COL862: Advanced Topics in Operating Systems

Summary
TensorFlow: A System for Large-Scale Machine Learning

Agastya Nanda
2016ANZ8513

Abstract

TensorFlow is a Machine Learning system that operates in large scale distributed environments where the machines may have CPUs, GPUs and TPUs (ASIC - Tensor Processing Unit). It allows for development of a variety of applications with a focus on training and inference on Deep Neural Networks. It uses dataflow graphs to represent computation, shared state and operations that mutate the state.

1 Keywords

Machine Learning, Heterogenous Systems, Distributed Systems

2 Background

2.1 Parameter Server

- Disjoint set of processes: Stateful Parameter Server, Stateless Worker Threads
- Parameter Servers maintain current version of model parameters
- Workers compute updates independantly and write back ”delta” updates to each parameter server
- Parameter Servers combine updates with current state

2.2 DistBelief

Predecessor of TensorFlow

- Parameter Server Model
- User defined Neural Network as DAG of layers and loss function
- Refining training algorithms or defining new ones is complex
- Support for GPU acceleration added later

3 Motivation

- High level scriptig interface for ease of use
- flexibility to experiment with different algorithms
- Common abstraction for heterogenous accelerators

4 Related Work

4.1 Single Machine Systems

- Caffe - Neural Networks on CPU+GPU. Easy to compose existing layers. Hard to add new layers.
- Theano - Similar to TensorFlow, but for single machine
- Torch - Allows for low level optimizations
- DryadLINQ - High level query language for sophisticated algorithms like MapReduce
- Spark - Extension of DryadLINQ with in-memory optimizations
- Dandelion - DryadLINQ + Code generation for GPU, FPGA

5 Execution Model

5.1 DataFlow Graph

- Nodes - Computation a.k.a. operations
- Edges - Input/Output to Node a.k.a tensors (n-dimension arrays)

5.2 Other Execution Model Details

- Specify Subgraphs to execute
- Executes independent operations in parallel
- Computation automatically partitioned across multiple devices
- Mutable State and Coordination via queues
- *kernels* executed on devices. Can have multiple kernels on a device with specialised implementations for a particular device (CPU/GPU/TPU) or data type
- Deferred executions va dataflow graph to offload chunks of work to accelerators
- Parameters distributed among **PS** tasks

6 Extensibility

- Embedding matrix used to convert large sparse matrices to smaller dense matrices
- User-level checkpointing for fault-tolerance
- Synchronous Replica coordination:
  - Asynchronous is scalable but can get stale values
  - Synchronous does not get stale values but can be slowed down by stragglers (slow processes)
  - Mitigate stragglers using backup workers

7 Test Your Understanding

- What are the disadvantages of the Parameter Server Model?
- How do backup workers mitigate stragglers?