Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 2, Issue. 3, March 2013, pg.46 – 51

SURVEY ARTICLE

Energy Awareness in HPC: A Survey

A. R. Surve[1], A. R. Khomane[2], S.D. Cheke[3]

[1]**Computer Science & Engineering, Walchand College of Engineering, Sangli**

a_nilsurve@rediffmail.com

[2]**Computer Science & Engineering, Walchand College of Engineering, Sangli**

amit.khomane@gmail.com

[3]**Computer Science & Engineering, Walchand College of Engineering, Sangli**

shailesh.cheke@gmail.com

Abstract— Along with the increased awareness of energy and cost related to it, power management becomes a big issue for high performance computing. Power control is becoming a key challenge for effectively operating a modern high end computing infrastructures such as server, clusters, data centers and Grids. In addition to reducing operating costs, precisely controlling power consumption is an essential aspect in the field of High performance computing. Building energy efficient computer infrastructure is next major goal of the high performance computing. In recent years many researchers have been taking keen interest into developing sustainable, energy efficient high performance architecture. We, in this paper we are surveying energy efficient techniques for cluster computing and try to classify these techniques.

Key Terms: - Cluster computing, High Performance Computing (HPC), Power Aware HPC, Energy efficient HPC, Green Computing, Green HPC.

I. Introduction

The tremendous increase in computer performance has come with an even greater increase in power usage. Outcome of it is power consumption has become a primary concern in high end computing. According to Eric Schmidt, Executive Chairman of Google, what matters most to Google "is not speed but power—low power, because data centers can consume as much electricity as a city" [1]. This does not conclude performance is secondary objective but makes power as a constraint in increasing performance.

Before going much deeper into taxonomy, let's give a look at variables that influent the energy efficiency in any computer system. The variables can be categories broadly into two areas based on where the question of energy efficiency is addressed. These are Hardware related variables and Software related variables. These categories further divided into sub classes. How power consumption for each category is reduced is given in Table 1 [7]. Detailed summary is given in Table 1.

Hardware	Silicon Process Technology	• Second generation strained silicon • Improved interconnects
	Chip Technology	• Dynamic sleep transistor • Demand based switching • On-die voltage regulation • Multi-core and clustered micro-architecture • Power Gating, Macro Fusion.
	Power Management	• Voltage Regulation Technology • Improved display power specs • Thermal design for advanced heat-sync technology
Software	Operating System	• Developing power conscious device drivers. • Tuning OS for less interference with a processor's low-power states. • Energy Aware Scheduling of Applications based on benchmarks.
	Applications	• Application code multi-threaded and multi-core ready. • Power monitoring and analysis tools. • Optimizing code for reducing CPU clock cycles. • Energy Aware Scheduling of Applications tasks.

Table 1. Variables Influencing Energy-Efficiency

Considering more about modern high end computing infrastructures, we try to differentiate study and research directions in energy awareness. Here we are inspecting systems such as compute servers; compute clusters, distributed virtualized infrastructures, data centers, computational Grids & Clouds. But this paper primarily focuses on energy awareness in cluster computing. Further, this paper is organized as *II. Research Directions in Energy Aware HPC, III. Cluster Computing, IV. Adaptive Resource Management, V. Programming Language Support, VI. DPM Vs DVFS-enabled methods.*

II. RESEARCH DIRECTIONS IN ENERGY AWARE HPC

In recent years a lot of work has been done concerning energy awareness in high end computing. These related research efforts are in several dimensions. In this paper, first we try to categories these studies of energy-aware high performance computing. A survey of techniques used in implementing power awareness in high performance computing can be categories based on methodologies adopted. These are hardware reconfiguration and operation, virtual machine migration and resource consolidation, programming language and runtime support, and adaptive workload distribution and system management pictorially shown in Figure 1[16].

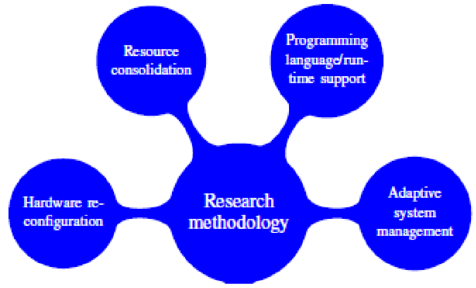

Figure 1. Taxonomy of Research Methodology in HPC

- *Hardware Reconfiguration:* System hardware can be reconfigured to low power states as per power requirements of the situation. This also gives benefit of minimum power on time since its quick to change states than to reboot the system.
- *Resource Consolidation:* In resource consolidation efficient use of resources is achieved by reducing total number of resources and consolidating more work on a resource. If multiple under-utilized resources are found then we try to consolidate their work on fewer numbers of resources and shutting down free resources. This may lead to complex configuration of data and applications in order to support the consolidation decisions. Dynamic check pointing can be the solution with overhead of memory and frequent updates to check point.
- *Programming Language support:* For power management using programming language support, programming modules are designed so that they give importance to power saving also. For example sequence matching tools in bioinformatics can be configured to use resources with power constraints.
- *Adaptive System Management:* A system that can change itself in response to changes in its environment in such a way that its performance improves through a continuing interaction with its surroundings.

