# A Study of Big Data Computing Platforms: Fairness and Energy Consumption

Zhaojie Niu, Bingsheng He
*Nanyang Technological University, Singapore*

## I. Introduction

In the current era of big data, data-intensive computing is a common paradigm in clusters and clouds. A lot of large-scale big data computing platforms have thereby emerged and become popular recently [1], [2], [3], [4]. Performance, fairness and energy consumption are three important concerns for cluster providers and users in the large-scale multi-tenant environments.

- Performance is de facto the very important factor in the big data platforms.
- Fairness is a key building block of any shared computing system which allows resource sharing more effectively.
- Energy consumption of the data center has reached 3% of all global electricity production while producing 200 million metric tons of $CO_2$ in 2014 [5]. In order to reduce carbon emission and financial burden on the electricity, a lot of data centers have been re-designed and powered with multiple energy sources, including renewable (green) energy from non-polluting sources and brown energy from traditional polluting sources [6].

Improving the performance is the common sense on those large-scale data processing frameworks and fruitful studies are proposed in this direction. In contrast, the fairness and energy consumption of those frameworks need further exploration and how the performance, fairness and energy consumption interact each other on big data computing frameworks is not well addressed. In our research, we study the fairness and the energy consumption of those big data computing systems. We find that there are tradeoff between these factors. We conduct detailed studies on the factors which impact the tradeoff between different factors. Based on the observations in our study, we propose workload-aware, energy-efficient and green-aware optimizations and implement them into Hadoop YARN. Particularly, in this thesis proposal, we propose to explore the following research problems. First, we explore the tradeoff between fairness and performance, and improve the performance of the state-of-the-art approach by up to 225% [7]. Second, we consider the energy efficiency, renewable energy supply as well as battery usage and reduce the brown energy consumption of existing systems by more than 25% [8]. Third, we will explore the relationship between fairness and energy consumption, and eventually we will develop multi-objective optimizations for performance, fairness and energy consumption.

## II. Related work

In this section, we review the related work on fairness and green energy usage.

### A. Fairness

Some fair schedulers [9], [10], [11] are proposed to resolve the resource fair allocation problem in the multi-tenant environment. These studies only consider single resource type in the cluster. In order to support fair allocation of multiple resources, new fairness definitions emerge. Dominant Resource Fairness applies the max-min fairness to multiple resource types in Hadoop YARN. Wang et al. [12] extend Dominant Resource Fairness especially for the heterogeneous environment. These studies resolve the resource fair allocation in different scenarios without considering the tradeoff between performance and fairness. Recently, some studies begin to consider this tradeoff. They theoretically analyze the fairness-efficiency tradeoff for different fairness definitions [13], [14]. Tetris [15] explores the performance-fairness tradeoff of Hadoop YARN from system view. Although these studies have observed the tradeoff between performance and fairness, the factors that impact this tradeoff is not explored in detail.

### B. Renewable energy-aware computing

There have been some research on green-aware scheduling systems. Some studies maximize the usage of renewable energy by delaying the execution of batch jobs [16], [17]. Goiri et al. have conducted a series of studies and develop a green data center prototype to manage both deferrable and non-deferrable workloads at the presence of renewable energy [18]. Chen et al. [19] propose ReinDB that integrates renewable energy supply into database systems. EU has founded a project called DC4Cities which proposes a technical and business related solution to optimize the usage of renewable energy in smart cities [20], [21]. Some attention has been paid to leverage battery to utilize the renewable energy efficiently [22]. Few of the previous studies have paid attention to the energy efficiency of the workload, particularly in the MapReduce/Hadoop cluster.

## III. Preliminary results

In this section, we first give an overview of my research work. Then, we introduce each of them in more details.

### A. *Overview*

Performance, fairness and energy consumption are the three key concerns in the cluster. We have observed that there is a tradeoff between these factors. We conduct detailed studies on the tradeoff and propose bi-criteria optimization algorithms to address the tradeoff between different factors, as shown in Figure 1. So far, we have proposed two frameworks which aim at the scenario ① and scenario ② which are shown in Figure 1. In scenario ①, we find that the tradeoff is related to the workload and propose a workload-aware scheduler which adaptively decides the most suitable scheduling policy at runtime according to the variation of the workload. In scenario ②, we consider the energy efficiency, renewable energy supply as well as battery usage in a MapReduce cluster and propose an energy-efficient and green-aware framework. Third, we will explore the scenario ③, and eventually we will develop multi-objective optimizations for performance, fairness and energy consumption.

Hadoop YARN supports plugable scheduling policies. Thus, we can easily implement and integrate our proposed scheduling algorithms into Hadoop YARN. We use complementary experimental approaches to evaluate our systems with both real deployment in a local cluster and trace-driven simulation. This cluster consists of 10 compute nodes, each with two Intel X5675 CPUs (6 CPU cores per CPU with 3.07 GHz), 24GB DDR3 memory and 500GB 7200RPM disk drivers. The nodes are connected with 10Gb/sec Ethernet. We synthesize a micro benchmark according to the daily pattern at Facebook [23]. The actual jobs are randomly generated from Hive benchmark [24], containing four types of applications, i.e., rankings selection, grep search (selection), uservisits aggregation and rankings-uservisits join. In order to evaluate our system at larger cluster, we implement a trace-driven simulator that replays the production traces from Google cluster [25]. In the following, we briefly present our preliminary studies on scenario ① and ②.

