

CS 491 Undergraduate Research

Sergio Servantez
Computer Science Department
Illinois Institute of Technology
Chicago, Illinois 60616
sservantez@hawk.iit.edu
Advisor: Dr. Zhiling Lan

Report 3: 3/3/19

The last couple weeks have been all about building knowledge on deep neural networks (DNN) and learning how to execute them on distributed systems. I spent a week at the SIAM CSE conference learning about adjusting hyperparameters to optimize performance and reduce communication costs on distributed systems. I have also begun diving into the TensorFlow and Horovod libraries. The materials I received from the Data Science Group at Argonne have proved to be very helpful in understanding how to use these libraries. These materials also helped me set up the shell environment so I am able to use the needed data science modules. As of today, I was able to submit our first deep neural network job on Theta. The job is still queued, but we should have the experiment results shortly. At this point, I am only trying to determine that the DNN training is executing properly. As such, I only ran the job on 128 nodes (1 PPN). Once we have determined that everything is running properly, we can begin scaling up and adjusting the hyperparameters.

Over the next couple weeks, I need to determine which tool or tools will best serve the project for profiling internode communicate, memory usage and power consumption of the DNN applications. My understanding is that all communications being executed by Horovod are using MPI messages. I believe Horovod has a built in tool for profiling these communications. This will be the first place I look. Once we've selected a profiling tool, I want to start running some additional tests with varying hyperparameters.