



Contribution ID: 5

Type: Oral

AI-based Data Popularity, Placement Optimization for a Novel Multi-tiered Storage System at BNL/SDCC Facility

Thursday 14 March 2024 17:10 (20 minutes)

Scientific experiments and computations, particularly in Nuclear Physics (NP) and High Energy Physics (HEP) programs, are generating and accumulating data at an unprecedented rate. Big data presents opportunities for groundbreaking scientific discoveries. However, managing this vast amount of data cost-effectively while facilitating efficient data analysis within a large-scale, multi-tiered storage architecture poses a significant challenge for the Scientific Data and Computing Center (SDCC).

The storage team is currently addressing optimization challenges related to data classification, placement, and migration in the existing multi-tier storage system. While users and administrators manually optimize storage by migrating data based on simple rules derived from human knowledge, decisions, and basic usage statistics, evaluating the placement of data in different storage classes with I/O-intensive workloads remains a complex task.

To overcome the aforementioned challenge and address existing limitations, we have developed a precise data popularity prediction model utilizing state-of-the-art AI/ML techniques. Additionally, we have designed a data placement policy engine based on data popularity, allowing us to migrate infrequently accessed data to more economical storage media, such as tape drives, while storing frequently accessed data on faster yet costlier storage media like HDD or SSD. This strategy optimally places data into the proper storage classes, maximizing storage capacity while minimizing data access latency for end users. This paper delves into the analysis of the data, demonstration patterns, tag files. Specifically, we detail the design and development of an accurate AI/ML prediction model to forecast future data popularity, based on an analysis of access patterns, facilitating optimal data movement and placement. Additionally, we provide insights into the implementation of a policy engine and data placement tool to execute automated migration actions. Finally, the evaluation of different strategies is illustrated, including those involving AI/ML models, etc.

Significance

References

Experiment context, if any

Primary author: HUANG, Qiulan (Brookhaven National Laboratory (US))

Co-authors: Mr LEONARDI, James (Brookhaven National Laboratory); Dr GARONNE, Vincent (Brookhaven National Laboratory); Dr YOO, Shinjae (Brookhaven National Laboratory)

Presenter: HUANG, Qiulan (Brookhaven National Laboratory (US))

Session Classification: Track 2: Data Analysis - Algorithms and Tools

Track Classification: Track 2: Data Analysis - Algorithms and Tools