

FA18 - ISL project

Objective: Enable system-wide file I/O and memory usage profiling for deep learning frameworks.

Approach: Write a script that collects all relevant data from /proc. Users can call it when running their deep learning applications.

How to get started: Here are two scripts ([memprof2.sh](#), [plotone2.sh](#)) that were developed ~3 years ago (or more) to collect and visualize the relevant data. They possibly need to be updated in order to work with the latest version of /proc.

To Do:

1. Request accounts on [Nano cluster](#).
2. Copy scripts to nano and learn how to use them on simple batch-scheduled applications
3. (possibly update scripts, if needed)
4. Go over TensorFlow getting started tutorial (<https://www.tensorflow.org/tutorials/>).
5. Figure out how to use the above scripts with TensorFlow tutorial examples, collect and visualize data.
6. Collect data from TensorFlow benchmarks (<https://www.tensorflow.org/guide/performance/benchmarks>)
7. Look into integrating the scripts with nano's system monitor (<https://nano.ncsa.illinois.edu:3000/d/3QVrDIFmz/nano-status?refresh=1m&orgId=1>).
 - Yan: InfluxDB takes HTTP requests with JSON-formatted data. Also check relevant Telegraf plugins (cpu, mem, disk, diskio, zfs, etc.) as these are currently being used in admin (non-public) dashboards and may provide some of the desired metrics.