Journal of Computer and System Sciences 81 (2015) 3–15

Contents lists available at ScienceDirect

Journal of Computer and System Sciences

www.elsevier.com/locate/jcss

A novel memory allocation scheme for memory energy reduction in virtualization environment

Xiaofei Liao, Hai Jin *, Shizhan Yu, Yu Zhang

Services Computing Technology and System Lab., Cluster and Grid Computing Lab., School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, 430074, China

ARTICLE INFO

Article history:
Received 31 January 2013
Received in revised form 11 June 2013
Accepted 22 January 2014
Available online 5 July 2014

Keywords:
Memory energy consumption
Virtualization
Memory allocation algorithm
Rank

ABSTRACT

Energy consumption is now a growing problem and leads to high costs in large data center based on virtualization technology. In data center, large memory consumes the main part of total energy, up to 41%. Observed from our experiments, we find that the dispersion characteristic of the virtual machine memory has a great impact on the overall energy consumption. This paper redesigns virtual machine's memory allocation scheme to reduce memory's energy consumption in virtual machine monitor. Two heuristic algorithms of allocating memory for virtual machines are proposed. Besides, based on the new memory allocations, we also design a dynamic memory management system for runtime environment. Experimental results show that our scheme achieves above 50% power reduction.

1. Introduction

The energy consumption problem in data center becomes more and more serious [1,2]. The memory system is now one of the most power-hungry parts in computing environment. Especially in high-end machines, the energy consumption of memory may transcend that of processor. For example, Lefurgy et al. [3] reported that in a high-end server equipped with 16 processors and 128GB main memory, the memory consumed 41% of total power while processors only used 28%. Therefore, reducing memory's energy consumption is very important in building energy-efficient data center.

Now, virtualization has become a powerful tool to simplify server deployment and automates resource management to optimize capacity and infrastructure management [4,5]. Especially for machines equipped with large main memory, a large number of virtual machines can be consolidated in a single computer, which greatly improves the resource utilization. Therefore, this paper aims to reduce memory's energy consumption in virtualization environment.

Modern memory chip consists of a number of memory blocks. Each memory block is called a rank in this paper. A rank is the smallest power control unit. During the runtime, a rank is allowed to be in different mode. In addition to normal mode, the rank can provide power-saving functions which cost little power but without data loss. These characteristics provide a physical basis for memory energy saving.

In virtualization environment, multiple virtual machines (VMs) are consolidated in one machine. A portion of physical memory is allocated to each VM. At any moment, only one VM is running on one physical processor, therefore only a part of whole memory is accessed. The basic idea to reduce memory energy is to transmit those ranks which have not been

* Corresponding author.
E-mail address: hjin@hust.edu.cn (H. Jin).

http://dx.doi.org/10.1016/j.jcss.2014.06.018
0022-0000/

accessed to power-saving mode as soon as possible. Jang et al. [6] have implemented this idea by changing rank mode in the context of switching VMs in order to cover the latency introduced by the conversion of rank mode.

However, we find that the distribution of VM may be very dispersing. In the worst case, the physical memory of one VM may scatter across all the memory ranks, therefore all memory ranks must be in active mode when the VM is running. It is not easy to achieve more memory's energy saving based on previous method in that case. Therefore, VM's physical memory distribution is very important in memory energy saving. However, current memory system in virtual machine manager (VMM) treats the whole physical memory equally and does not know the existence of ranks. Memory of a VM is managed by buddy system, which does not consider the energy consumption any more. In order to save memory energy, it is necessary to reduce the dispersion of VM's memory by redesigning memory allocation system in VMM.

In this paper, we propose a novel idea to reduce the energy consumption by redesigning virtual machine's memory allocation scheme. The main part of this idea contains two heuristic methods of allocating memory for virtual machines and a dynamic memory management system to save memory energy in dynamic environment. Experiments shows that the novel scheme has a good performance on energy saving.

The remainder of this paper is mainly organized as follows. Section 2 introduces the background and motivation. Section 3 describes the details of memory allocation system in static environment. Section 4 describes the dynamic memory management system in dynamic environment. Experimental results are then presented in Sections 5. Section 6 introduces the related works about memory energy savings. Conclusion and future works are given in Section 7.

2. Background and motivation

2.1. Synchronous dynamic random-access memory (SDRAM)

SDRAM is widely used for main memory, including single-data-rate (SDR), double-data-rate (DDR) and the ram-bus (RDRAM) architectures. We focus on the DDR3 in this paper since it is the main-stream DRAM architecture in today's computer systems. The smallest memory unit of DDR3 is called rank, which contains a number of DRAM chips. A DRAM chip is composed of a number of subcomponents. Therefore a rank can be in power-saving state if some subcomponents are disabled.

A memory rank has mainly four states: I/O state, Precharge state (PRE), Power down state (PD) and Self Refresh state (SR). I/O state is the state when the memory is reading and writing while PRE state is the state when memory is waiting for the next I/O operation. PD and SR are the so called power-saving states when several subcomponents are disabled and power consumption is greatly reduced. For convenience, we call I/O state and PRE state as active state while PD state and SR state as power-saving state. Since power of a rank in active mode is much higher than in power-saving mode, the key of reducing memory consumption is to schedule as many as possible memory ranks into power-saving mode. However, power transition incurs a large delay and increases power consumption, resulting in performance loss during the runtime. Therefore, power transition is processed during the switching context to cover the latency [7].

2.2. Virtualization

Our works are based on the Xen virtual machine monitor (VMM) [8]. Xen is a popular open-source hypervisor. The main function of Xen provides several isolated and secure environments for multiple virtual machines (VMs). Split driver model [9] in Xen allows the drivers of certain devices to be used in special guest OS (driver domain or Domain 0), which is responsible for serving other guest Oses (called guest domain or Domain U). The inter-communication between Domain 0 and Domain U is based on shared memory pages.

Current Xen uses a credit scheduler [10,11] as default strategy to schedule VMs. At any moment, only a part of VMs is running together. Therefore, only a part of the whole memory units can be accessed by the running virtual machines while the rest memory units are not needed. It can reduce the whole energy consumption of memory system to put the rest memory units into power-saving mode. This idea has been implemented by Jae-Wan Jang [6]. They first acquire the set of virtual machines which run from the virtual machine scheduler in VMM, and then transmit the ranks only related to those virtual machines to active state before they start to run while keeping other ranks in power-saving mode. In most cases, rank mode conversion is finished before those VMs start to run. It aims to reduce the impact on system performance introduced by memory access delay.

2.3. Motivation

However, we observe that the effect of memory' energy-saving is not so satisfying in our experiments based on the work of Jae-Wan Jang [6]. We track the operations of memory allocation and reclaim in VMM and acquire memory distributions information. Table 1 shows the memory distributions of virtual machines. The experiment is conducted on an experiment machine, which has dual Quad-Core Intel Xeon(R) 1.6GHz processors, 4GB DDR3 RAM containing 8 memory ranks. Three VMs are running on it, each VM is running 64-bit CentOS 5.5 distribution with 1GB memory and the Xen hypervisor is Xen 4.0.0.