Multi-layer Power Saving System Model Including Virtualization Server and Many-core Server

Joon-young Jung, Dong-oh Kang, Heong-jik Lee, Jang-ho Choi, Chang-seok Bae
Human Computing Research Section
Electronics and Telecommunications Research Institute
Deajeon, Korea
e-mails: {jyjung21, dongoh, leehj, janghochoi, csbae}@etri.re.kr

***Abstract*— Electric power has become an important part of our life and the power consumption is increasing rapidly. In data center, the power consumption is also increasing and it is essential to reduce the power consumption. Therefore, this paper proposes the requirement functions, the platform, and the structure of a Multi-layer Power Saving System (MPSS), including virtualization servers and many-core servers. Nowadays, the servers in data centers are sufficiently powerful to support many virtual machines. Therefore, the proposed system should be able to manage virtualization servers. We implement this system to monitor and control the power consumption using Intelligent Platform Management Interface (IPMI), the Advanced Configuration and Power Interface (ACPI), and the Dynamic Voltage and Frequency Scaling (DVFS). Through experimental results, the power consumption is reduced about 23%, when the Time-based Dynamic Power Management (TDPM) policy is used.**

Keywords-Data Center; Green IT; Management Server; Power Saving; Virtualization Server.

I. INTRODUCTION

There have been many studies to improve energy efficiency. Alonso et al. [1] presented a holistic approach to have a more efficient consumption behavior without lowering the threshold of comfort that consumers are used to. Power consumption is also increasing in data centers. Therefore, it is becoming a key design issue in data centers [2]. Even if the peak performance is required during short intervals, it should be supported for customer satisfaction using a lot of servers in the data center. Even if system components are not always required to be in the active state, it should be ready for customer satisfaction. So, an energy-efficient server has the ability to enable and disable these components [3]. The Dynamic Power Management (DPM) reconfigures clusters to provide the requested services with minimum active components and loads dynamically [4][5]. The Voltage and Frequency scaling can reduce the power consumption because the consumption energy of a processor is proportional to V^2 and F, where V and F are the operating voltage and frequency, respectively. The Dynamic Voltage Scaling (DVS) controls the operating voltages and frequencies of the processors according to their workload intensities. The DVF works well in web servers typically because they have a very large capacity compared to the average workload [6]. When a server is idle, it uses about 60% less power than when it is in use. And servers commonly work at around a 5% to 20% utilization rate because of the response time. An average of 10% of the servers in data centers were unused [7]. Vondra et al. [8] presented Time Series Forecasting, which enabled to predict the load of a system based on past observations. The prediction can be used to decide how many servers to turn off at night.

Modern servers in data centers are powerful enough to use virtualization to present many smaller virtual machines, where each machine runs a separate operating system [9]. In a virtualization system, several users connect to a single server and use applications independently [10]. The technology for a server virtualization exists. A hypervisor is the hardware virtualization technique that allows multiple operating systems to run concurrently on a host computer. Thus, multiple users can work together with a Virtual Machine (VM) and Input/Output (I/O) devices simultaneously. Nowadays, the virtualization server is generally used for efficiency in the data center. However, energy saving policies have not considered the virtualization server [14][15]. Therefore, this paper proposes 3-tier Power Saving System (3-PSS) to reduce the power consumption in the data center which includes virtualization servers and many-core servers. The 3-PSS can monitor and control the power consumption of guest OSs in virtualization servers.

The rest of the paper is organized as follows. In Section II, we propose MPSS modeling. In Section III, we provide the implementation. Section IV provides the experimental result, and some concluding remarks are finally given in Section V.

II. MPSS MODELING

A. MPSS Functions

There are several mandatory functions for the MPSS, as shown in Figure 1. The monitoring manager should monitor the power consumption of the servers in the data center. The power consumption data are also saved to analyze the power consumption pattern. The maximum value is used to decide a power capping value. The power consumption of the VM, single server, rack, cluster, and data center levels may be monitored. The emergency condition should be handled by using the power threshold alerting function. Whenever the

ISBN: 978-1-61208-365-0

target power budget cannot be maintained, an alert message is sent to the user.

