Sustainable Computing: Informatics and Systems 4 (2014) 52–65

Contents lists available at ScienceDirect

Sustainable Computing: Informatics and Systems

journal homepage: www.elsevier.com/locate/suscom

Performance-controlled server consolidation for virtualized data centers with multi-tier applications☆

Yefu Wang[a], Xiaorui Wang[b,*]

[a] University of Tennessee, Knoxville, United States
[b] The Ohio State University, United States

ARTICLE INFO

Article history:
Received 8 August 2012
Received in revised form 24 August 2013
Accepted 19 February 2014

Keywords:
Power management
Data centers
Server consolidation
Multi-tier web applications
Virtualization
Control theory

ABSTRACT

Modern data centers must provide performance assurance for complex system software such as multi-tier web applications. In addition, the power consumption of data centers needs to be minimized to reduce operating costs and avoid system overheating. Various power-efficient performance management strategies have been proposed based on dynamic voltage and frequency scaling (DVFS). Virtualization technologies have also made it possible to consolidate multiple virtual machines (VMs) onto a smaller number of active physical servers for even greater power savings, but at the cost of a higher overhead. This paper proposes a performance-controlled power optimization solution for virtualized server clusters with multi-tier applications. While most existing work relies on either DVFS or server consolidation in a separate manner, our solution utilizes both strategies for maximized power savings by integrating feedback control with optimization strategies. At the application level, a novel multi-input–multi-output controller is designed to achieve the desired performance for applications spanning multiple VMs, on a short time scale, by reallocating the CPU resources and conducting DVFS. At the cluster level, a power optimizer is proposed to incrementally consolidate VMs onto the most power-efficient servers on a longer time scale. Empirical results on a hardware testbed demonstrate that our solution outperforms pMapper, a state-of-the-art server consolidation algorithm, by having greater power savings and smaller consolidation overheads while achieving the required application performance. Extensive simulation results, based on a trace file of 5415 real servers, demonstrate the efficacy of our solution in large-scale data centers.

1. Introduction

In recent years, power has become one of the most important concerns for enterprise data centers hosting thousands of high-density servers and providing outsourced business-critical IT services. A well-known approach to reducing server power consumption is to transition the hardware components from high-power states to low-power states whenever performance allows. For example, a widely used power-efficient server design is to have run-time measurement and control of the desired application performance by adjusting the CPU power states using dynamic voltage and frequency scaling (DVFS). However, while this approach can effectively reduce the dynamic power of the system with a small overhead (*e.g.*, under 100 μs in AMD Athlon [2] and under 50 μs in IBM POWER7 [3]), it cannot minimize the system leakage power for maximized power savings.

Recently, many data centers have begun to adopt server virtualization strategies for resource sharing. Virtualization technologies such as VMware and Xen can consolidate applications previously running on multiple physical servers onto a smaller number of physical servers, effectively reducing the power consumption of a server cluster by shutting down unused servers. More importantly, live migration [4] allows the movement of a virtual machine (VM) from one physical host to another with a reasonably short downtime [5]. This function makes it possible to use server consolidation as an online management approach, *i.e.*, having run-time estimation of resource requirements of every VM and dynamically re-mapping VMs to physical servers using live migration. When a more power-efficient VM-server mapping is found, unused servers can be put into the sleep mode for reduced power consumption. A recent study [6] shows that server consolidation can achieve a power saving of 34%, which is significantly more than that achieved by DVFS (17%). However, the down side of consolidation is that it

☆ This is a significantly extended version of a workshop paper [1].

* Corresponding author. Tel.: +1 614 247 1977.
E-mail addresses: xwang@ece.osu.edu, wang.3596@osu.edu (X. Wang).

http://dx.doi.org/10.1016/j.suscom.2014.02.001
2210-5379/

migrates VMs from one host to another, which may incur a high overhead in terms of both time and server resources. Therefore, server consolidation can only be conducted on a long time scale to amortize the migration overheads.

While power consumption must be minimized, an important goal of data center operators is to meet the service-level agreements (SLAs) required by customers, such as response time and throughput. SLAs are important to operators of data centers because they are the key performance metrics for customer service and are part of customer commitments. Therefore, it is important to guarantee the SLAs of the applications while minimizing the power consumption of a data center.

