Chapter 37
Power Consumption Aware Cluster Resource Management

Simon Kiertscher
Potsdam Institute for Climate Impact Research, Germany

Bettina Schnor
University of Potsdam, Germany

Jörg Zinke
University of Potsdam, Germany

ABSTRACT

In 2007, the Green500 list was introduced, which compares supercomputers by performance-per-watt. Since supercomputers consist of thousands of nodes, energy-saving is a growing demand. Compute clusters are often managed by a so-called Resource Management Systems (RMS), which have load information about the whole system. For clusters with changing compute demands, this can be used to switch on/off nodes according to the current load situation and save energy this way. Here, the authors present energy-saving techniques that work on the management level and measurements that show that speed scaling is not a good means for energy saving. Further, they give an overview of some important standards and specifications related to energy saving, like ACPI and IPMI. Finally, the authors present their energy-saving daemon called CHERUB. Due to its modular design, it can operate with different Resource Management Systems. Their experimental results show that CHERUB's scheduling algorithm works well, i.e. it will save energy, if possible, and avoids state flapping.

INTRODUCTION

Clusters are a popular hardware platform for compute-intensive applications, but the power consumption of these machines has already reached an unacceptable amount. The High Performance Computing (HPC) community is aware of this conflict. As a complement to the list of the 500 fastest machines, the Green500 list (Green500, 2011) compares supercomputers by their performance-per-watt since November 2007. For example, in the Top500 (Top500, 2011) list of June 2010, the first place is held by the Jaguar cluster, which is located at the National Laboratory at Oak Ridge and was installed in 2009. It has 225.162 cores and a performance of 1.759.000

GFlops. That system requires 7MW under full load. This amount of energy results in costs of approximately 7 million USD per year (Dongarra, 2010). Because of the high-energy consumption, Jaguar is only ranked 56 in the Green500 list of June 2010. Only a half year later, Tianhe-1A overtakes Jaguar with 2.566.000 GFlops. This is nearly 33% more compute power, but Tianhe-1A also needs only 4MW, which are 42% less energy consumption. It is ranked 1 on the Top500 (Nov 2010) and rank 11 on the Green500 (Nov 2010), while Jaguar only gets ranked 2 on the Top500 and a weak rank of 88 in the Green500 in those Lists.

In 2006, all data and computing centers of the USA consumed about 61 billion kWh. This equates to about 4.5 billion USD and 1.5% of the whole American power consumption. The U.S. Environmental Protection Agency has raised this numbers and their prognosis is a doubling of these numbers until 2011 (USEPA, 2007). Thus, data and computing centers are starting to take care of their operations per joule. Besides massively reducing the input power by using lower powered embedded systems, a data center has the following options:

1. Virtualization of clusters: Aggregate a lot of weak machines into a few strong computers, and optionally simulate the cluster. The benefit is fewer machines require less energy, and result in less heat production. Thus, one requires less air conditioning systems, which usually require also a lot of energy.
2. Shutdown of unused hardware: Machines that are powered off do only consume a minimum of voltage and produce (nearly) no heat.

The second option seems to be very efficient but requires some kind of monitoring system to decide which machines are not needed and can be powered off. This is a strategy, which is not well suitable for clusters working. However, it is well suited for clusters in institutes with varying workloads (see section "Experimental Results").

This chapter presents a design for a daemon, which implements such a monitoring system. We start with explaining important background like Resource Management Systems and possibilities to remotely control a cluster. Then, we show two measurement series in which we not only observed how much energy a compute node of a cluster needs in various modes but also how efficient it is to clock down cores or use Hyperthreading. Next, we describe a modular design for an energy-saving daemon and an open source prototype implementation called CHERUB. Furthermore, we present test results, which show the benefits of CHERUB.

BACKGROUND

First, we give a short introduction into the field of Resource Management Systems. We will also present some energy aware research already done in this field. The concept of an energy-saving daemon is also suited for Server Load Balancing (SLB) clusters managed by dispatchers like Linux Virtual Server (LVS, 2010), since the only requirement is a central approach for resource management, but in this chapter we will concentrate on the High Performance Computing clusters.

Resource Management Systems

The general purpose of Resource Management Systems (RMS) in compute clusters is to handle distribution and execution of jobs/requests on cluster nodes. They all have in common a central approach, i.e., there exists a central decision component (daemon). The general architecture of a RMS is shown in Figure 1.

A RMS typically uses four hardware components:

- **Master Node:** Usually runs the server programs and the central scheduler of the RMS.

Related Content

Tourism and Its Impact on Quality of Life: Evidence from Kamarpukur, West Bengal, India
Nilanjan Rayand Dillip Kumar Das (2017). *International Journal of Sustainable Economies Management (pp. 51-66).*
www.irma-international.org/article/tourism-and-its-impact-on-quality-of-life/181253

Analysis of the Seasonal Variation of the Physical and Chemical Characteristics of the Water of the Roosevelt River, Amazonas, Brazil
Marcelo Rodrigues dos Anjos, João Ânderson Fulan, Nadja Gomes Machado, Paulo André Martins da Silva, Maria Clécia Gomes Sales, Elder Leite Barbosaand Harumy Sales Noguchi (2022). *International Journal of Social Ecology and Sustainable Development (pp. 1-14).*
www.irma-international.org/article/analysis-of-the-seasonal-variation-of-the-physical-and-chemical-characteristics-of-the-water-of-the-roosevelt-river-amazonas-brazil/290396

Value Added Tax Administration and its Implications for Tax Performance
Isaac Okoth Randa (2022). *International Journal of Sustainable Economies Management (pp. 1-24).*
www.irma-international.org/article/value-added-tax-administration-and-its-implications-for-tax-performance/309400

Evaluation of Labor Market Programs During Recession in the North Great Plain Region of Hungary
Andrea Szabo (2012). *International Journal of Sustainable Economies Management (pp. 38-51).*
www.irma-international.org/article/evaluation-labor-market-programs-during/75192

Evaluation and Visualization of Blockchain-Based Healthcare
Riten Bhagraand Gaurav Choudhary (2022). *Futuristic Trends for Sustainable Development and Sustainable Ecosystems (pp. 147-164).*
www.irma-international.org/chapter/evaluation-and-visualization-of-blockchain-based-healthcare/307675