Wrangler: Predictable and Faster Jobs in Distributed Processing Systems using Machine Learning

Neeraja J. Yadwadkar (neerajay@eecs.berkeley.edu), Bharath Hariharan, Ganesh Ananthanarayan, Joseph Gonzalez, and Randy Katz

http://www.istc-cc.cmu.edu/
Parallel Data Analytics

Job queue

Master

Slaves

Job completed
Stragglers
Impact of Stragglers:

We measure the potential in speeding up jobs in the trace using the following crude analysis: replace the progress rate of every task of a phase that is slower than the median task with the median task’s rate. If this were to happen, the average completion time of jobs improves by 47%, 29% and 36% in the Facebook, Bing and Yahoo! traces, respectively; small jobs ($\leq 10$ tasks) improve by 49%, 38% and 41%.

Dolly, NSDI’13
Speculative Execution

Job queue

Replicating

Job completed

Wasted Resources

Wasted Time in detecting stragglers

$T_S$: In progress
Existing Approaches

- Wasted Resources
- Wasted Time in detecting stragglers

Approaches:
- Replicate
  - Speculative Execution (OSDI’08)
- Wait
  - LATE (OSDI’08)
- Dolly (NSDI’13)
- Mantri (OSDI’10)
- Wrangler
Our proposal: Wrangler [SoCC’14]
Load-Balancing with Wrangler

Workload: FB2010
Wrangler Improves Job Completions

Workload: CC_b

Baseline: Speculative Execution

But, we built a model for every node!
However....

Real-world production clusters could contain over 1000 nodes

- Scalability!
  - Need to train too many models separately
  - Prohibitively long training data capture duration

Idea

Share data across nodes and workloads: Multi Task Learning [SDM’15]
Further Improved Job Completions

Workload: FB2009

We need only a sixth of training data!

....i.e., 4 hours → 40 minutes!!
Wrangler: Predictable and Faster Jobs in Distributed Processing Systems using Machine Learning

Neeraja J. Yadwadkar (neerajay@eecs.berkeley.edu), Bharath Hariharan, Ganesh Ananthanarayanan, Joseph Gonzalez, and Randy Katz

http://www.istc-cc.cmu.edu/