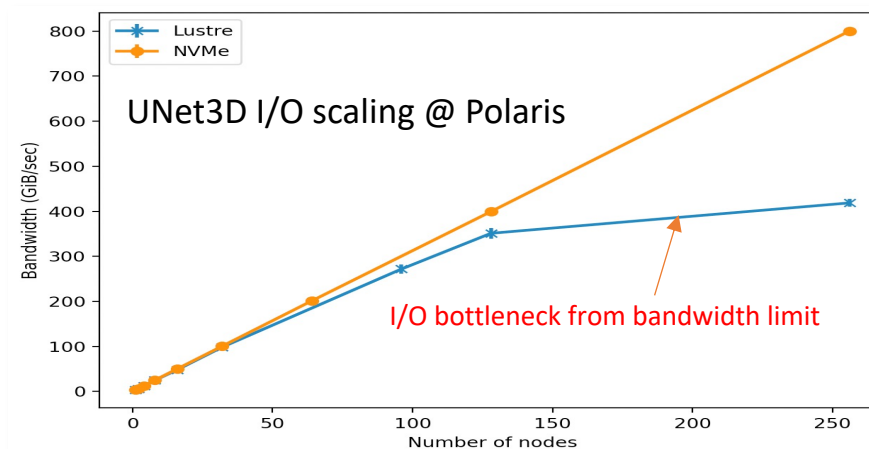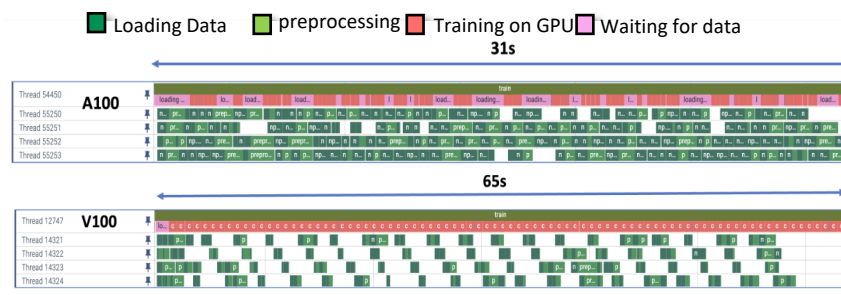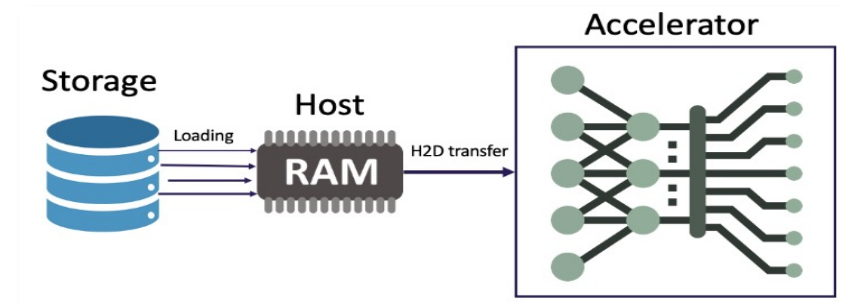# Data management, I/O and scaling for Deep Learning

Huihuo Zheng - Computer Scientist (LCF)

Deep learning I/O profiling and benchmarking

- **Developing profiling tools** to characterize data movement pattern and identify bottleneck - DLIO profiler

- **Benchmarking I/O performance** for different workloads on different file systems, dataset formats, storage, file organizations – DLIO benchmark, MLPerf Storage Working Group

- **Developing data pipeline library** for efficient data pipeline supporting parallel I/O (universal caching engine)

- **Optimizing data movement to support complex workflows**

https://github.com/argonne-lcf/dlio_benchmark.git
https://mlcommons.org/en/groups/research-storage/
https://github.com/hpc-io/vol-cache

Data management, I/O and scaling for Deep Learning

Huihuo Zheng - Computer Scientist (LCF)

Scaling Deep learning at exascale

- Profiling & benchmarking distributed frameworks and communication libraries (Horovod, DDP, DeepSpeed; oneCCL, NCCL)

- Working with science teams to scale their applications

Collaboration opportunities

- Optimizing I/O intensive workloads

- Use cases with complex workflows

- Minyang Tian, et al, Physics-inspired spatiotemporal-graph AI ensemble for gravitational wave detection, arXiv:2306.15728
- Chaturvedi, P.; Khan, A.; Tian, M.; Huerta, E. A.; Zheng, H. Inference-Optimized AI and High Performance Computing for Gravitational Wave Detection at Scale. *Frontiers in Artificial Intelligence* 2022, *5*
- Khan, A.; Huerta, E. A.; Zheng, H. Interpretable AI Forecasting for Numerical Relativity Waveforms of Quasicircular, Spinning, Nonprecessing Binary Black Hole Mergers. *Physical Review D* 2022, *105* (2), 024024.