Contribution ID: **200** Type: **Talk**

® I/O behavior of scientific deep learning workloads

Wednesday, 6 December 2023 14:10 (20 minutes)

Deep learning has been shown as a successful method for various tasks, and its popularity results in numerous open-source deep learning software tools. Deep learning has been applied to a broad spectrum of scientific domains such as cosmology, particle physics, computer vision, fusion, and astrophysics. Scientists have performed a great deal of work to optimize the computational performance of deep learning frameworks. However, the same cannot be said for I/O performance. As deep learning algorithms rely on big-data volume and variety to effectively train neural networks accurately, I/O is a significant bottleneck on large-scale distributed deep learning training.

In this talk, I aim to provide a detailed investigation of the I/O behavior of various scientific deep learning workloads running on the Theta cluster at Argonne Leadership Computing Facility. In this talk, I present DLIO, a novel representative benchmark suite built based on the I/O profiling of the selected workloads. DLIO can be utilized to accurately emulate the I/O behavior of modern scientific deep learning applications. Using DLIO, application developers and system software solution architects can identify potential I/O bottlenecks in their applications and guide optimizations to boost the I/O performance leading to lower training times by up to 6.7x.

Student or Postdoc?

Primary author: DEVARAJAN, Hariharan (Lawrence Livermore National Laboratory)

Presenter: DEVARAJAN, Hariharan (Lawrence Livermore National Laboratory)

Session Classification: Special