

# Energy Efficient Resource Usage in Data Centers: GreenDoop

D. Fernández, A. Fernández-Montes, Juan A. Ortega

Departamento de Lenguajes y Sistemas Informáticos

Universidad de Sevilla

Spain

damianfdezcerero@gmail.com, afdez@us.es, jortega@us.es

## Abstract

The spread of cloud computing has increased recent data center size and usage. Because of this, the optimization of data centers is becoming the focus of most important companies in this sector. In order to improve data center energy efficiency, we propose a software infrastructure for minimizing power usage of data centers based on policies for switching on/off servers, load balancing and data organization; keeping the minimum number of machines on, while affecting as little as possible performance and reliability.

## 1 Introduction

The growing interest in energy efficiency has led big companies to enhance its efforts in reducing power consumption of data center infrastructures due to the increase in data centers energy consumption in the past ten years, which is 1.5% of worldwide energy consumption [Kooomey, 2008; 2011]. Thanks to these efforts and researchers results, current energy consumption trends have slowed down in the past five years. Many of these efforts have been directed to areas unrelated to software, such as cooling [El-Sayed, 2012], hardware scaling [Fan, 2007] or power distribution [Femal, 2005]. This work, in contrast, focuses on software systems that support data centers, such as parallel computing systems or distributed file systems in order to develop a software product that minimizes the energy consumption.

## 2 Summary of Proposals

In order to create a real software system we have to choose an existing distributed computing framework to work with. That system is Apache Hadoop, and our goal is to modify it by creating a software solution to coordinate and minimize energy consumption at data center level by applying the designed energy efficiency policies to its two major modules: HDFS and YARN.

### 2.1 Modifications to HDFS

HDFS is responsible for serving data. In order to minimize the number of machines running needed to maintain redundancy benefits and data reliability we will apply different

modifications: **Restructuration of the Data Center** to categorize block replicas, **Design a new Balancing Policy** to increase server use and **Create a Buffer and Caching System** to handle versioning issues, prevent inconsistencies, hardware errors and bad response times.

### 2.2 Modifications to YARN

The purpose of YARN is running MapReduce applications and resource management [Tomasic, 2012]. According to queuing theory, the arrival of events to a set of resources is described by exponential distribution and the arrival of consecutive events to a set of resources is described by a Gamma distribution. Our proposal is to start from these distributions to fit the situation: to minimize energy consumption.

## 3 Expected Results

GreenDoop has the objective creating a highly configurable and useful enterprise tool that can notably reduce current Hadoop based data centers energy consumption.

## References

- [El-Sayed, 2012] Nosayba El-Sayed. Temperature management in data centers: Why some (might) like it hot. Technical report, Department of Computer Science, University of Toronto, 2012.
- [Fan, 2007] Xiaobo Fan. Power provisioning for a warehouse-sized computer. Technical report, Google Inc., 2007.
- [Femal, 2005] M.E. Femal. Boosting data center performance through non-uniform power allocation. In *Automatic Computing, 2005. ICAC 2005. Proceedings. Second International Conference on*. Dept. of Comput. Sci., North Carolina State Univ., 2005.
- [Kooomey, 2008] Jonathan Kooomey. Worldwide electricity used in data centers. *Environmental Research Letters*, 2008.
- [Kooomey, 2011] Jonathan G. Kooomey. Growth in data center electricity use 2005 to 2010. Technical report, Stanford University, 2011.
- [Tomasic, 2012] Ivan Tomasic. Bigdata and mapreduce with hadoop. In *Proceedings of the 1st International Conference on Cloud Assisted ServiceS*, 2012.