



Contribution ID: 12

Type: **not specified**

Using AI/ML for Data Placement Optimization in a Multi-Tiered Storage System within a Data Center

Tuesday 5 November 2024 14:30 (30 minutes)

Scientific experiments and computations, particularly in High Energy Physics (HEP) programs, are generating and accumulating data at an unprecedented rate. Effectively managing this vast volume of data while ensuring efficient data analysis poses a significant challenge for data centers. This paper aims to introduce machine learning algorithms to enhance data storage optimization across various storage media, providing a more intelligent, efficient, and cost-effective approach to data management. We begin by outlining the data collection and preprocessing steps used to explore data access patterns. Next, we describe the design and development of a precise data popularity prediction model using AI/ML techniques. This model forecasts future data popularity based on an analysis of access patterns, enabling optimal data movement and placement. Additionally, the paper evaluates the model's performance using key metrics such as F1 score, accuracy, precision, and recall, alongside a comparison with the Least Recently Used (LRU) strategy. The model achieves an optimal prediction accuracy of up to 92% and an optimal F1 score of 0.47. Finally, we present a prototype use case, leveraging real-world file access data to assess the model's performance.

Desired slot length

Speaker release

No

Authors: Mr DELEON, Calos (Stony Brook University); Mr LEONARDI, James (Brookhaven National Laboratory); HUANG, Qiulan (Brookhaven National Laboratory (US))

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

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

Session Classification: Storage & Filesystems

Track Classification: Storage & Filesystems