The policy manager analyzes power consumption data. It is possible to forecast the power consumption using the historical data. The policy manager can dynamically make and change the policy to reduce the power consumption. The policy may be classified by workload, such as CPU-intensive, I/O-intensive, and memory-intensive workload. The policies may be used to save the power consumption in the data center [16].

The control manager should control the energy resources. It controls the power consumption of VM, node, rack, cluster, and data center using Intelligent Platform Management Interface (IPMI), the Dynamic Voltage and Frequency Scaling (DVFS), and the Advanced Configuration and Power Interface (ACPI).

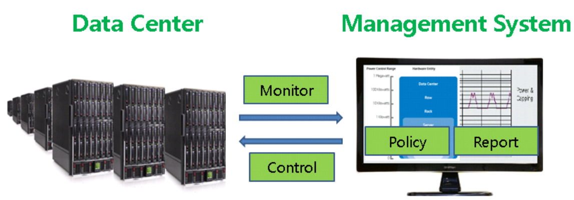


Figure 1. MPSS mandatory functions.

The reporting manager shows the power consumption at the VM, node, rack, cluster, and data center levels. This can be used for detailed analysis of the power consumption behavior of a particular workload or at a particular time interval. The report should be shown periodically and can show the power consumption data as a combination of location and service.

B. *MPSS Platform*

As shown in Figure 2, MPSS platform consists of the server platform, the power management interface, and the DPM platform. The server platform consists of the hardware, the host Operating System (OS), the hypervisor, and the guest OS.

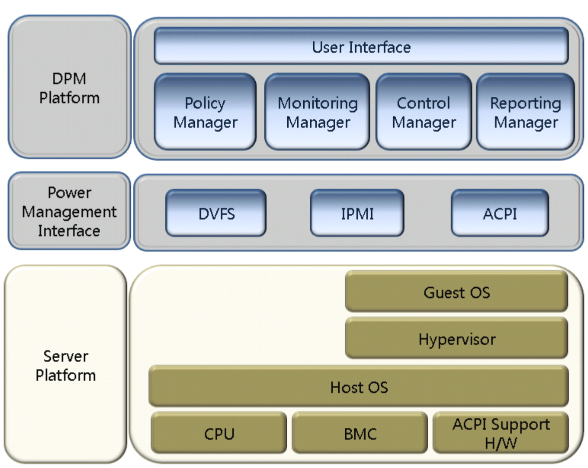


Figure 2. MPSS platform.

The Baseboard Management Controller (BMC) provides the intelligence behind intelligent platform management. The BMC manages the interface between the system management software and the platform management hardware. A hypervisor or Virtual Machine Monitor (VMM) is a piece of computer software, firmware or hardware that creates and runs virtual machines. The proposed platform in Figure 2 is type 2. Type 2 hypervisors are easy to use because they typically require no modification of the guest OS. The power management interface connects with the server platform to monitor and control the power consumption by using IPMI, ACPI, and DVFS. The IPMI is a standardized computer system interface used by system administrators to manage a computer system and monitor its operation [11]. The ACPI establishes industry-standard interfaces by enabling OS-directed configuration, power management, and thermal management of mobile, desktop, and server platforms [12]. DVFS may refer to dynamic voltage scaling and dynamic frequency scaling [13]. Dynamic voltage scaling changes the voltage value used in a component depending upon the circumstances. Dynamic frequency scaling changes the frequency of a microprocessor either to conserve power or to reduce the amount of heat generated by the chip. Voltage and frequency scaling are often used together to save power. The DPM platform consists of the policy manager, monitoring manager, control manager, and the reporting manager.

C. *MPSS Structure*

MPSS structure is shown in Figure 3. One power management server cannot monitor and control all nodes in the data center because there are lots of nodes in the data center. Therefore, MPSS has a 3-layer hierarchy structure, that is, it has a DPM master server, a DPM group server, and nodes.

