Towards Energy-Efficient Reactive Thermal Management in Instrumented Datacenters

Ivan Rodero, Eun Kyung Lee, Dario Pompili, Manish Parashar
Center for Autonomic Computing
Rutgers University, NJ, USA

Marc Gamell
Open University of Catalonia
Barcelona, Spain

Renato J. Figueiredo
Center for Autonomic Computing
University of Florida, FL, USA

Abstract—**Virtual Machine (VM) migration is one of the most common techniques used to alleviate thermal anomalies (i.e., hotspots) in cloud datacenter's servers of by reducing the load and, therefore, decreasing the server utilization. However, there are other techniques such as voltage scaling that also can be applied to reduce the temperature of the servers in datacenters. Because no single technique is the most efficient to meet temperature/performance optimization goals in all situations, we work towards an autonomic approach that performs energy-efficient thermal management while ensuring the Quality of Service (QoS) delivered to the users.**

In this paper, we explore ways to take actions to reduce energy consumption at the server side before performing costly migrations of VMs. Specifically, we focus on exploiting VM Monitor (VMM) configurations, such as pinning techniques in Xen platforms, which are complementary to other techniques at the physical server layer such as using low power modes. To support the arguments of our approach, we present the results obtained from an experimental evaluation on real hardware using High Performance Computing (HPC) workloads on different scenarios.

I. INTRODUCTION

The growing scale of consolidated virtualized datacenters has made issues related to power consumption, air conditioning, and cooling infrastructures critical concern in terms of the growing operating costs. Furthermore, power and cooling rates are increasing eight-fold every year [1] and are becoming a dominant part of IT budgets. Addressing this problem is an important and immediate task for enterprise datacenters. Current work in the field of thermal management explores efficient methods of extracting heat from the datacenter [2]. Researchers also proposed software-based approaches [3], [4] in order to manage temperature by efficiently scheduling workload in low-temperature regions. Virtual Machine (VM) migration is one of the most common techniques used to alleviate thermal anomalies (i.e., hotspots) in cloud datacenter's servers by reducing the load and, therefore, by decreasing the server utilization. Switching off idle servers is also a typical policy along with VM migration for power management. In some situations, however, the VM migration process may lead to unacceptable operational temperatures.

Along with VM migration there are other mechanisms that can be used to reduce the temperature of the servers in datacenters. However, not always the same mechanism is the most efficient to meet the desired optimization goals. The long-term goal of our approach is to autonomically manage datacenters using the information from sensors and taking decisions at different levels (through controllers) based on the optimization goals (e.g., performance, energy efficiency, cost). To do this, we consider an architecture composed by layers belonging to different abstract components with different responsibilities but with the same common objectives. Specifically, the architecture is composed by four layers: environment layer (which detects, localizes, characterizes, and tracks thermal hotspots using scalar sensors such as temperature and humidity), physical resource layer (which manages the hardware and software components of servers), virtualization layer (which instantiates, configures, and manages VMs), and application layer (which is aware of the workload's and application's characteristics and behavior). For reactive thermal management we specifically focus on the possible interactions between the environment layer and the virtualization and physical resource layers based on temperature and power consumption.

In this paper, we explore ways to take actions at the server side before performing costly migrations of VMs, including techniques that can be used as temporary actions that may facilitate migrations without incurring in additional penalty in terms of server thermal behavior (e.g., using pinning before performing VM migrations). We aim at optimizing the energy efficiency of datacenters while ensuring the Quality of Service (QoS) delivered to the users. Specifically, we focus on exploiting VM Monitor (VMM) configurations, such as pinning techniques (i.e., CPU affinity) in Xen platforms. Note that this approach is complementary to other techniques at the physical server layer such as using low power modes. The techniques that we consider in this work try to ensure an acceptable thermal behavior of the datacenter's hardware components (i.e., the temperature is under a certain threshold).

Although our experiments focus on the impact of thermal management techniques on a single server, we believe that the obtained results are sufficiently representative to be extrapolated to consolidated virtualized datacenters where components at different layers need to coordinate to achieve the common objective of improving datacenter management efficiency. To support the arguments of our approach, we present the results obtained from an experimental evaluation on real hardware using High Performance Computing (HPC) workloads in different scenarios. The results state that pinning is an effective mechanism to react to thermal anomalies and show tradeoffs between VM migration and pinning depending on the system characteristics and optimization goals. From the knowledge distilled from this study we conclude that an autonomic system could leverage different thermal management techniques to optimize the datacenter's energy efficiency while ensuring the QoS delivered to the users. The main contributions of this papers are: (1) study different reactive thermal management techniques for virtualized and instrumented datacenters from the energy perspective towards the design of an autonomic approach, (2) study the tradeoffs between performance, energy efficiency, and thermal efficiency of the techniques for HPC workloads, and (3) propose pinning as a technique for energy-efficient thermal management.

The rest of the paper is organized as follows. Section II surveys related work. Section III discusses temporal-spatial characteristics of hotspots and describes some reactive thermal management approaches. Section IV presents and discusses the experimental results. Finally, Section V concludes the paper and outlines future research directions.

