

An integrated framework to measure the energy consumption of a parallel application on a computational grid

D.B Srinivas†, Puneeth R.P††, Rajan M.A†††, Sanjay. H.A††††

†Nitte Meenakshi Institute of Technology, Bangalore,India

††Nitte Meenakshi Institute of Technology, Bangalore,India

†††Tata Consultancy Services, Bangalore,India

††††Nitte Meenakshi Institute of Technology, Bangalore,India

Summary

Grid computing is the forerunner for deploying high performance scientific parallel applications. Execution of parallel application on computational grids generally consumes enormous amount of energy. Energy requirements for computation grids are increasing in nonlinear fashion due to massive digitization of the physical world. As a result of this Capital Expenditure (CAPEX) and Operational Expenditure (OPEX) of computational grid requirements are also increasing. One way to reduce this is to minimize the energy consumption. To envisage this, it is essential to understand the energy consumption by parallel applications. As part of this work, we proposed a novel energy measurement framework based on software approach for parallel applications. Further we validate the framework for measuring energy consumption of a parallel Message Passing Interface (MPI) application on a computational grid with varying workloads.

Key words:

computational grid, parallel application, energy, clock cycles..

1. Introduction

Computational grid is composed of a set of heterogeneous nodes that are spatially distributed for parallel applications. Hence grid computing is an attractive domain for High Performance Computing (HPC) applications. In recent years, new computationally intensive scientific applications are becoming more and more prominent [1]. Due to the technological advancements in Information Communication Technologies (ICT), people around the world are increasingly using digital applications for their day-to-day activities. To support this, telecom operators are building large data centers that are interconnected across the world using computational grids. Further it is estimated that Internet users are demanding high bandwidth and which is growing at a rate faster than that of processor speed. Therefore to enable this, the best way is to interconnect the high speed computing nodes to Internet in an efficient manner [2]. From these requirements, it is evident that grid computing is a good paradigm to enable an efficient way of managing large data centers to enhance users' digital experience. To

accomplish this, large number of parallel applications are running on computational grids.

To run parallel applications, computational grids demand high energy. Further, some of the major consumers of these grids are high-demanding scientific parallel applications (weather forecasting, genome, social networks, etc) and large data centers (for telecom applications, wherein high virtualization is required). Parallel applications demand high availability of the grids. Consequently grids draw high energy continuously till the completion of a job and thus incur huge OPEX cost to maintain grid. This is evident from the following facts.

According to 2007 EPA report, U.S. data centers alone consumed 61 billion kWh in 2006; which is enough energy needed to power 5.8 million average households. To add to the worries, it is estimated that IT industry energy consumption is projected to grow exponentially in coming years. Major chunk of energy consumption in IT industry is due to their large data centers setup and they consume most of their power budget. Thus reducing data center energy consumption is a major issue in the IT industry at present and in future. Further, in 1998, heat loads for dense rack-mount servers hovered around 5000W per rack and by 2006 it increased to 32,000 W per rack [3].

From these observations, it is evident that optimizing energy consumption is very important. In general this can be achieved in two ways: (i) Designing energy-efficient grid processors. (ii) Minimizing the energy leakages (due to underutilization of grid, improper scheduling etc.).

The former approach is feasible for new data centers. For existing data centers the latter approach is preferred. To determine energy leakages, energy consumption needs to be measured. In this article, we work towards understanding energy consumption of a parallel application on a computational grid. In general energy consumption depends on the following parameters.

1) Supply voltage: Variation in supply voltage leads to a large fluctuations in energy usage. (2) Temperature: Though in general temperature effect is considered negligible, it becomes a significant factor for high performance parallel applications. (3) Electrical specifications: Components with different specifications in

a system and their usage characteristics causes reasonable amount of variations in energy consumption. (4) Frequency: The goal of modern processors is to deliver good performance by keeping energy consumption within reasonable limits. In general, application performance is directly proportional to processor frequency. Operating processor at high frequency affects health of the processor: such as lifetime of a processor; reliability of the processor etc. Hence, as a result of this, it incurs additional CAPEX to manage a computational grid. Accordingly, there are techniques wherein energy consumption of an application can be reduced without performance penalty by operating a processor at an optimal frequency.

Therefore by configuring the above parameters suitably significant amount of energy consumption can be reduced. To enable this, proper energy measurement technique is essential.

