294944