II. Related Work

Several approaches have been proposed for addressing cooling and thermal issues. Moore et al. [5] propose a method to infer a model of thermal behavior to automatically reconfigure the thermal load management systems, thereby improving cooling efficiency and power consumption. They also propose in [3] thermal management solutions focusing on scheduling workloads while considering temperature-aware workload placement. Bash et al. [6] propose a policy to place the workload in areas of a datacenter that are easier to cool, which results in cooling power savings. Tang et al. [4] formulate and solve a mathematical problem that maximizes the cooling efficiency of a datacenter. Bianchini et al. [7] propose emulation tools for investigating the thermal implications of power management. In [8], they present C-Oracle, a software prediction infrastructure that makes online predictions for datacenter thermal management based on load redistribution and Dynamic Votage and Frequency Scaling (DVFS). Raghavendra et al. [9] propose a framework that coordinates and unifies five individual power management solutions (consisting of HW/SW mechanisms).

A large body of work in datacenter energy management addresses the problem of the request distribution at the VM management level in such a way that the performance goals are met and the energy consumption is minimized. Zhao et al. [10] present an experimental study of virtual machine migration focused on the VM-based resource reservation problem looking at performance. Song et al. [11] propose an adaptive and dynamic scheme for adjusting resources (specifically, CPU and memory) between virtual machines on a single server to share the physical resources efficiently. Kumar et al. [12] present vManage, a practical coordination approach that loosely couples platform and virtualization management aiming at improving energy savings and QoS and at reducing VM migrations. Soror et al. [13] address the problem of optimizing the performance of database management systems by controlling the configurations of the virtual machines in which they run. Laszewski et al. [14] present a scheduling algorithm for VMs in a cluster to reduce power consumption using DVFS.

Several research efforts propose methods to jointly manage power and performance at the physical resource layer. One of the most used techniques in the last decades to save energy is DVFS. Researchers have developed different DVFS scheduling algorithms and mechanisms to save energy while provisioning resources under deadline restrictions. Chen et al. [15] address resource provisioning and propose power management strategies with SLA constraints based on steady-state queuing analysis and feedback control theory. They use server turn on/off and DVFS for enhancing power savings. Ranganathan et al. [16] highlight the current issue of under utilization and over-provisioning of the servers. They present a solution of peak power budget management across a server ensemble to avoid excessive over-provisioning considering DVFS and memory/disk scaling. Nathuji et al. [17] investigate the integration of power management and virtualization technologies. In particular they propose VirtualPower to support the isolated and independent operation of virtual machine and control the coordination among virtual machines to reduce the power consumption. Rusu et al. [18] propose a cluster-wide on/off policy based on dynamic reconfiguration and DVFS. They focus on power, execution time, and server capacity characterization to provide energy management. Kephart et al. [19][20] address the coordination of multiple autonomic managers for power/performance tradeoffs by using a utility function approach in a non-virtualized environment.

VMM configurations (e.g., within the Xen hypervisor) has been used in different approaches such as is [21]. However, to the best of our knowledge, none of the existing approaches has exploited VMM configurations (i.e., pinning techniques) to mitigate the effects of thermal anomalies with energy efficiency and performance as optimization goals.

III. Reactive Thermal Management in Datacenters

In this section, we briefly discuss the characteristics of thermal hotspots that majorly cause thermal inefficiency in datacenters. Then, we introduce three different techniques to mitigate the effects of hotspots that are evaluated in the following section.

A. Characteristics of Thermal Hotspots

Hotspots can be detected using internal and external temperature sensors when measurements cross specific temperature values defined as hotspot threshold. Hotspots are difficult to localize accurately in space and are hard to predict in time: this is because the heat transfer along the airflows (convection) and through the server blades and racks (conduction) are phenomena difficult to model. In addition, hotspots change their positions in time and space depending on several factors such as distribution and intensity of running workloads, server characteristics, airflow circulation, etc. Thus, it is crucial to understand the characteristics of hotspots before taking actions towards energy-efficient thermal management.

Fig. 2 shows the temporal correlation of the measured temperature data collected from 13 TelosB sensor nodes (deployed in front of the outlet fans of each server in a dual rack system), each placed on 13 vertically arranged servers as showed in Fig. 1. These TelosB sensor nodes are wirelessly connected and built with IEEE 802.15.4 compliant CC2420 radio (2.4 GHz) with a few sensors of temperature, humidity, and light. The temperature sensors on the TelosB nodes and internal sensors nearby the CPU are used to measure temperature in order to observe heat propagation in space and time. Figs. 2a and 2b correspond to the set of experiments of using TelosB nodes in which only servers 7-13 and 7-10 are running, respectively. These results show that some idle servers that are in spatial proximity to operational servers experience an increase in temperature much higher than that of other idle servers. These results show that an hotspots affecting a server may in fact be caused by another server running on the same rack or on a different one.

Fig. 3 shows the thermal behavior of a single server in the presence of an hotspot and the corresponding power consumption. It also shows the impact of VM migration on server's temperature and power consumption. The increase of the external temperature, which is controlled with a heat source (from second 600 to 1400) results on the increase of the server's internal temperature. The figures illustrate the impact of migrating one VM on the server's temperature and power consumption (at second 900 - reacting to the hotspot). In comparing the internal server temperature (Fig. 3a) and power consumption (Fig. 3b) we observe that there is a correlation between the two metrics.