Thus in this article, we propose a novel framework to measure energy consumption of a parallel application on a computational grid, which in turn can be used as a tool to understand energy leakages. By using this, grids can be managed efficiently with respect to energy usage. As we understand, the proposed energy measurement framework for parallel applications is novel.

Proceeding of the paper is organized as follows. Section 2 describes related work. Section 3 defines a system model and experimental methodology to measure the energy consumption of a parallel application. In Section 4 experimental setup and results are discussed. Finally in Section 4, we summarize with conclusion and future work.

2. Related Work

In [4] energy consumption of the linear program in independent processing systems is approximated. Bona M. et al. proposed [5] a system of methods for energy estimation at instruction level for Very long Instruction Word (VLIW) processors. The proposed method helps to enable low power software and hardware effectively for Instruction Level Parallelism (ILP) processors.

In [6] authors highlighted the need for higher processor performance at different frequencies of a processor. Van Bui et al. [7] presented a Tuning and Analysis Utilities (TAU), a portable application profiling tool kit for performance analysis. It obtains various details related to performance of parallel applications. They used Thermal Design Power (TDP) metric as maximum power value for a processor. Vicente Blanco, et al. [8] proposed systematic user driven framework to obtain analytical model of MPI (Message Passing Interface) applications on parallel systems. This approach consists of two phases. In the first phase, instrumentation of the source code is performed by using CALL, which is a profiling tool for interacting with the code, which obtains different performance metrics and

stores the performance information in XML files. Using this information, an analytical model of the performance behavior is done in the second phase.

Energy consumption by data centers are simulated using data center simulators [9][10]. This helps to understand energy consumption patterns in data centers. Hybrid applications are [11] used to investigate performance models and application characteristics (performance counters) which affects power consumption of system, processor and memory. Here it uses Multi-core Application Modeling Infrastructure (MuMI). MuMI uses Performance Application Programming Interface (PAPI) and PowerPack to provide systematic measurement and modeling of power consumption. Further performance power tradeoffs in Multi-core systems. Study of basic MPI-I/O primitives running on a PC cluster based on the Network File Systems (NFS) and Parallel Virtual File Systems (PVFS) [12] are discussed. This work helps application developers to tune the file system configuration and selects the best I/O routine to improve I/O performance.

In [13] divide and conquer approach is used. Here large problems are divided into sub programs, which are then solved concurrently to minimize response time by taking advantage of non-local resources and overcoming memory constraints. The main goal is to form a cluster-oriented parallel computing architecture for MPI based applications which demonstrates the performance gains achieved through parallel processing using MPI.

3. Proposed Energy Measurement Framework for Parallel Applications in a grid

In this section we propose energy measurement framework for parallel application running on a computational grid along with experimental setup.

3.1. System Model

Proposed energy measurement framework for parallel application in a grid is as shown in Figure 1. It consists of n number of clusters ($C_1, C_2, \dots, C_n$) with each computational cluster C_i contains m_i number of processing elements. Note that p_{ij} denotes power of processing element j in the cluster C_i . Aggregated energy consumed by a grid for running parallel application is given by

$$E = \sum_{i=1}^n \sum_{j=1}^{m_i} p_{ij} \left(\frac{\Delta t_{ij}}{s_{p_{ij}}} \right) \quad (1)$$

Δt_{ij} : Total number of processor cycles utilized by p_{ij} processing element

$s_{p_{ij}}$: Operating speed of a processing element j in the cluster C_i .

In general Power (P) is measured as

$$P = D_w ((F_{min}/F_{max}) (V_{min}/V_{max})^2) \quad (2)$$

where D_w , F_{min}/F_{max} and V_{min}/V_{max} are default wattage, Minimum/Maximum operating frequency and Minimum/Maximum operating voltage of a processor respectively.

Using (2) in (1), we get

$$E = \sum_{i=1}^n \sum_{j=1}^{m_i} D_{wij} (F_{ij}^{min}/F_{ij}^{max}) (V_{ij}^{min}/V_{ij}^{max})^2 \left(\frac{\Delta t_{ij}}{S_{pij}} \right) \quad (3)$$

where D_{wij} , $F_{ij}^{min}/F_{ij}^{max}$ and $V_{ij}^{min}/V_{ij}^{max}$ are default wattage, Minimum/Maximum operating frequency and Minimum/Maximum operating voltage of a processing element j in the cluster C_i respectively.