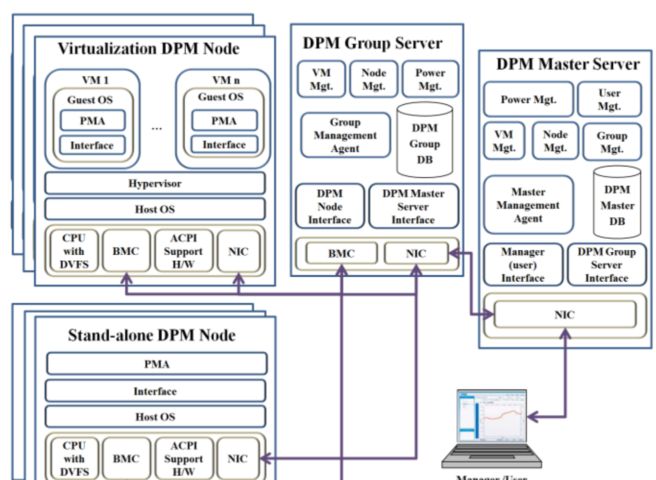


Figure 3. MPSS structure.

A DPM group server can monitor and control hundreds of nodes, which consists of virtualization DPM nodes and stand-alone DPM nodes. The DPM master server can monitor and control nodes through the DPM group server.

The DPM master server manages the power consumption for VMs, nodes, racks, and clusters in the data center using IPMI, ACPI, and DVFS.

A virtualization DPM node is detailed in Figure 4. It contains hardware and software components, including a BMC board, the ACPI support hardware, a CPU with DVFS, a Network Interface Card (NIC), a hypervisor and several VMs. Several VMs can be run in a DPM node. Each VM has a Power Management Agent (PMA), that includes DVFS, IPMI and ACPI modules, to save the power consumption and control resources. Each VM also has a hardware interface layer that connects the PMA and hardware using the hypervisor. The VM uses a group server interface to connect with a DPM group server. Each VM sends power consumption information and receives control messages through this interface. Each DPM node has a BMC board and uses OpenIPMI library to collect power consumption information. The DPM group master manages and periodically requests DPM nodes information. The DPM group server sends this information whenever the DPM master server requests it. The DPM master server monitors and collects DPM nodes information in the Data Base (DB). The DPM master server platform with Policy Based Management (PBM) and DPM analyzes the monitored data and controls nodes in the data center via IPMI, ACPI, and DVFS.

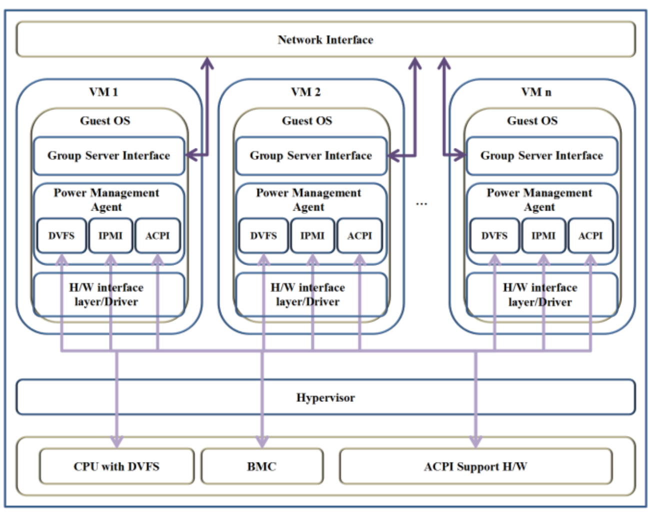


Figure 4. Block diagram of virtualization DPM node.

The PBM manages policies and rules that save the power consumption. It can make a policy by using historical data automatically. System manager can make a policy directly via a user-friendly user interface. DPM monitors energy resources in real time and controls them dynamically in order to save power consumption. The user monitors and controls the data center using the web-based User Interface (UI) program. All commands except the system on/off command, which uses a BMC interface, make use of the TCP/IP packet.