B. Reactive Thermal Management Approaches

A typical mechanism used to to reduce the server's temperature (e.g., to react to a hotspot) consists in decreasing the heat generated by the server, which, based on our measurements is

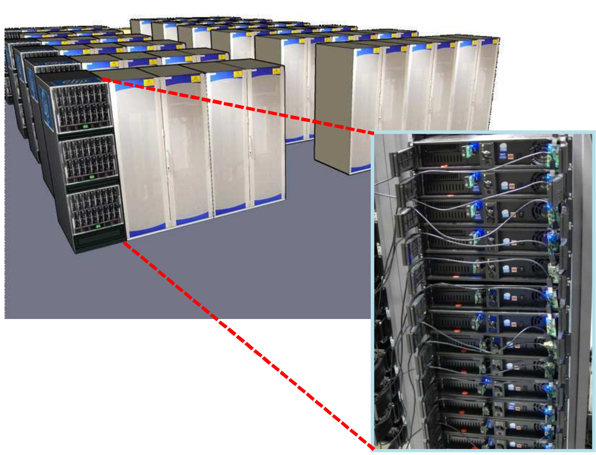


Fig. 1: 26 TelosB motes deployed on a rack to observe the characteristics of hotspots.

highly correlated with the server's power consumption. As the CPU[1] is the most power consuming component of a server, we consider CPU power (as a simplification) to describe the different techniques analyzed in this paper. Eq. (1) shows a simplified dynamic power dissipation model for CPU, where C is the capacitance of the processor (that we consider fixed), α is an activity factor (also known as switching activity), and V and f are the operational voltage and frequency, respectively.

$$P_{cpu} \sim C \times \alpha \times V^2 \times f. \qquad (1)$$

Therefore, the power consumption of a server can be decreased by either reducing the activity of CPUs or reducing the frequency/voltage of CPUs (via DVFS). In the following, we discuss different techniques that use this model to aim at reducing the server's power consumption (and thus the heat generated). We do not take into account cooling or placement issues; rather, we focus on the energy efficiency of thermal management approaches on the servers considering HPC workloads and virtualized environments under the assumption that cooling and placement do not change.

1) VM Migration: This technique consists in moving a running VM, its guest OS, and all of its applications to a different server. Migrating VMs reduces the CPU activity α in (1) and, if a CPU is freed and the OS implements dynamic CPU power management, it can also reduce the frequency/voltage. We assume that the OS power management is enabled by default. When we need to migrate a VM or multiple VMs to react to a hotspot, one of the following four scenarios is possible:

- Another server is available to host the VM(s) that are being migrated: migration can be performed at the penalty of some overhead (energy, latency, bandwidth).
- A server has been powered down: we can perform the migration after powering the server up at the additional penalty of booting.
- All servers already have some load but some of them have higher thermal efficiency: we can perform the migration but with possible penalty due to the the resource sharing between migrated workload and the existing workload in the destination server.
- Any other server is available to host new VM(s): in this case, we can only suspend the VM(s) until a server becomes available.

2) Processor DVFS: This is an effective technique to reduce processor power dissipation supported in most of current processors. DVFS reduces the processor frequency/voltage but not the activity factor. However, CPU-intensive workloads running at low frequencies may experience a significant penalty on their execution time. Within a single server we can use DVFS in two ways:

- *By reducing the frequency/voltage of all CPUs simultaneously*: the workload execution may increase depending of the frequency/voltage reduction and workload

[1] We use the term CPU to refer to each core of a multi-core processor

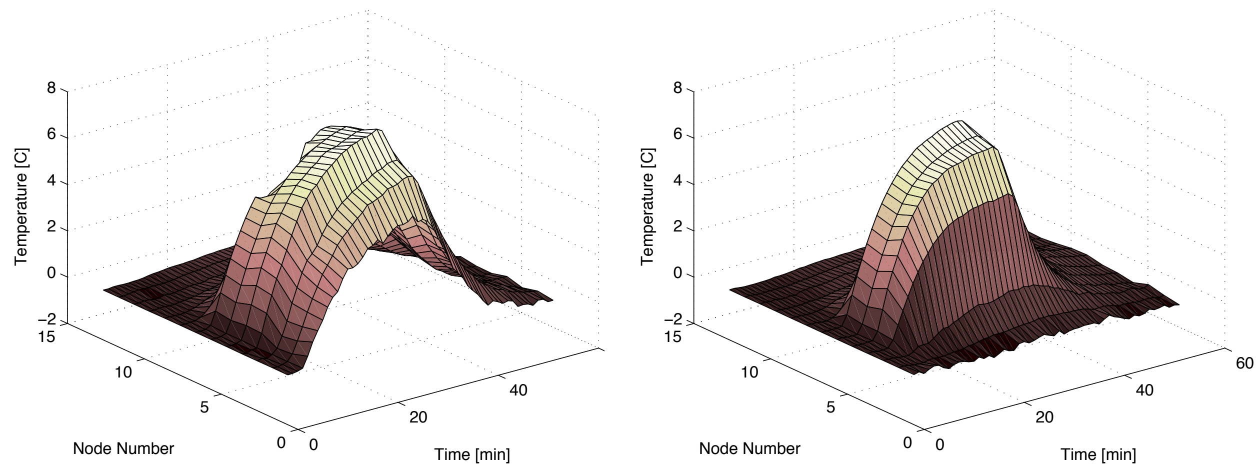
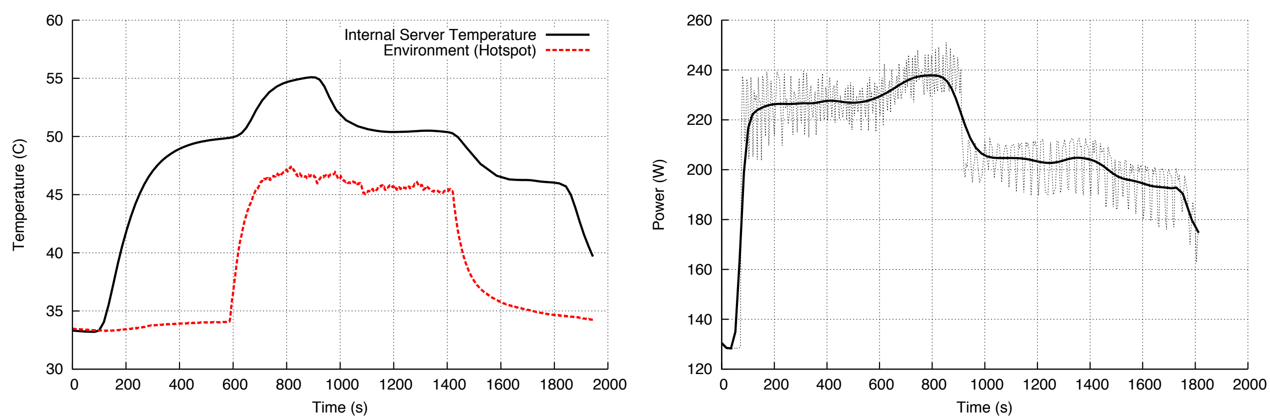


