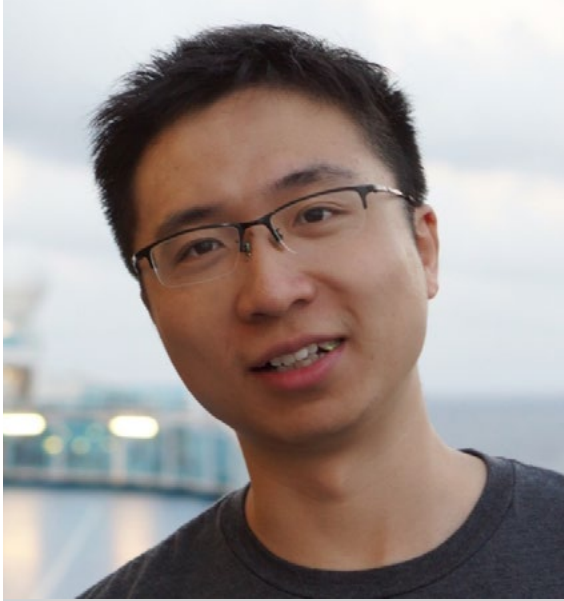




DISSERTATION DEFENSE



JUNCHENG GU

Efficient Resource Management for Deep Learning Clusters

Monday, August 2, 2021

10:00am - 12:00pm

[Virtual](#) (Passcode: 946819)

ABSTRACT: Deep Learning (DL) is gaining rapid popularity in various domains, such as computer vision, speech recognition, etc. With the increasing demands, many clusters, in both public and private clouds, have been established to run DL jobs (e.g., preparing datasets, and model training). However, the resource management techniques in those DL clusters have not been adapted to the new features of DL jobs, which leads to resource inefficiency and may dramatically hurt jobs' performance.

To narrow this gap, this thesis proposes a suite of resource management techniques for DL clusters to enhance resource efficiency and performance of DL jobs. Before model training, application-specific datasets have to be prepared from a large volume of raw data. Those data-processing jobs require lots of memory resource for good performance. We design INFINISWAP, a decentralized memory disaggregation system for those memory-intensive applications. It can opportunistically harvest and expose unused memory in the cluster to the machines who are running out of local memory. Model training is compute-intensive and requires powerful and expensive GPUs. It is common to perform distributed DL training to leverage multiple GPUs in parallel. To bring better training performance using limited GPU resource in a DL cluster, we present TIRESIAS, a GPU cluster manager tailored for distributed DL training jobs. By considering the job-level information and the cluster status, it can efficiently schedule and place DL training jobs to reduce their job completion time (JCT). Other than GPU resource, lots of CPU resource are allocated for model aggregation in distributed DL training. However, the bursty model aggregation tasks can not keep the assigned CPUs saturated all the time. We propose AUTOPS, an elastic model aggregation framework. It provides a shared model aggregation service to all the training jobs in the cluster, and brings higher CPU utilization without affecting training performance too much.

CO-CHAIRS: Profs. Kang Shin and Mosharaf Chowdhury