Guaranteeing the desired SLAs with minimized power consumption introduces several major challenges. First, complex system software, such as web applications, commonly has multi-tier installations. Thus, an application may span multiple VMs. This characteristic calls for advanced multi-input–multi-output (MIMO) control solutions to manipulate multiple VMs simultaneously. Second, web applications often face significant, unpredictable workload variations. To guarantee SLAs, a control solution must respond to a workload variation quickly by adjusting system resource allocation (*e.g.*, CPU time on each server). This cannot be achieved by migrating some VMs among the servers because a VM migration typically requires seconds, or even minutes, to finish. Finally, servers in a cluster may be manufactured by different hardware vendors and have different power efficiencies, *i.e.*, some servers are more power-efficient than others. A power management solution must be able to utilize this kind of heterogeneity to further reduce the total power consumption.

While existing work on data center power management relies on either DVFS or server consolidation in a *separate* manner, it is important to integrate application-level performance control and cluster-level server consolidation for maximized power savings for two reasons. First, most existing work of server consolidation assumes that the resource requirements of the VMs are known a priori or can be estimated through measuring the resource utilization. However, resource utilization commonly differs from resource demands, especially when servers are overloaded. Thus, it is preferable to have an application-level performance controller that can dynamically allocate system resource and conduct DVFS in response to short-term application resource demand variations. Second, a performance controller based on CPU resource allocation and DVFS may fail to guarantee the application performance when servers are overloaded due to significant long-term workload increases. Therefore, it is preferable to have a cluster-level consolidation algorithm to dynamically migrate VMs among physical servers to resolve the overload problem.

The potential large size of virtualized datacenters requires that the integration of SLA guarantee, DVFS and server consolidation should be implemented in a scalable manner. In this paper, we propose a scalable integrated management solution to minimize power consumption for virtualized data centers while providing application-level performance assurance. Each application-level performance controller adopts a MIMO control strategy to maintain the desired performance and reduce power consumption through DVFS and dynamic CPU resource reallocation. The cluster-level power optimizer then consolidates the VMs onto the most power-efficient servers and places unused servers into the sleep mode for power savings on a much longer time scale, to amortize the migration overhead. Specifically, the contributions of this paper are four-fold:

- We formulate the performance assurance problem for multi-tier applications running in multiple VMs as a MIMO control problem in a distributed manner at the application level. We then design a MIMO performance controller based on advanced Model Predictive Control (MPC) theory.
- We design a novel cluster-level power optimizer that conducts server consolidation based on the application-level resource demands profiled by the MIMO performance controllers. Compared with existing solutions, our solution can achieve more power savings with smaller consolidation overheads. We present detailed analysis of the computational overhead of our consolidation algorithm and demonstrate that the overhead is sufficiently small for deployment in real data centers.
- We integrate server consolidation with DVFS and CPU resource allocation to handle both long-term and short-term workload variations. The key benefit of such an integration is to achieve maximized power savings with application-level performance assurance. We introduce the system architecture of our integrated power management solution and the implementation details of each component.
- We present both empirical results on a hardware testbed and simulation results (based on a trace file of 5415 real servers) to demonstrate that our solution can effectively reduce cluster power consumption while achieving the desired performance for multi-tier applications in virtualized data centers. In addition, we show that our solution outperforms a state-of-the-art server consolidation solution, pMapper [7].

The rest of the paper is organized as follows. Section 2 introduces the overall architecture of our power management solution. Section 3 presents the modeling, design, and analysis of the performance controller. Section 4 discusses our power optimizer. Section 5 describes the implementation details of our testbed and simulator. Section 6 presents the results. Section 7 highlights the distinction of our work by discussing the related work with Section 8 concluding the paper.

2. System architecture

In this section, we provide a high-level description of our system architecture, which includes an application-level response time controller and a cluster-level power optimizer.

As shown in Fig. 1, our power management solution includes two levels. At the application level, for every application running in the cluster, there is a performance controller that dynamically controls the performance of the application by adjusting the CPU resource (*i.e.*, fraction of CPU time) allocated to the virtual machines running the application. We choose to control the 90-percentile response time of each multi-tier web application as an example SLA metric, but our management solution can be extended to control other SLAs such as average or maximum response times. We assume that the response time of a web application is independent from that of another web application. This is usually a reasonable assumption because they may belong to different customers. Hence, we choose to have a response time controller for each multi-tier application. A server-level CPU resource arbitrator then collects the CPU resource demands of all VMs hosted on the server, allocates the CPU resource to the VMs, and uses DVFS to save power, if the server has more CPU resources than the VMs require.

At the cluster level, a power optimizer collects the CPU resource demands of all the VMs running in the cluster, then uses a server consolidation algorithm to find the most power-efficient VM-server mapping while satisfying the CPU resource requirements of all the VMs. The power optimizer then sends VM migration commands to the VM migration interfaces on every server, if necessary. It also puts selected servers into the sleep/active mode.

Server consolidation makes it possible to put unused servers into the sleep mode, which can typically save more power than