Fig. 2: Temporal and spatial measurement data (temperature) among 13 nodes



(a) Temperature

(b) Power Consumption

Fig. 3: Server's thermal and power behavior in the presence of an hotpost

characteristics (e.g., I/O-intensive workloads may not be significantly penalized).

- *By reducing the frequency/voltage of a subset of CPUs*: if different VMs are independent they may complete their workload at different times. If VMs are coupled (e.g., MPI applications) the workload running in slower CPUs may penalize the workload running in faster CPUs. However, some architectures have restrictions (e.g., by paired CPUs).

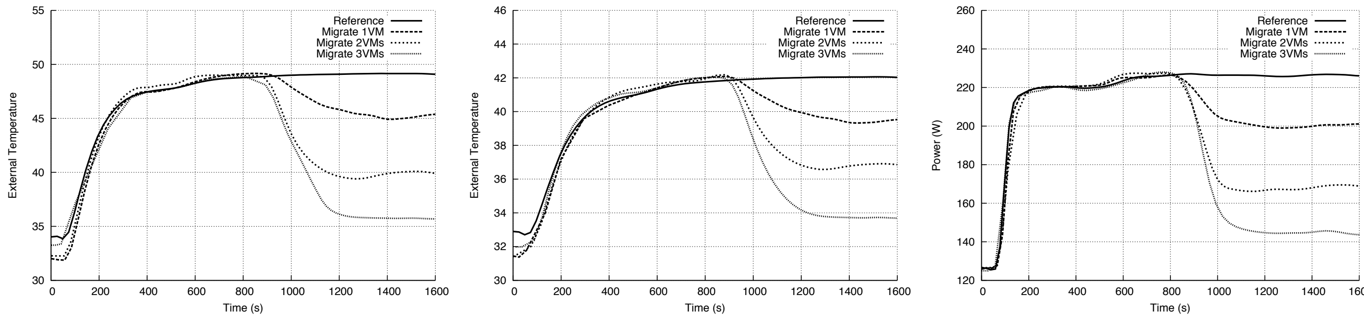
3) VMM Configuration (Pinning): We propose to use this technique to react to hotspots as alternative to VM migration and DVFS. VMMs may allow *virtual* CPUs (vCPUs) of VMs to be assigned to *physical* CPUs (pCPUs) in two different approaches: without and with affinity (pinning). In the former, the VMM determines how vCPUs are assigned to pCPUs; in the latter, the VMM allows hard assignments of the vCPUs to one or more pCPUs. Pinning techniques are typically used where the characteristics of the workload would benefit from executing on specific CPUs (e.g., cache locality).

We propose to reduce the activity of one or more CPUs by pinning the VMs to the other CPUs. As we assume that the OS performs by default dynamic CPU power management, when a CPU is freed from VMs activity, its frequency/voltage can also be reduced. However, the activity of the running CPUs may be increased, resulting in a penalty in the workload's execution performance caused by higher resource sharing.