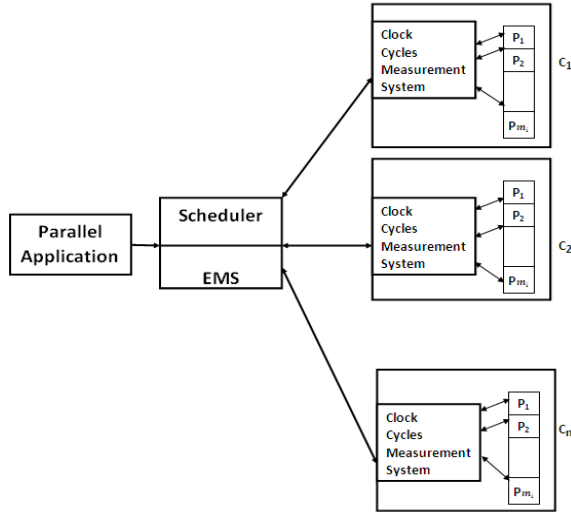


Figure 1. Energy measurement framework.

Proposed energy measurement framework consists of a scheduler which allocates parallel application to computing recourses. When a parallel application is executed on clusters of nodes using the proposed framework, the clock cycles measurement system (using Linux perf tool) measures clock cycles on each cluster. Finally Energy Measurement System (EMS) collects clock cycles from all the clusters and computes energy consumption using equation (3).

3.2. Experimental Methodology

To validate the proposed energy measurement framework, we follow the steps described as shown in Figure 2 to measure clock cycles and compute energy consumption. Further we run a computationally intensive scientific parallel application Molecular Dynamics simulation of Lennard-Jones system [14]. It is a systolic algorithm using MPI, consists of γ number of particles and they are

distributed equally among the processes running on parallel nodes.

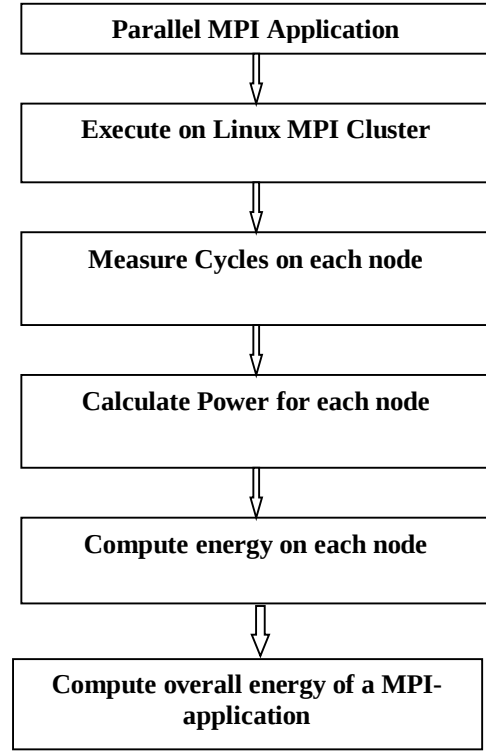


Figure 2. Energy measurement steps.

4. Experimental Setup and Results

To evaluate the framework we run Lennard-Jones Molecular Dynamics (MD) parallel application with the following experimental set up. It consists of a cluster comprising three nodes connected via the Ethernet. Configurations of these nodes are as shown in Table 1. Performance of this parallel application is measured through clock cycles and tabulated in Table 2. It is evident that the complexity of the algorithm is proportional to the number of particles and hence clock cycles utilized by the application are increases with number of particles. It is evident that from Figure 3 and Figure 4 that as number of particles in MD simulation increases elapsed time increases and consequently energy consumption by the application is also increases.

Table 1. Specifications of the machines used for the experiments.

CPU model	Process or operating speed	D_w (processor)	F_{min}	F_{max}	V_{min}	V_{max}
Intel(R) Core(TM) 2 DUO CPU E8200	2.66 GHz	64 W	2.66 GHz	2.87 GHz	0.85 V	1.36 V

Table 2. Clock cycles on each node.

Number of particles	Number of clock cycles on node ₁	Number of clock cycles on node ₂	Number of clock cycles on node ₃
300	127271257976	127265486080	127265673704
600	376022202536	376009927728	376009722984
900	766593535048	766573109584	766572330304
1200	1258964941928	1258934559208	1258933119880
1500	1932391394240	1932346173704	1932344043744

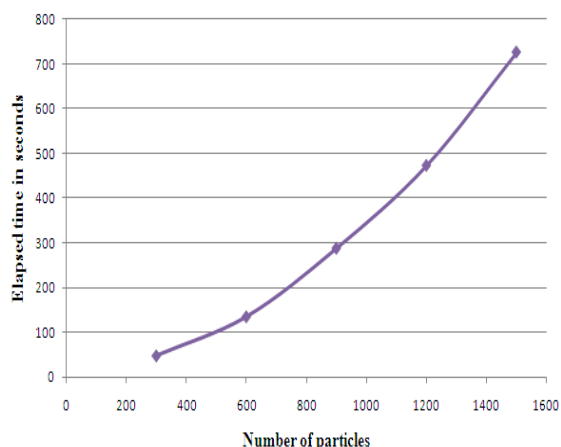


Figure 3. Elapsed time at different workloads.