III. IMPLEMENTATION

We implemented the power saving system in the data center. The DPM master server can monitor nodes in the data center, as shown in Figure 5.

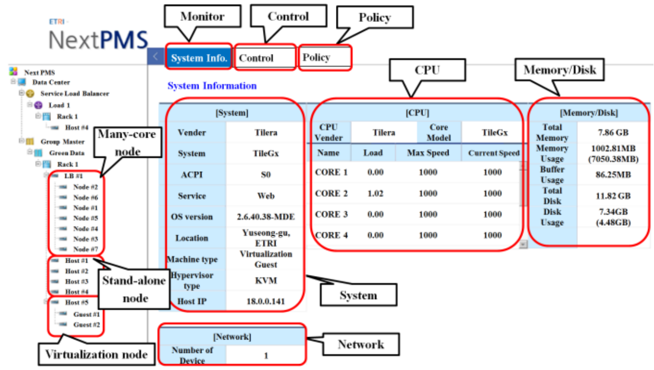


Figure 5. Node system information.

There are three kinds of nodes, that is, the many-core node, the stand-alone node, and the virtualization node. If one node is selected, the system information of that node, CPU, memory, network, and disk, is described in detail. The usage information of that node can be monitored, as shown in Figure 6.

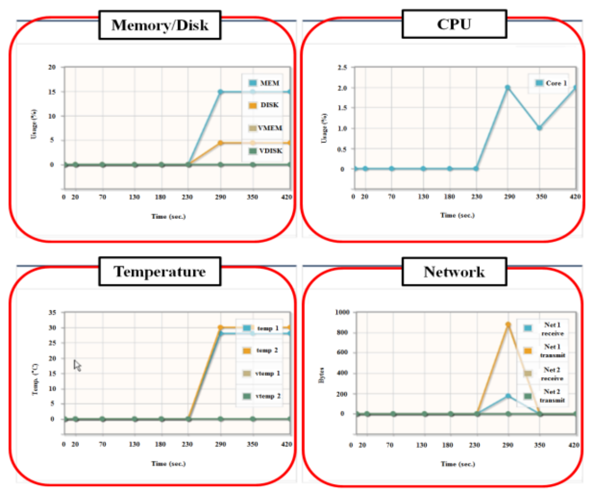


Figure 6. Node usage information graph.

The nodes in the data center can be controlled to save the power consumption, as shown in Figure 7.

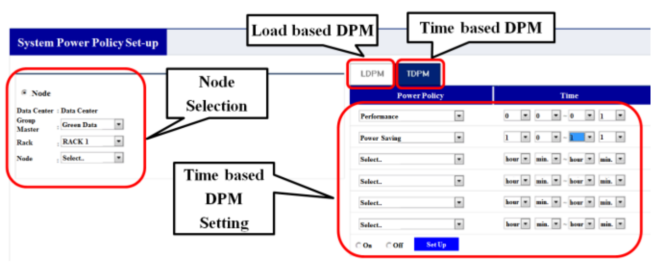
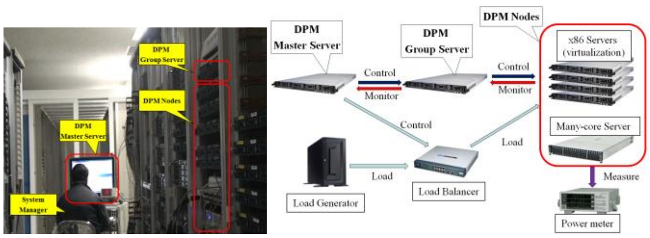


Figure 7. TDPM control interface.

The selected node can be controlled by the LDPM and the TDPM policies. The node can be turned off using the IPMI command, and turned on using the IPMI over LAN command.