IV. Experimental Evaluation

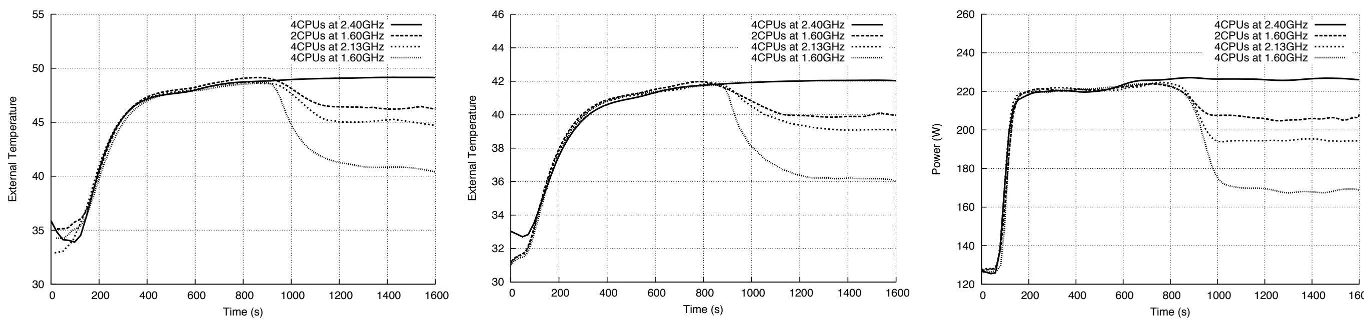
A. Experimental Environment

The experiments were conducted using two Dell servers, each with a Intel quad-core Xeon X3220 processors which operate at four frequencies ranging from $1.6\mathrm{G}Hz$ to $2.4\mathrm{G}Hz$, $4\mathrm{G}B$ of memory, two hard disks, and two $1\mathrm{G}b$ Ethernet interfaces. This is intended to represent a general-purpose rack server configuration, widely used in virtualized datacenters. The servers run CentOS Linux operating system with a patched 2.6.18 kernel with Xen version 3.1. To empirically measure the "instantaneous" power consumption of the servers



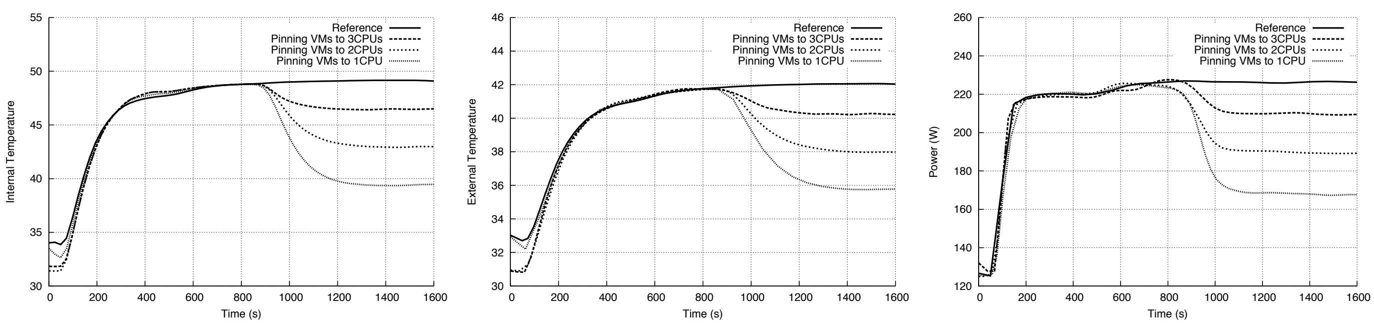
(a) Internal temperaure (b) External temperature (c) Power comsuption

Fig. 4: Thermal behavior and power consumption using VM migration



(a) Internal temperaure (b) External temperature (c) Power comsuption

Fig. 5: Thermal behavior and power consumption using processor DVFS



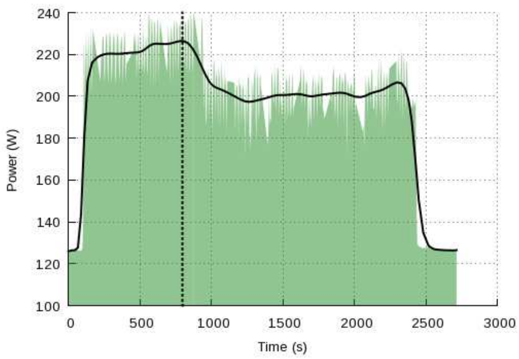
(a) Internal temperaure (b) External temperature (c) Power comsuption

Fig. 6: Thermal behavior and power consumption using pinning techniques (pinning 4 VMs to different number of CPUs - from 3 to 1 CPUs)

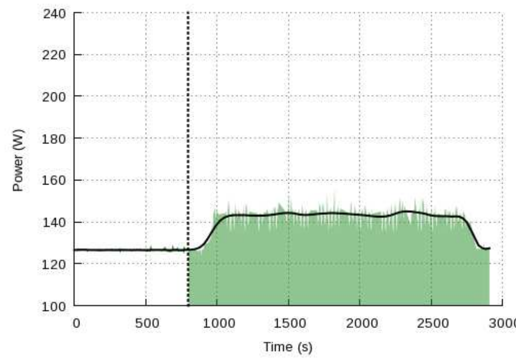
we used a *"Watts Up? .NET"* power meter. This power meter has an accuracy of $\pm 1.5\%$ of the measured power with sampling rate of $1\mathrm{H}z$. The meter was mounted between the wall power and the server. We estimate the consumed energy by integrating the actual power measures over time. We used TelosB motes to measure both internal and external temperatures of the server as described in Section III-A. Because the servers used for our experiments are not production machines, we used a Sunbeam SFH111 heater (directed to the servers) in order to emulate a thermal hotspot.