III. CLUSTER COMPUTING

In coming chapters we will give importance to power management in Cluster computing. We will present some of the major works in each Methodology of power management in cluster computing. In Cluster computing, the methodologies adopted are as shown in the Figure 2 [16].

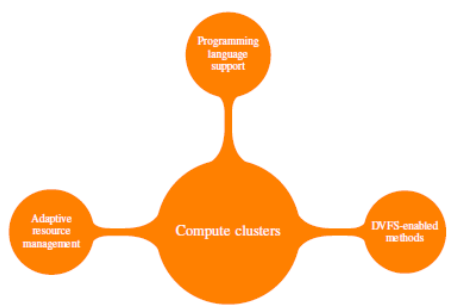

Figure 2. Classification of Energy aware Cluster systems

- *Adaptive Resource Management:* A system which adapts changes in its environment to improve the performance parameters in consideration. In our case energy efficiency.

- *Programming Language Support:* Programming languages can modified or equipped with modules can give importance to parameters that are of interest like energy efficiency.

- *DVFS-enabled Methods:* Dynamic Voltage and Frequency Scaling is a framework to change the frequency and operating voltage of processors based on system requirement at given time.

IV. Adaptive Resource Management

A system that can change itself in response to changes in its environment in such a way that its performance improves through a continuing interaction with its surroundings. In adaptive Resource Management we try to change behavior of the resources based on the system parameters.

Cluster wide load balancing and unbalancing is to eventually spread the work load among the available nodes or to consolidate the work to fewer nodes [2] [3] [4] [5]. Adapting through idleness if the system and changing its power state to low power is implemented in [12] [13]. In [2] implemented load concentration or unbalancing operation that saves the power consumed by the powered-down nodes, but can degrade the performance of the remaining nodes and potentially increase their power consumption.

Thus, load concentration involves an interesting performance vs. power tradeoff. The implementations were performed in two ways: (1) at the application level for the network server; and (2) at the OS level for the cycle server. The experiment was conducted on 8 node system. The result provided states as mentioned in Table 2 [2], with setting the threshold for performance degradation is less than 20% in both the cases.

Reduction in (%)	Implementation 1	Implementation 2
Power	86	43
Energy	86	32

Table 2. Results of Experiment

V. Programming Language Support (Scheduling)

Programming language and matching tools can be modified to give special importance to power policies of the system. MPI programs (or other parallel modules) or sequence matching tools are considered my many researchers for power management implementation. More about algorithms taxonomy is given in Figure 3.

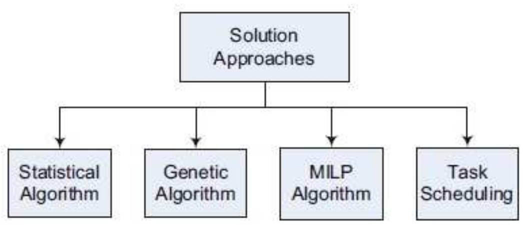

Figure 3. Solution Approaches

Some runtime systems are developed for power management of scientific applications [8][9]. These implementations develop some online system information logging, analysis and application prediction.

The paper [10] discusses power prediction techniques for MPI applications based on task aggregation and [11] proposes power aware Hybrid MPI/OpenMP programming pattern based on the interaction between communication and computation without using DVFS.

Programming approaches that can be used to address this problem in the software – application layer as illustrated in Figure 3.Research in this direction focuses on the software – application area and specifically tries to address the question of energy aware scheduling of application tasks. For power aware scheduling the scheduling logic is written into the scheduler considering power as one of the constraints for scheduling. Figure 4. shows the simple system.

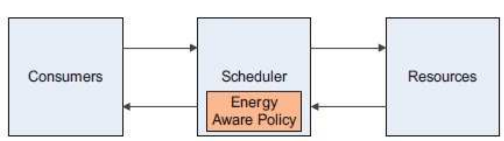

Figure 4. Energy Aware Scheduling System

Energy aware Policy can take any of the approach shown in the Figure 3. In Bioinformatics domain, for matching sequences, taking energy awareness into consideration is presented in [7].

Well-known sequence matching tool BLAT is used to match sequence with deadline and for energy awareness EAS (Energy Aware Scheduling) is proposed [7]. EAS system is two phase system (1) an Offline phase and (2) an online phase. Offline phase parallelize the sequence to understand speedup (run profile) of the program. EAS engine then use this information to generate initial schedule. In online phase is a feedback mechanism is incorporated. As scheduled tasks are completed, their actual execution time (AET) is used to adjust the resources required for scheduling remaining tasks using the least number of nodes while meeting a given deadline. These techniques have shown satisfactory results.

VI. DPM VS DVFS- ENABLED METHODS

In this we examine two types of hardware mechanisms allowing us to manage CPU energy consumption: Dynamic Voltage and Frequency Scaling (DVFS) and Dynamic Power Management (DPM).