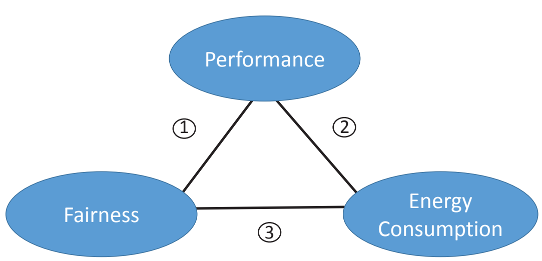


Figure 1: The research scope of my thesis proposal.

### B. *Gemini [7] (Scenario ① in Figure 1)*

We have observed that there is a tradeoff between performance and fairness because of resource contention between users/jobs and the tradeoff is related to the workload. In order to address this tradeoff, we propose an workload-aware scheduler called Gemini for Hadoop YARN. The system overview of Gemini is shown in Figure 2. We first develop a workload characteristic model with the regression approach to estimate the performance improvement and the fairness loss given the workload. Next, we leverage this model to guide the resource allocation at runtime to optimize the performance of the cluster given user-defined threshold on fairness loss. Particularly, instead of using a static scheduling policy, Gemini adaptively decides the most suitable policy with the variation of the current running workload. We implement Gemini in Hadoop YARN. We compare the performance improvements of both schedulers under the same fairness loss with the Facebook workload. Gemini achieves up to 125% performance improvement compared with Tetris [15] at the cost of the same fairness loss. The performance gain is achieved by taking special considerations on the variation of the performance-fairness tradeoff during the computation. Gemini applies an adaptive scheduling policy according to the variation of the workload. Instead, Tetris utilizes a static approach over the entire execution of the workload. More experimental results can be found in our paper [7].

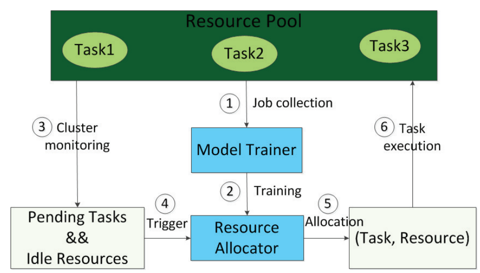


Figure 2: System overview of Gemini. (The figure is adopted from [7]).

### C. *GreenMR [8] (Scenario ② in Figure 1)*

Interests have been growing in integrating renewable energy into data centers, which attracts a lot of research efforts in developing green-aware systems and algorithms. However, little attention was paid to the efficiency of each joule expended by data center workloads. Actually, not all joules are equal in the sense that the amount of work that can be done by a joule varies significantly in data centers. Ignoring this fact results in great energy waste. In this thesis proposal, we investigate how to leverage such joule efficiency to maximize the profits of renewable energy for MapReduce framework. We develop energy-efficient and green-aware scheduling algorithms with a particular focus on the critical aspects of joule efficiency in a MapReduce cluster, including energy efficiency of MapReduce workloads, the supply of renewable energy and the battery using. We further develop an effective performance-energy cost model to guide our green-aware scheduling. We implemented GreenMR, an energy-efficient and green-aware MapReduce framework, on top of Hadoop YARN.

The architecture of our system is shown in Figure 3. The system comprises of a Hadoop cluster, a charge controller, an inverter, batteries and a switch. The charge controller monitors the charing/discharing operations and the switch is connected with both green and brown sources. GreenMR performs cost-aware optimizations to minimize the brown energy consumption of the cluster. It first generates a basic execution plan with special considerations on the joule efficiency of MapReduce workload. Next, it leverages green-aware job transformations and battery assisted green shifting algorithms to further reduce the brown energy consumption of the execution plan. Compared with GreenHadoop [17], GreenMR achieves more than 25% reduction on the brown energy consumption, by embracing these energy-efficient optimizations. More experimental results can be found in our paper [8].

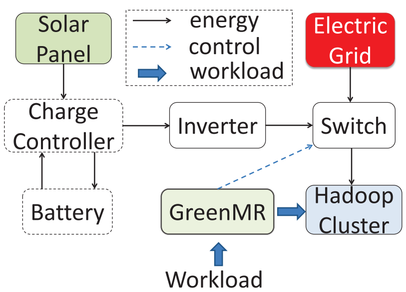


Figure 3: Architecture of GreenMR. (The figure is adopted from [8]).

