NOVEL CLASSIFICATIONS OF SCHEDULING STRATEGIES IN BIG DATA SYSTEMS

Xiao Meng
CS846 Fall 2016

Photo credit: Matthew Munoz
OUTLINE

• Introduction
• Motivation
• Research Questions
• Research Methodology
• Goal & Progress & Contribution
• Future Work
• References
INTRODUCTION

- Big Data Systems: Hadoop, Spark
- Big Data Infrastructure: racks, rooms, data centers
- MR-style computation: tenants, jobs, tasks

Photo credit: Google
INTRODUCTION - A MOTIVATING EXAMPLE

• “On average, the shuffle phase accounts for 33% of the running time in these jobs. In addition, in 26% of the jobs with reduce tasks, shuffles account for more than 50% of the running time, and in 16% of jobs, they account for more than 70% of the running time.”

• With optimized schedule, however, a 1.9X overall improvement is observed.

Credit: Chowdhury, M., et al. Managing data transfers in computer clusters with orchestra. In SIGCOMM 2011, ACM.

Figure 1: CDF of the fraction of time spent in shuffle transfers in Facebook Hadoop jobs with reduce phases.
MOTIVATION

• Scheduling algorithms play a critical role in managing clusters and meeting quality requirements

• Scheduling algorithms achieve various goals by controlling the order and distribution of users/tenants, jobs and tasks execution

• Given the rise of Spark, and emergence of workload patterns, what is the landscape of schedulers now?
RESEARCH QUESTIONS

• What are the common requirements that scheduling algorithms are optimizing for, and how they are solved with TCS?

• How the rise of Spark and emergence of workload patterns have affected scheduling research?

• How can we make sense of the entire space of scheduling algorithms for big data systems, and make suggestions accordingly?
RESEARCH METHODOLOGY

• Literature collection:
  • Via academic search engines: Google Scholar
  • Via digital libraries: ACM/IEEE/Springer Library
  • Via targeted conferences/journals: Databases, Cluster Computing, Operating Systems…
RESEARCH METHODOLOGY

• Classification methodology:
  • Multi-dimensional framework

• Dimension selection:
  • Extracted from previous work [1]
  • Based on the goal of this project
GOAL

With this project, ideally achieving the following goals:

• To gain a comprehensive understanding of the landscape for schedulers in big data systems

• To concisely and accurately classify schedulers based on various dimensions

• To make reasonable recommendations for choosing schedulers in a systematic fashion
PROGRESS

• Four dimensions selected

• 1. Optimization Goal
  • Fairness/Availability/Data locality
  • Energy efficiency/Resource utilization

• 2. Scheduling entities
  • Users/Jobs/Map tasks/Reduce tasks
PROGRESS

• Four dimensions selected

• 3. Hadoop/Spark Availability
  • Hadoop/Spark/Both/Others/None

• 4. Workload Pattern
  • Dependency-aware/Dependency-oblivious
  • Dependency-aware: SS(shared scan)/DAG (directed acyclic graph)
# PROGRESS

<table>
<thead>
<tr>
<th></th>
<th>DAG-aware</th>
<th>SS-aware</th>
<th>Dependency-oblivious</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Both</strong></td>
<td>Quartet [3]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>None</strong></td>
<td></td>
<td></td>
<td>BalancedPools [4]</td>
</tr>
</tbody>
</table>
With this project, several contributions can be claimed:

- The first survey to consider the scheduler availability in different big data systems
- The first survey to include workload-specific schedulers
- A manual to follow for big data system admins, and an overview to reference for big data researchers
FUTURE WORK

• Short term:
  • Consider adding a new dimension: centralized/distributed scheduler
  • Complete a comprehensive classification of algorithms based on four dimensions, with emphasis on Spark availability and dependency-awareness
  • Propose a recommendation workflow (e.g., a flowchart) for big data systems users
FUTURE WORK

• Long term:
  • Implementation of cross-platform schedulers
  • Empirical study of performance in a real cluster
  • Propose new schedulers for a combination of goals
REFERENCES


(Only slides-related papers are listed here)