Methods and Approaches for Energy Saving considered in the Ener-G project

[Extended Abstract]

Michel Steichen
University of Kaiserslautern
Germany
steichen@informatik.uni-kl.de

Bernd Reuther
University of Kaiserslautern
Germany
reuther@informatik.uni-kl.de

Gergö Lovasz
University of Passau
Germany
gergoe.lovasz@uni-passau.de

Paul Müller
University of Kaiserslautern
Germany
pmueller@informatik.uni-kl.de

1. INTRODUCTION

This extended abstract provides a short overview of the actual research work being tackeld in the Ener-G project with the main focus on preliminary results.

Ener-G is a research project in the second phase of German-Lab. G-Lab and Ener-G projects are both funded by the BMBF and will last until the first quarter of 2011. To provide a realistic possibility for testing and evaluation purposes, partner institutes of the G-Lab project were equipped with an high performance cluster.

In a first step Ener-G is intending to exemplarily reduce the energy consumption of these G-Lab clusters and thus to guarantee a sustainable and power-saving operation In a second step this research knowledge is used to optimize, in the sense of energy-saving, the development and deployment of servers and network infrastructure for the Future Internet.

Ener-G is subdivided in eight work packages from which two have been completed successfully and the third one is actually running. In the first work package the actual state of the art regarding energy efficiency was determined. The aim of the second work package was to take concrete measurements on the G-Lab clusters to identify the energy consumption of the hardware components and to figure out how useful the exisiting monitoring systems on G-Lab are. In the actual work package the focus is on the development of energy consumption models which can be used to systematically optimize the energy consumption.

In the first section a short overview of the state of the art with focus on energy saving is given. In the second section the problems occurring from some software solutions are pointed out. Finally an overview of the main characteristics and advantages, that can be gained from the employment of energy consumption models, is given.

2. STATE OF THE ART

This section is aiming to give an overview of existent mechanisms and research topics in the domain of energy efficiency. Methods for energy saving can be categorized in hardware and software solutions.

Hardware techniques for power saving are a manufacturer specific matter. Thus the customer's choice is dependent from the products offered by the hardware manufactures. Fortunately many hardware producer are thinking in terms of Green IT and therefore lots of power-saving technologies are actually available on the market. Most of these hardware techniques are implementing two main techniques to reduce the power consumption of their products:

A first method to save power is to dynamically adapt the tunable parameters of hardware components. For example the clock speed of the CPU can be dynamically changed considering the actual needed load and computational requirements of the user. If less computing power is required the clock speed is stepwise decreased. The other way round the clock speed is immediately increased to catch load peaks. Common technologies are for example the Cool'nQuiet and PowerNow solutions from AMD. Intel's CPUs use a similar technique branded SpeedStep and Demand-Based Switching.

A second common method for energy saving is to switch off unused functional units. The well known and common ACPI standard supports the feature to power down hardware components. Temporarily unused components like for instance RAM or hard drives are put in a sleep mode or completely

switched off.
A more fine-grained approach for powering down unused units is the power gating technology. The particular difference here is that special cut-off circuits included on the chip substrate are responsible for lowering the energy consumption.

The most promising software solution for energy optimization is virtualization. Several virtual machines are run on top of the server hardware. This way the hardware can be shared in a more optimal way between many users. However the most important advantage is that virtual machines from less busy servers can be migrated to other servers which are already running above a certain load level. Thus servers ressource can be aggregated and idle servers are powered down. The criteria for server consolidation is based on the observation that there is no significant difference regarding the energy consumption of a server running above 30% of its capacity and another server running at full load. An additional benefit of virtual machine migration is that the real network traffic of distributed applications can be reduced when they are running on the same server inside different virtual machines.

3. PROBLEMS RELATED TO MIGRATION

During the process of migration some new challenges have to be considered. If for example currently used services which are contained in virtual machines are migrated, the applications using these services must cope with such a suddenly changing network infrastructure. A possibility to solve this problems is to develop new protocols that are natively supporting such scenarios. Transparency towards migration is an important and crucial feature that has to be included in future protocols.

An alternative to cope with this problem is to virtualized the complete network toplogy that is actually used by the applications and services. Tunneling techniques can then be used to maintain the connection between an application and its services thus making the migration process transparent. The monitoring system has to cope with the same problem. As the process of migration is strongly dependent on the monitoring data it must be guaranteed that the monitoring system stays available during and after the process of migration.

Furthermore it must be possible to decide when migration can lead to enery saving. This requires that the relationship between measurable parameters (like CPU load, I/O rate, network traffic, ...) and overall energy consumption is known. This knowledge has to be integrated in a model which estimates the trends for future ressource consumption. Instead of migrating virtual machines services can be migrated thus ensuring the transparency of migration.

4. APPROACH FOR OPTIMIZATION OF ENERGY CONSUMPTION

In order to develop appropriated energy consumption models empirical data, reflecting the correlation between resource use and energy consumption, have to be collected. These informations can be gathered by running dedicated hardware benchmarks. To fully load certain hardware components (e.g. RAM, hard drive, ...) special applications for benchmarking will be developped. Important and decisive parameters that are strongly related to the energy consumption are identified during these benchmark runs. Beside the dynamic parameters also static characteristics have to be included inside the consumption model description. For example the boot time of a server is a static parameter that is important to be considered before shutting down a whole server. Thus the consumption model must also be able to deliver good prediction regarding the future server capacity. Furthermore the reliability and performance of the server's hardware components can influence the consumption model. Stochastical methods like the event tree analysis can help to model the reliability of hardware resources. These results can then be included in the energy consumption model. Finally the services are categorized in energy classes. The energy consumption model helps to make the optimal choice for selecting an energy efficient service.

All decisions in this project are based on the minimum principle. As an opposite to the economical principle, Ener-G is looking forward to minimize the input (energy and resources) for a fixed output. In other word a defined goal, this could be a QoS requirement, is achieved by using a strict minimum of hardware and energy resources.