Characterization of Hadoop Jobs using Unsupervised Learning

Sonali Aggarwal, Shashank Phadke, Milind Bhandarkar

[sonali@cs.stanford.edu, {sphadke, milindb}@yahoo-inc.com]
Hadoop at Yahoo!

- Behind Every Click!
- 38,000+ Servers
- Largest cluster is 4000+ servers
- 1+ Million Jobs per month
- 170+ PB of Storage
- 10+ TB Compressed Data Added Per Day
- 1000+ Users
Hadoop Growth

- 38K Servers
- 170 PB Storage
- 1M+ Monthly Jobs

Graph shows the growth of Hadoop servers and storage from 2006 to 2010, reaching 38K servers and 170 PB storage.
Hadoop Clusters

- Hadoop Dev, QA, Benchmarking (10%)
- Sandbox, Release Validation (10%)
- Science, Ad-Hoc Usage (50%)
- Production (30%)
Benchmarking Hadoop - Part I

- Micro benchmarks
  - Sort
  - TestDFSIO
  - SmallJobs
  - NNBBench
  - ...

Benchmarking Hadoop - Part 2

- Application Kernels - GridMix - 1
- Most common “real” applications
- Machine Learning, ETL, Graph Algorithms
- Synthetic data
GridMix

- GridMix Version 2
- GridMix Version 1 + Load Generator
- Synthetic Scheduling
- GridMix Version 3
- Real Scheduling + Synthetic Load Generator
GridMix V3

- Collect production job metrics + Submission statistics
- Generate workload
- Execute on benchmarking cluster
  - With different versions of Hadoop
GridMix Problems

- Benchmarking cluster has less resources than production clusters
- Impossible to simulate large intervals
- Different production clusters have different workload characteristics
- Workload characteristics change over time
Our Solution

• Insight: Most applications are executed periodically
• Automatically determine prominent application characteristics
• Simulate only prominent applications
Our Approach

• Fetch Job metrics, and Job configurations (job.xml)
• Extract “features”
• Determine number of job clusters
• Find centroids of job clusters
## Counters for attempt_200904270516_5709_m_000000_0

### FileSystemCounters
- FILE_BYTES_READ: 15,201,458
- HDFS_BYTES_READ: 143,611,125
- FILE_BYTES_WRITTEN: 25,452,741

### Map-Reduce Framework
- Combine output records: 2,755,686
- Map input records: 1,717,449
- Spilled Records: 2,755,686
- Map output bytes: 200,041,337
- Combine input records: 16,226,718
- Map output records: 14,560,207
Job Metrics

- Number of Map Tasks
- Number of Reduce Tasks
- Slots per task
- InputFormat / OutputFormat
- Type of output and intermediate data compression
Task Metrics

- HDFS Bytes Read/Written
- File Bytes Read/Written
- Combiner Records Ratio
- Shuffle Bytes per Reduce Task
- ...
- Mean & StdDev across tasks: Job Metric
Job Clustering

- Rescale features
- Find and eliminate correlated features
- Random sampling of jobs
  - \( \sqrt{N} \)
- Determine Number of clusters
  - Minimize “within group sum of squares”
Feature Vectors

- Numeric Features
  - e.g. Number of tasks
  - Rescale, mean=0, stddev=1

- Nominal Features
  - e.g. Type of Compression
Number of Clusters
K-Means Clustering

• Find initial seeds
• Hierarchical Agglomerative Clustering
• Euclidean distance between jobs
Job Clusters
Centroids
Largest Cluster

- nMaps: 79, nReduces=28
- HDFS Bytes read / Map = 44.82 MB
- HDFS Bytes written / Reduce = 54.85 MB
- Input Records per Map = 334K
- Output Records per Reduce = 235K
Validating GridMix v3

- Clustering real workload
- Clustering corresponding GridMix3 workload
- Compare clusters
- Results in same number and sizes of clusters
Conclusions

• GridMix v3 accurately simulates real workload
• Number of prominent types of Hadoop jobs on Yahoo! production cluster is 8
• Job Clustering can be used to generate benchmark suites
Future Work

• Automate the process
• Incorporate in GridMix version 4
• Extend to Pig job chains
• Extend to Oozie workflows
• Contribute to Open Source