B. Results

We have evaluated the three VM management techniques discussed in Section III-B. Figs. 4, 5, and 6 show the thermal behavior and power consumption of a server when running a HPC workload with the different VM management techniques considered in this paper. Specifically, we used the HPL linpack benchmark, which uses intensively different resources such as CPU and memory. All the techniques and configurations have been evaluated under the same conditions. The experiment consists of running HPL in 4 VM instances with the same configuration and applying a given technique 800 seconds



(a) Source server (running 4 VMs)



(b) Destination server (initially idle)

Fig. 7: Power consumption of source/destination servers when migrating one VM. Migration starting time is depicted with the dashed vertical line.

after starting the experiment (which is long enough to reach the steady state). The figures focus on the initial part of the experiment to show better the trends and only plot the Bezier curves for readability. We obtained the external temperature of the server with a sensor placed in the back side of the server, and the internal temperature (in °C) of the server with a sensor placed inside the server. The internal sensor provides the average temperature of the server's components (not only CPU temperature, which can be obtained from its internal sensors).

We can appreciate in the figures that internal and external temperatures are strongly correlated and follow similar patterns. However, the internal temperature is almost 8°C higher than the external temperature. Temperature and power are also correlated but variations in temperature are slower than in power. Overall, the higher reduction of temperature and power are obtained using VM migration. The higher the number of VMs migrated, the more significant the decrease of temperature and power. The reduction of temperature and power is more moderate using DVFS than using VM migration. The highest reduction in temperature and power is obtained operating all 4 CPUs at $1.60\mathrm{G}Hz$. This is consistent with (1), where the higher operational frequency the higher temperature and power consumption. However, running 2 CPUs at $1.60\mathrm{G}Hz$ and 2 CPUs at $2.40\mathrm{G}Hz$ obtains slightly worse results than running all CPUs at $2.13\mathrm{G}Hz$.

The reduction of temperature and power using pinning is lower than using VM migration but higher than using DVFS. The reduction of temperature and power is higher when the VMs are pinned to fewer CPUs. Although using different techniques we obtain different reductions of temperature and power, the plots obtained with the three different mechanisms follow similar patterns. Hence, we can conclude that all of them can reduce the temperature and power consumption effectively. Due to space limitations, we do not provide further study of the thermal behavior of the considered techniques and we focus on the tradeoffs between performance, energy efficiency, and thermal efficiency (i.e., temperature reduction).

In order to measure the energy consumed by the servers in the experiments that perform VM migrations, we have considered the energy consumed by the original server during the whole execution of the experiment plus the energy consumed by the destination machine from when the VMs start being migrated to the destination machine (i.e., shadowed area in Figs. 7a and 7b).

In the case that the destination machine already hosts other VMs, we take into account the energy consumed from the increase of power in relation to the power consumed by the server before the migration. Although we can find other intermediate scenarios, in our experiments we only consider these two scenarios as they provide simplified but meaningful performance bounds.

Although VM migration seems better in terms of thermal efficiency, we look at the tradeoffs between thermal efficiency and other dimensions such as performance and energy efficiency. Figure 8 illustrates the tradeoff between thermal efficiency (reduction of temperature) and performance (workload's response time) when pinning vCPUs of VMs to different number of physical pCPUs.

Table I summarizes the obtained results. It shows i) the obtained makespan (the time needed to complete the workload, which may include the migration overhead), ii) the energy consumed, iii) the Energy Delay Product (EDP), which is a good metric for energy efficiency because it captures the effect of energy management on performance [22], and iv) the reduction of temperature ("Temp. ↓" in Table I). For the reference execution (regular execution without any specific technique)

TABLE I: Experimental results using different VM management techniques. "Destination Empty" means that the destination server is running but idle, "Destination Full" means that the destination server has 4 VMs running on the 4 CPUs, and "Destination Off" means that de destination server is switched off.

Technique - Configuration	Makespan(s)	Energy(J)	EDP	Temp ↓
Reference (regular execution)	2,239 s	501,023 J	1,121,790,497	-
Migrate 1 VM - Destination Empty	+19.60%	+52.82%	+82.79%	4°C
Migrate 1 VM - Destination Full	+37.51%	+31.92%	+81.41%	4°C
Migrate 1 VM - Destination Off	+26.52%	+57.40%	+83.10%	4°C
Migrate 2 VMs - Destination Empty	**+14.15%**	**+51.86%**	**+73.36%**	**9°C**
Migrate 2 VMs - Destination Full	+37.51%	+30.92%	+80.04%	9°C
Migrate 2 VMs - Destination Off	+21.08%	+56.44%	+73.68%	9°C
Migrate 3 VMs - Destination Empty	+16.43%	+51.71%	+76.65%	15°C
Migrate 3 VMs - Destination Full	+37.60%	+25.89%	+73.12%	15°C
Migrate 3 VMs - Destination Off	+23.35%	+56.29%	+76.96%	15°C
DVFS 4 CPUs @1.60GHz	+78.38%	+46.18%	+160.76%	8°C
DVFS 4 CPUs @2.13GHz	**+53.46%**	**+36.28%**	**+109.15%**	**4°C**
DVFS 2 CPUs @1.60GHz	+60.02%	+51.14%	+141.87%	3°C
Pinning VMs to 3 CPUs	**+18.57%**	**+16.23%**	**+37.83%**	**2.5°C**
Pinning VMs to 2 CPUs	+55.69%	+37.03%	+113.35%	6°C
Pinning VMs to 1 CPU	+165.74%	+108.89%	+455.11%	10°C