DPM allows temporarily suspending its operation and putting it in low-power states. Whereas DVFS allow us to dynamically throttle the voltage and frequency of the CPU. As we throttle or suspend the processor we inevitably sacrifice execution speed, and so the key challenge in designing DVFS/DPM control algorithms is to find the setting that saves the most power while sacrificing the least speed. Dynamic Voltage and Frequency Scaling (DVFS) is accepted as a technique to reduce power and energy consumption of microprocessors [17].

DVFS efficiently reduce processor power dissipation [14][15]. DVFS saves processor energy consumption by varying the frequency and voltage of a microprocessor in real-time according to processing needs. The DVFS approach is limited by the structure of the applications. Only configurable applications can be inserted DVFS instructions for power control. Lowering only the operating frequency f_{clk} (operating frequency) can reduce the power consumption but the energy consumption remains the same because the computation needs more time to finish. Whereas DPM can be applied to any configuration but need to analyze system usage over the period of time.

VII. SUMMARY AND DISCUSSION

In power aware cluster computing extensive research is being carried out. But there are still some open questions [7], such as,

- How to monitor the “energy-aware” aspects of a parallel computing system in multiple scales? There are multiple interesting system indexes, for example, power, temperature, and humidity. This monitoring process will be taken in multiple scales.
- What is the energy usage profile and compute cost model for high performance applications in a large scale distributed compute system? There are still no effective methods to estimate energy usage and cost for users.
- Developing effective tools, middleware and software framework to support energy efficient high performance applications in a parallel computing system.

REFERENCES

[1] J. Markoff and S. Lohr. Intel’s huge bet turns iffy. New York Times Technology Section, September 29, 2002. Section 3, Page 1, Coumn 2.

[2] Eduardo Pinheiro, Ricardo Bianchini, Enrique V. Carrera, and Taliver Heath. Load balancing and unbalancing for power and performance in cluster-based systems. In In Workshop on Compilers and Operating Systems for Low Power, 2001.

[3] C. Hsu and W. Feng. A power-aware run-time system for high performance computing. In Proceedings of the 2005 ACM/IEEE conference on Supercomputing. IEEE Computer Society Washington, DC, USA, 2005

[4] A. Bestavros, M. Crovella, J. Liu, and D. Martin. Distributed Packet Rewriting and its Application to Scalable Server Architectures. In Proceedings of the International Conference on Network Protocols, October 1998

[5] Michael J. Litzkow and Marvin Solomon. Supporting Checkpoint and Process Migration Outside the UNIX Kernel. In Usenix Conference Proceedings, pages 283–290, San Francisco, CA, Jan 1992

[6] Vincent W. Freeh and David K. Lowenthal. Using multiple energy gears in mpi programs on a power-scalable cluster. In Keshav Pingali, Katherine A. Yelick, and Andrew S. Grimshaw, editors, PPOPP, pages 164–173. ACM, 2005.

[7] Sachin Pawaskar and Hesham H. Ali, A Dynamic Energy-Aware Model for Scheduling Computationally Intensive Bioinformatics Applications, HPCS 2010.

[8] Chung-Hsing Hsu and Wu chun Feng. A power-aware run-time system for high-performance computing. In SC, page 1. IEEE Computer Society, 2005.

[9] T. Li and L. K. John. Run-time modeling and estimation of operating system power consumption. In SIGMETRICS International conference on Measurement and modeling of computer systems, pages 160–171, New York, NY, USA, 2003. ACM.

[10] Dong Li, Dimitrios S. Nikolopoulos, Kirk W. Cameron, Bronis R. de Supinski, and Martin Schulz. Power-aware mpi task aggregation prediction for high-end computing systems. In IPDPS, pages 1–12, 2010.

[11] Dong Li, Bronis R. de Supinski, Martin Schulz, Kirk W. Cameron, and Dimitrios S. Nikolopoulos. Hybrid mpi/openmp power-aware computing. In IPDPS, pages 1–12, 2010.

[12] Zhuo Liu, Aihua Liang, Limin Xiao, Li Ruan . Workload-aware Power Management of Cluster Systems.DCABES. 2010.

[13] Hewlett-Packard Corporation, Intel Corporation, et al. “Advanced Configuration and Power Interface Specification (Revision 4.0),” June 2009.

[14] Joseph L. Hellerstein, Yixin Diao, Sujay Parekh, and Dawn M. Tilbury.Feedback Control of Computing Systems. John Wiley & Sons, 2004.

[15] C. Hsu and W. Feng. A power-aware run-time system for high performance computing. In Proceedings of the 2005 ACM/IEEE conference on Supercomputing. IEEE Computer Society Washington, DC, USA, 2005.

[16] Chang Cai, Lizhe Wang, Samee U. Khan, Jie Tao. Energy-aware High Performance Computing – A Taxonomy Study.

[17] C.H Hsu, Compiler-Directed Dynamic Voltage and Frequency Scaling for CPU Power and Energy Reduction, Ph. D Dissertation, the State University of New Jersey, USA 2003.

