Profiling and Analysis of Power Consumption for Virtualized Systems and Applications

Ziming Zhang and Song Fu
Department of Computer Science and Engineering
University of North Texas
ZimingZhang@my.unt.edu; Song.Fu@unt.edu

I. INTRODUCTION

Power consumption and energy efficiency have been a great concern in large-scale computing systems such as data centers and compute clouds. Moreover, the quantity of power consumption and resources utilized for cooling in large-scale computer systems will be beyond the boundary we can control [1]. There are at least two levels at which power/energy saving are being made. The first is at the computer architecture level. Techniques have been proposed to reduce processor energy usage through multi-core processors, designing energy proportional hardware, and increasing the number of performance and sleep states. A second level of energy optimizations is at the system layer, either at a single server or across multiple servers, through energy efficient use of the underlying hardware.

In addition, advancements in virtualization technologies enable us to get more computing power from the underutilized capacity of physical servers. Virtualization has become a key enabling technology for cloud computing. Virtual machines provide an additional layer of abstraction for clouds. Performance isolation and fast virtual machine instantiation make resource management easier and more efficient. The service availability can be enhanced significantly by means of virtual machine checkpointing and migration mechanisms. In order to efficiently utilize resources on different physical servers, multiple virtual machine instances are created for a database application. They run cooperatively and exploit cloud resource to meet the customer’s requirements. Power management in virtualized environments becomes an interesting and popular research topic [2].

II. METHODOLOGY

To improve the energy efficiency of computational applications in compute cloud systems, it is critical to profile the power consumption of real systems and applications at appropriate granularity. A large number of factors influence power and energy consumption in cloud computing systems, including each hardware component's electrical specification, the database usage characteristics, and the system software. In addition, power consumption and application performance are tightly coupled and often conflicting and complex. Improving energy efficiency without negatively affecting the database performance is challenging. Thus, we need direct power measurement to understand how each system component contributes to the total system power consumption. Coarse-grain database and server-level profiling can be used to identify how, when, and where power is consumed by cloud systems and running applications.

In coarse-grain profiling, we treat a cloud server as a unit and measure the impact of different configuration and scheduling strategies on the power consumption of the server. This approach can significantly reduce the complexity of managing large-scale data center systems at the back-end of clouds. In addition, it provides valuable information about the extent to which different factors on the power dynamics. To address the need for coarse-grained energy profiling for cloud computing, we propose a power-performance profiling framework and use it to evaluate energy efficiency and power-aware techniques for computational applications in clouds. The framework is a combination of hardware (including sensors and digital meters) and software (including VMs, instrumentation APIs, benchmarks, and analysis tools) that achieve automatic power and energy profiling at server granularity. It provides correlations between system activities and power/energy consumption of a server or multiple servers. The proposed framework uses direct or derived measurements to isolate components within server power profiles. Specifically, we isolate CPU, memory, disks, VMs, and applications. We change the configuration or settings of each preceding component or a combination of settings of multiple components, and measure the corresponding power/energy consumption.

In our experiments, we run a series of benchmark applications on a prototype cloud system. We collect
power consumption data by changing the configurations of some hardware components (i.e., CPU, memory and disks), which influence system power consumption more significantly [3]. Moreover, all the three hardware components have several different running modes which can be changed in order to alter system power consumption. In addition, multiple parallel applications run in user domains (virtual machines) and we change the number of virtual machines on each server to measure the impact of workload and it distribution on the system power/energy consumption.

III. EXPERIMENTAL SETTINGS AND RESULTS

We have implemented a prototype of our power/energy profiling framework and tested it in a prototype cloud system. Servers in the cloud have two types of configuration. Table 1 lists the configuration of servers used in the experiments. We run parallel applications on the cloud servers. In our experiments, we use NPB (NASA Parallel Benchmark) [4], which consists of a set of scientific computation applications. We create virtual machines on each cloud server and run NPB tasks on the virtual machines. We also change the settings of hardware components (e.g., CPU frequency, memory frequency, disk rotation speed) and the number of virtual machines running on each cloud server, and measure the power/energy consumption for different combinations of system configuration and settings.

Figures 1 and 2 show the results from our power consumption profiling.

IV. CONCLUSIONS

In this paper, we present a power/energy profiling framework for collecting and analyzing energy consumption data in compute clouds. We implement a prototype of our profiling framework and test it in a cloud computing environment. By measuring the power/energy consumption with various combinations of system configuration and settings, we build knowledge of the extent to which each factor influences the system power dynamics. As a future work, we plan to explore the collected profiling data and design resource management mechanisms that consider both the performance requirements of applications and energy budget of a cloud system for green computing.

ACKNOWLEDGMENT

We would like to thank the anonymous reviewers for their constructive comments and suggestions. This research was supported in part by U.S. NSF Grant CNS-0915396 and LANL Grant IAS-1103.

REFERENCES