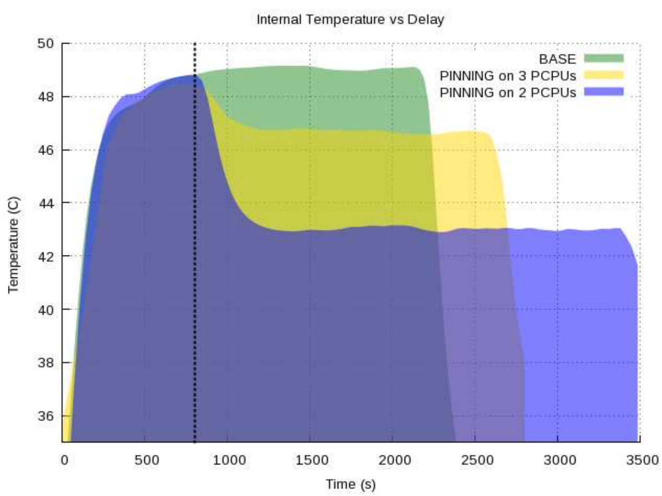


Fig. 8: Example of tradeoff between thermal efficiency and performance of pinning.

we present absolute values and for the different techniques we present the results relative the reference execution (in the form of % increased), except for temperature reduction. The best results of each technique are shown in bold.

As we commented previously, VM migration is the technique that achieves higher reduction of temperature. The price of migration overhead, we also obtained the shorter makespan migrating two VMs. However, it does not achieve the highest energy efficiency. The configuration where the destination server is empty is better than when the destination server is full in term of makespan but is worse in terms of energy efficiency. The configuration where the destination is full is the best case in term of energy efficiency (except for migration of 2 VMs) but the worse in terms of makespan. However, when the destination server is switched off, the penalty of starting up the server on both makespan and energy efficiency is significant. When the destination server is full, the number of migrated VMs does not influence significantly either the makespan or the energy efficiency. However, the higher the number of migrated VMs, the higher the reduction of temperature. Hence, the potential increase of temperature and penalty in existing VMs on the destination server should be analyzed in order to identify the tradeoffs between increasing the number of migrated VMs and the negative effects on the destination machine.

In most of the cases, DVFS obtains worse results than the other techniques. To obtain a similar reduction of temperature with DVFS, the penalty on both makespan and energy efficiency is higher. As we commented previously, running all CPUs at 2.13GHz works better than running 2 CPUs at 1.60GHz. In fact, if the workloads of the different VMs were dependent, the execution time of the workload when using 2 CPUs at 1.60GHz and using 4 CPUs at 1.60GHz would be similar. Pinning the VMs to 3 CPUs penalizes makespan only 18.57% (which is similar to the penalty when performing VM migration) but with the highest energy efficiency. However, the reduction of temperature is lower than with VM migration. The higher reduction of temperature is achieved when pinning the VMs to 2 CPUs; however, in such scenario the makespan increases significantly and becomes comparable to running all CPUs at 2.13GHz (which is the best case for DVFS). Pinning the VMs to only one CPU achieves a high decrease of temperature but the penalty on both makespan and energy efficiency (due to resource sharing and the associated problems such as context switches) is not acceptable. Thus, the threshold for applying the pinning technique is 2 CPUs for the HPL workload, which results on a temperature decrease of 6°C. This temperature decrease may be sufficient to react to many moderate hotposts.

Overall, the obtained results show that, depending on the temperature reduction required to mitigate the effects of a hotspot and the optimization goals (i.e., performance or energy efficiency), VM migration and pinning are the most effective techniques. The results also show that when there are available servers to migrate VMs and the main objective is optimizing performance (i.e., minimizing the makespan), it may be better migrating VMs rather than other techniques. However, when the focus is energy efficiency, pinning techniques may be a preferable technique in favor of VM migration. Furthermore, when VM migration is not feasible, pinning is the most effective mechanism to reduce the server's temperature while balancing between performance and energy efficiency.

V. Conclusions and Future Work

In this paper, we have studied different reactive thermal management techniques for virtualized and instrumented datacenters from the energy perspective. We have also studied the tradeoffs between performance, energy efficiency, and thermal efficiency of the techniques for HPC workloads. Specifically, we considered VM migration, DVFS at the server side, and pinning techniques as mechanisms to alleviate the cloud datacenter's servers from thermal anomalies (i.e., hotspots). The results of our evaluation conducted on real hardware using HPC applications showed that pinning is an effective mechanism to react to thermal anomalies under certain conditions. We have also showed that there are tradeoffs between the different analyzed mechanisms depending on the system characteristics and optimization goals.

Current and future research efforts include considering different workloads types with different demand for resources and resource utilization patterns (e.g., memory- or I/O-intensive workloads) and different lengths in order to define models for VM allocation at the datacenter level. We also plan to coordinate our reactive thermal management techniques with other layers (e.g., application layer) using a cross-layer design approach. Finally, we plan to implement an autonomic management system to take decisions and to act at the different layers.

Acknowledgments

This research was conducted within the NSF Center for Autonomic Computing and is supported in part by The Extreme Scale Systems Center at ORNL, by the Department of Defense, and by an IBM Faculty Award. This material was based on work supported by the National Science Foundation, while co-author M. Parashar was working at the Foundation. Any opinion, finding, and conclusions or recommendations expressed in this material; are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

[1] "Report to congress on server and data center energy efficiency," U.S. Environmental Protection Agency, Tech. Rep., August 2007.