## IV. Summary

This thesis proposal describes the scope of our research work which mainly aims at the fairness and energy efficiency of the large-scale data computing frameworks. Our preliminary studies show the interplay between fairness, performance and energy efficiency and performs workload-aware, energy efficient and green-aware optimizations on top of Hadoop YARN. The experimental results demonstrate the effectiveness of our proposed systems. In the future, we plan to extend our study in two major aspects; 1) study the tradeoff between the fairness and energy consumption; 2) extend our study to geo-distributed environment.

## V. Acknowledgment

This work is supported by the National Research Foundation Singapore under its Interactive Digital Media (IDM) Strategic Research Programme, a MoE AcRF Tier 2 grant (MOE2012-T2-2-067) and a MoE AcRF Tier 1 grant (MOE 2014-T1-001-145) in Singapore.

## References

[1] J. Dean and S. Ghemawat, "Mapreduce: simplified data processing on large clusters," *OSDI*, 2004.
[2] B. Hindman, A. Konwinski, M. Zaharia, A. Ghodsi, A. D. Joseph, R. H. Katz, S. Shenker, and I. Stoica, "Mesos: A platform for fine-grained resource sharing in the data center." in *NSDI*, 2011.
[3] V. K. Vavilapalli, A. C. Murthy, C. Douglas, S. Agarwal, M. Konar, R. Evans, T. Graves, J. Lowe, H. Shah, S. Seth *et al.*, "Apache hadoop yarn: Yet another resource negotiator," in *SoCC*, 2013.
[4] M. Zaharia, M. Chowdhury, T. Das, A. Dave, J. Ma, M. McCauley, M. J. Franklin, S. Shenker, and I. Stoica, "Resilient distributed datasets: A fault-tolerant abstraction for in-memory cluster computing," in *NSDI*, 2012.
[5] "Industry outlook: Data center energy efficiency http://www.datacenterjournal.com/industry-outlook-data-center-energy-efficiency/."
[6] R. Bianchini, "Leveraging renewable energy in data centers: present and future," in *HPDC*, 2012.
[7] Z. Niu, S. Tang, and B. He, "Gemini: An adaptive performance-fairness scheduler for data-intensive cluster computing." in *CloudCom*, 2015.
[8] Z. Niu, B. He, and F. Liu, "Not all joules are equal: Towards energy-efficient and green-aware data processing frameworks." in *IC2E*, 2016.
[9] "Hadoop mapreduce 1.0 - fair scheduler," http://hadoop.apache.org/docs/r1.2.1/fair_scheduler.html.
[10] M. Isard, V. Prabhakaran, J. Currey, U. Wieder, K. Talwar, and A. Goldberg, "Quincy: fair scheduling for distributed computing clusters," in *SOSP*, 2009.
[11] A. Ghodsi, M. Zaharia, S. Shenker, and I. Stoica, "Choosy: Max-min fair sharing for datacenter jobs with constraints," in *Eurosys*, 2013.
[12] W. Wang, B. Li, and B. Liang, "Dominant resource fairness in cloud computing systems with heterogeneous servers," in *INFOCOM*, 2014.
[13] C. Joe-Wong, S. Sen, T. Lan, and M. Chiang, "Multiresource allocation: Fairness–efficiency tradeoffs in a unifying framework," *INFOCOM*, 2012.
[14] W. Wang, C. Feng, B. Li, and B. Liang, "On the fairness-efficiency tradeoff for packet processing with multiple resources," in *CoNEXT*, 2014.
[15] R. Grandl, G. Ananthanarayanan, S. Kandula, S. Rao, and A. Akella, "Multi-resource packing for cluster schedulers," in *SIGCOMM*, 2014.
[16] Í. Goiri, K. Le, M. E. Haque, R. Beauchea, T. D. Nguyen, J. Guitart, J. Torres, and R. Bianchini, "Greenslot: scheduling energy consumption in green datacenters," in *SC*, 2011.
[17] Í. Goiri, K. Le, T. D. Nguyen, J. Guitart, J. Torres, and R. Bianchini, "Greenhadoop: leveraging green energy in data-processing frameworks," in *Eurosys*, 2012.
[18] Í. Goiri, W. Katsak, K. Le, T. D. Nguyen, and R. Bianchini, "Parasol and greenswitch: Managing datacenters powered by renewable energy," in *ACM SIGARCH Computer Architecture News*, 2013.
[19] C. Chen, B. He, X. Tang, C. Chen, and Y. Liu, "Green databases through integration of renewable energy." in *CIDR*, 2013.
[20] S. Klingert, F. Niedermeier, C. Dupont, G. Giuliani, T. Schulze, and H. de Meer, "Renewable energy-aware data centre operations for smart cities–the dc4cities approach."
[21] E. CREATE-NET and P. UNI, "Dc4cities," 2014.
[22] S. Govindan, A. Sivasubramaniam, and B. Urgaonkar, "Benefits and limitations of tapping into stored energy for datacenters," in *ISCA*, 2011.
[23] "Facebook swim trace," https://github.com/SWIMProjectUCB/SWIM.
[24] "Hive performance benchmarks," https://issues.apache.org/jira/browse/HIVE-396.
[25] "Google cluster data," https://code.google.com/p/googleclusterdata.