DTIO: Unifying I/O for HPC and AI

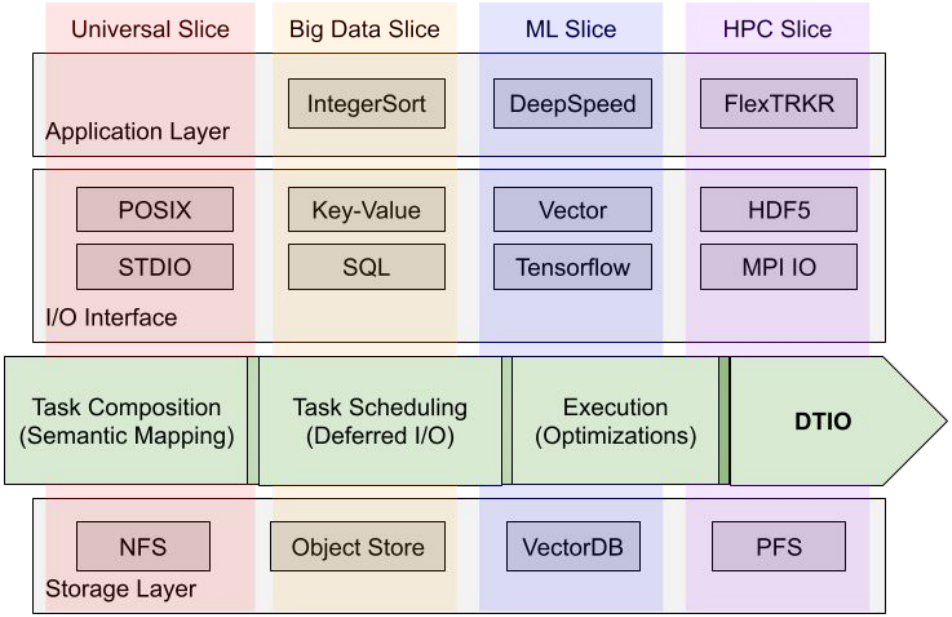
Speaker: Keith Bateman, Illinois Tech
Advisors: Anthony Kougkas, Xian-He Sun
Collaborators: Bogdan Nicolae, Franck Cappello

Introduction: Problem Statement

- HPC, Big Data Analytics, and Machine Learning have become increasingly intertwined, as popular models have driven discovery in scientific fields (e.g., GenSLMs, OpenFold).
- HPC I/O infrastructure involves a Parallel File System and MPI-IO or HDF5, while ML workloads may utilize tensors or a distributed vector database, and Big Data tends to utilize key-value and object stores.
- Different stacks have different semantics and functionality, therefore...
- **Problem Statement**: There is a need for a system which unifies the existing I/O stacks for the triple convergence of HPC, Big Data, and ML.

Methodology

- DTIO utilizes **Task Composition**, **Task Scheduling**, and **Execution** components to integrate various I/O interfaces and storage solutions.
- **Unified I/O**: Utilize a *task-based* infrastructure in order to provide a scalable and flexible data representation that can easily encapsulate existing interfaces.
- **Deferred Consistency Model**: DataTasks should be executed as-needed, which should ideally occur during compute, so that performance can be improved compared to strict interfaces like POSIX.
- **Hierarchical Integration**: DataTasks should be copied across different destinations to improve locality for integrated workflows with online and offline consumers.



Anticipated Benefits

- Preliminary results show that overheads are comparable to direct I/O.
- While we do not expect *unified I/O* to contribute a performance increase, it will increase the accessibility of data across domains (i.e., consumers can access data that they may not have the requisite interface for).
- We expect *deferred consistency* to give memory performance in the best case and, in the worst case, disk performance.
- We expect *hierarchical integration* to improve performance by permitting direct in-memory interface conversion for online consumers and utilizing faster storage for offline consumers.