[2] R. K. Sharma, C. E. Bash, C. D. Patel, R. J. Friedrich, and J. S. Chase, "Balance of power: Dynamic thermal management for internet data centers," *IEEE Internet Computing*, vol. 9, no. 1, pp. 42–49, 2005.

[3] J. Moore, J. Chase, P. Ranganathan, and R. Sharma, "Making scheduling "cool": temperature-aware workload placement in data centers," in *Annual Conf. on USENIX Annual Technical Conf.*, 2005, p. 5.

[4] Q. Tang, S. K. S. Gupta, and G. Varsamopoulos, "Energy-efficient thermal-aware task scheduling for homogeneous high-performance computing data centers: A cyber-physical approach," *IEEE Trans. Parallel Distrib. Syst.*, vol. 19, no. 11, pp. 1458–1472, 2008.

[5] J. D. Moore, J. S. Chase, and P. Ranganathan, "Weatherman: Automated, online and predictive thermal mapping and management for data centers," in *Intl. Conf. on Autonomic Computing*, 2006, pp. 155–164.

[6] C. Bash and G. Forman, "Cool job allocation: Measuring the power savings of placing jobs at cooling-efficient locations in the data center," in *USENIX Annual Technical Conf.*, 2007, pp. 363–368.

[7] T. Heath, A. P. Centeno, P. George, L. Ramos, Y. Jaluria, and R. Bianchini, "Mercury and freon: temperature emulation and management for server systems," in *ASPLOS*, 2006, pp. 106–116.

[8] L. Ramos and R. Bianchini, "C-oracle: Predictive thermal management for data centers," in *Intl. Symp. on High-Performance Computer Architecture*, 2008, pp. 111–122.

[9] R. Raghavendra, P. Ranganathan, V. Talwar, Z. Wang, and X. Zhu, "No "power" struggles: coordinated multi-level power management for the data center," *SIGOPS Oper. Syst. Rev.*, vol. 42, no. 2, pp. 48–59, 2008.

[10] M. Zhao and R. J. Figueiredo, "Experimental study of virtual machine migration in support of reservation of cluster resources," in *VTDC '07: Proceedings of the 2nd international workshop on Virtualization technology in distributed computing*, 2007, pp. 1–8.

[11] Y. Song, Y. Sun, H. Wang, and X. Song, "An adaptive resource flowing scheme amongst vms in a vm-based utility computing," in *IEEE Intl. Conf. on Computer and Information Technology*, 2007, pp. 1053–1058.

[12] S. Kumar, V. Talwar, V. Kumar, P. Ranganathan, and K. Schwan, "vmanage: loosely coupled platform and virtualization management in data centers," in *Intl. Conf. on Autonomic Computing*, 2009, pp. 127–136.

[13] A. A. Soror, U. F. Minhas, A. Aboulnaga, K. Salem, P. Kokosielis, and S. Kamath, "Automatic virtual machine configuration for database workloads," in *ACM SIGMOD Intl. Conf. on Management of data*, 2008, pp. 953–966.

[14] G. Laszewski, L. Wang, A. J. Younge, and X. He, "Power-aware scheduling of virtual machines in dvfs-enabled clusters," in *Cluster Computing*, 2009, pp. 1–10.

[15] Y. Chen, A. Das, W. Qin, A. Sivasubramaniam, Q. Wang, and N. Gautam, "Managing server energy and operational costs in hosting centers," in *ACM SIGMETRICS Intl. Conf. on Measurement and modeling of computer systems*, 2005, pp. 303–314.

[16] P. Ranganathan, P. Leech, D. Irwin, and J. Chase, "Ensemble-level power management for dense blade servers," *SIGARCH Comput. Archit. News*, vol. 34, no. 2, pp. 66–77, 2006.

[17] R. Nathuji and K. Schwan, "Virtualpower: coordinated power management in virtualized enterprise systems," in *ACM SIGOPS Symp. on Operating Systems Principles*, 2007, pp. 265–278.

[18] C. Rusu, A. Ferreira, C. Scordino, and A. Watson, "Energy-efficient real-time heterogeneous server clusters," in *IEEE Real-Time and Embedded Technology and Applications Symp.*, 2006, pp. 418–428.

[19] J. O. Kephart, H. Chan, R. Das, D. W. Levine, G. Tesauro, F. Rawson, and C. Lefurgy, "Coordinating multiple autonomic managers to achieve specified power-performance tradeoffs," in *Intl. Conf. on Autonomic Computing*, 2007, p. 24.

[20] R. Das, J. O. Kephart, C. Lefurgy, G. Tesauro, D. W. Levine, and H. Chan, "Autonomic multi-agent management of power and performance in data centers," in *Intl. joint Conf. on Autonomous agents and multiagent systems*, 2008, pp. 107–114.

[21] M. Steinder, I. Whalley, D. Carrera, I. Gaweda, and D. Chess, "Server virtualization in autonomic management of heterogeneous workloads," in *IEEE Symp. on Integrated Network Management*, 2007, pp. 139–148.

[22] R. Gonzalez and M. Horowitz, "Energy dissipation in general purpose microprocessors," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 9, pp. 1277–1284, 1996.