IV. EXPERIMENTAL RESULTS

The 3-PSS is tested in the data center, as shown in Figure 8. The DPM nodes consist of stand-alone nodes, virtualization nodes, and many-core nodes. The DPM group server connects with DPM nodes using 1Gbps LAN.



(a) Test environment (b) Test configuration

Figure 8. 3-PSS test.

The DPM master server monitors and controls DPM nodes through the DPM group server. A load generator makes loads and sends those loads to a load balancer which distributes loads to DPM nodes. The load balancer is controlled by the DPM master server.

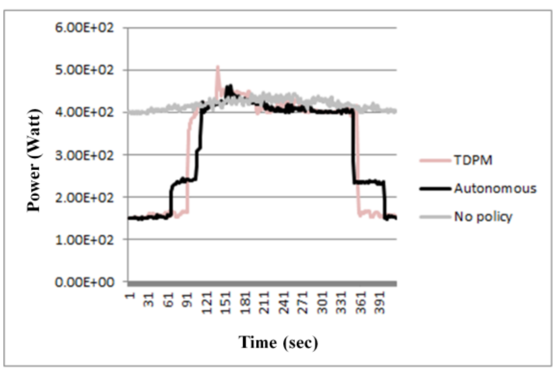


Figure 9. The result of time based DPM test.

Figure 10. The load scheme for time based DPM test.

We test TDPM with four x86 nodes. The result of the TDPM test is shown in Figure 9. The load increases for 4 minutes and then decreases for 3 minutes, as shown in Figure 10. We make these loads using the Apache JMeter. These loads scenario supposes that 7 users request Web services every 30 seconds for 210 seconds. We use Linux Stress tool to consume the power of Web servers when loads are requested from the Apache JMeter.

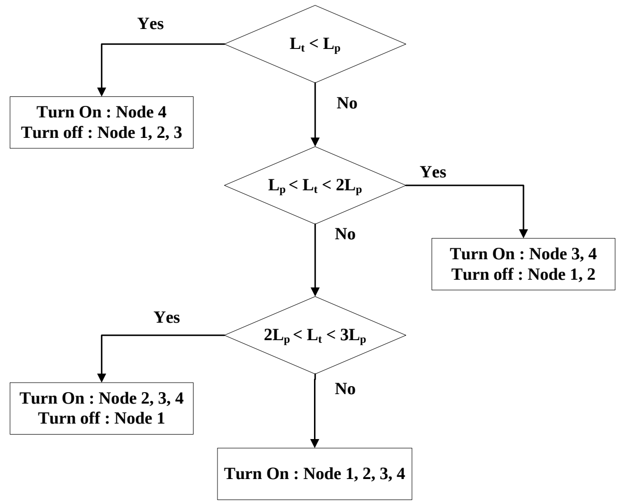


Figure 11. The autonomous policy based on load.

When the TDPM policy is used, one node is on while three nodes are off from 0 to 92 seconds, all four nodes are on from 92 to 348 seconds, then finally, one node is on and three nodes are off from 348 to 420 seconds. When the autonomous policy is used, the load profile (L_p), which is the loads that can be treated one node, is predefined by the DPM master server. The DPM master server monitors the total loads (L_t), which is the loads that be generated by the load generator, and controls the DPM nodes to save power consumption. The autonomous policy of this experiment is shown in Figure 11. If $L_t < L_p$, the DPM mater server turns on the node 4 and turns off the node 1, node 2, and the node 3. If $L_p < L_t < 2L_p$, it turns on the node 3 and the node 4 and turns off the node 1 and the node 2. If $2L_p < L_t < 3L_p$, it turns on the node 2, node 3, and the node 4 and turns off the node 1. If $3L_p < L_t$, it turns on all nodes. The results of the autonomous policy is that one node is on from 0 to 62 seconds and from 396 to 420 seconds, two nodes are on from 62 to 106 seconds and from 345 to 396 seconds, three nodes are on from 106 to 113 seconds, and four nodes are on from 113 to 345 seconds. In this case, the power used is over 400 watts, when the policy is not used. However, the reduction in power consumption is about 23%, when the TDPM policy or the autonomous policy is used.

