



Contribution ID: 31

Type: **Talk**

Accelerating Big Data Processing and Associated Deep Learning on Modern HPC Clusters and Clouds

Monday, 4 December 2017 14:30 (30 minutes)

The convergence of HPC, Big Data, and Deep Learning is becoming the next game-changing business opportunity. Apache Hadoop, Spark, gRPC/TensorFlow, and Memcached are becoming standard building blocks in handling Big Data oriented processing and mining. Modern HPC bare-metal systems and Cloud Computing platforms have been fueled by the advances in multi-/many-core architectures, RDMA-enabled networking, NVRAMs, and NVMe-SSDs during the last decade. This talk will provide an in-depth overview of the architecture of Hadoop, Spark, gRPC/TensorFlow, Memcached, and the associated Deep Learning frameworks. We will examine the challenges in re-designing networking and I/O components of these middleware with modern interconnects, protocols (such as InfiniBand, RoCE) and storage architectures. This talk will provide case studies of the new designs for several Hadoop/Spark/gRPC/TensorFlow/Memcached components and their associated benefits. Through these, we will also examine the interplay between high-performance interconnects, storage (HDD, NVM, and SSD), and multi-core platforms (e.g., Xeon x86, OpenPOWER) to achieve the best solutions for these components and applications on modern HPC clusters and clouds. We also present in-depth case-studies with modern Deep Learning tools (e.g., Caffe, TensorFlow, BigDL) running over RDMA-enabled Hadoop, Spark, and gRPC.

HPC content

The convergence of HPC, Big Data, and Deep Learning is becoming the next game-changing business opportunity. Modern HPC bare-metal systems and Cloud Computing platforms have been fueled by the advances in multi-/many-core architectures, RDMA-enabled networking, NVRAMs, and NVMe-SSDs during the last decade. This talk will discuss how to re-design networking and I/O components of the popular Big Data processing middleware and associated Deep Learning frameworks with modern high-performance interconnects, protocols (such as InfiniBand, RoCE) and storage architectures over HPC clusters and clouds.

Primary author: Dr LU, Xiaoyi (The Ohio State University)

Presenter: Dr LU, Xiaoyi (The Ohio State University)

Session Classification: Cognitive Computing and Machine Learning

Track Classification: HPC Techniques and Computer Science