V. CONCLUSION AND FUTURE WORK

For saving the power consumption in the data center, we considered the MPSS functions and propose the 3-PSS platform, and structure for the data center including virtualization servers and many-core servers. We implemented this system to monitor and control the power consumption using IPMI, ACPI, and DVFS. Through experimental results, the power consumption has been found to be reduced by about 23%, when the TDPM or the autonomous policy is used.

For further work, we plan to develop MPSS without performance degradation. The performance is an essential factor in the data center. Therefore, the performance degradation should not occur due to the power consumption saving. We will also estimate the number of nodes that can be managed by a single server.

ACKNOWLEDGMENT

This work was supported by the IT R&D program of MSIP/IITP, [10040696, Development of energy saving green server under 1mW/MIPS for cloud service].

REFERENCES

[1] I. G. Alonso, M. R. Fernández, J. J. Peralta, and A. C. García, “A Holistic Approach to Energy Efficiency Systems through Consumption Management and Big Data Analytics,” International Journal on Advances in Software, vol. 6, numbers 3 & 4, 2013, pp. 261–271.

[2] E. N. Elnozahy, M. Kistler, and R. Rajamony, "Energy-Efficient Server Clusters," Proc. 2nd Workshop Power-Aware Computing Systems, 2003, pp. 179–196.

[3] L. Benini, A. Boliolo, and G. D. Micheli, “A Survey of Design Techniques for System-Level Dynamic Power Management”, IEEE Transactions on VLSI Systems, vol. 8, no. 3, June 2000, pp. 299–316.

[4] J. Lorch and A. Smith, “Software strategies for portable computer energy management,” IEEE Personal Commun., vol. 5, June 1998, pp. 60–73.

[5] L. Benini and G. De Micheli, Dynamic Power Management: Design Techniques and CAD Tools. Norwell, MA: Kluwer, 1998.

[6] P. Bohrer, et al., “The Case for Power Management in Web Server,” Power-Aware Computing, January 2002, pp. 261–289.

[7] P. Foster, PC and Server Power Management Software, Pike Research, 1Q 2010.

[8] T. Vondra and J. Sedivy, “Maximizing Utilization in Private IaaS Clouds with Heterogenous Load through Time Series Forecasting,” The International Journal on Advances in Systems and Measurements, vol. 6, numbers 1&2, 2013, pp. 149–165.

[9] P. Barham, et al., “Xen and the art of virtualization,” Proc. ACM Symposium OSP, 2003, pp. 164–177.

[10] J. Nieh, S. J. Yang, and N. Novik, "A comparison of thin-client computing architectures," Technical Report CUCS-022-00, November 2000.

[11] Intel, Hewlett-Packerd, NEC, and Dell, “Intelligent Platform Management Interface Specification Second Generation,” October 2013.

[12] Intel, Hewlett-Packerd, Microsoft, Phoenix, and Toshiba, “Advanced Configuration and Power Interface Specification revision 5.0,” December 2011.

[13] L. Sueur, Etienne, and G. Heiser, "Dynamic voltage and frequency scaling: The laws of diminishing returns," Proc. of the 2010 international conference on Power aware computing and systems. USENIX Association, 2010, pp. 1-8.

[14] Elnozahy, Mootaz, M. Kistler, and R. Rajamony. "Energy conservation policies for web servers," Proc. of the 4th conference on USENIX Symposium on Internet Technologies and Systems, vol. 4, 2003, pp. 8-8.

[15] Kephart, et al., "Coordinating Multiple Autonomic Managers to Achieve Specified Power-Performance Tradeoffs," International Conference on Autonomic Computing, vol. 7, June 2007, pp. 24.

[16] C. Wissam, and C. Yu, "Survey on power management techniques for energy efficient computer systems," Laboratory Report, Mobile Computing Research Lab, 2002.