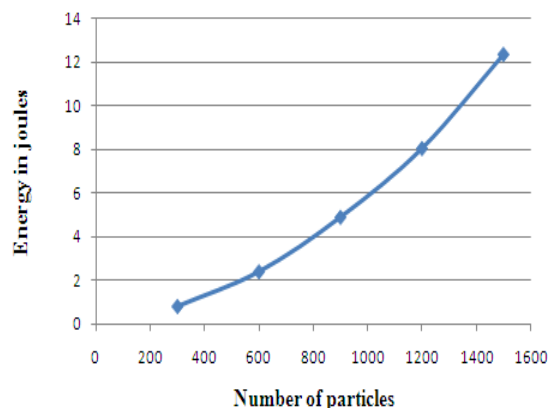


Figure 4. Energy consumption with different workloads.

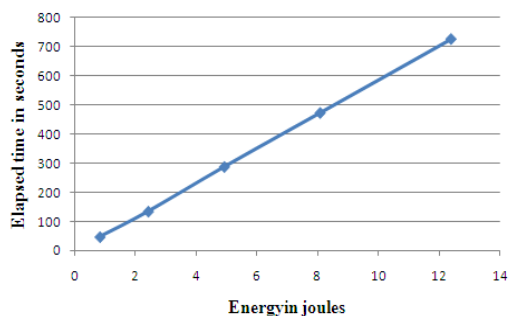


Figure 5. Energy consumption with respect to elapsed time.

Further from equation (3) we infer that energy is propositional to elapsed time and same is experimentally demonstrated in Figure 5. Finally we demonstrated that the energy consumption of a parallel application is measured by using processor clock cycles.

5. Conclusion and Future Work

In this article we presented a framework to measure the energy consumption of a parallel MPI application on a computational grid. A software approach is used to measure the number of clock cycles of processor on a cluster of processors. The energy measurement is carried out with various workloads on a set of homogeneous nodes. Further we extend this work for heterogeneous set of nodes. Importance of this work is reasoned for understanding, analyzing and scheduling of a parallel application on a computational grid. Future work includes designing energy-efficient scheduler for parallel applications which minimizes the energy consumption on a computational grid.

References

- [1] Fang Dong, JunzhouLuo, Aibo Song, Jiuxin Cao and Jun Shen, An Effective Data Aggregation Based Adaptive Long Term CPU Load Prediction Mechanism on Computational Grid, Journal of Future Generation Computer Systems, Elsevier BV, Vol. 28, No. 7, pp. 1030-1044, 2012.
- [2] D. Laforenza. Grid programming: some indications where we are headed, ACM Parallel Computing Special Issue: Advanced environments for parallel and distributed computing, Vol. 28, Issue 12, pp.1733-1752, 2002.
- [3] Report to Congress on Server and Data Center Energy Efficiency, https://www.energystar.gov/ia/.../prod.../EPA_Datacenter_Report_Congress_Final1.pdf.
- [4] S. Nikolaidis and Th. Laopoulos "Instruction-Level Power Consumption Estimation Embedded Processors Low-Power Applications", International Workshop on intelligent Data Acquisition and Advanced Computing Systems Technology and Applications, Vol. 24 Issue 2, pp. 133 – 137, June 2002.
- [5] Bona M. Samit D. Sciutot C. Silvanog V AZaccariar R. Zafalont ,” Energy Estimation and Optimization of Embedded VLIW Processors based on Instruction Clustering in DAC2002”, pp. 886 - 891, New Orleans, Louisiana, USA.
- [6] Intel Corp. Datasheet: Intel R 4 Pentium R 4 Processor with 512-KB L2 cache on 0.13 Micron Process at 2GHz–3.06 GHz. Available at <http://www.intel.com/design/pentium4/-Datashts/298643.htm>, Nov. 2002.
- [7] Van Bui , Boyana Norris, Kevin Huck, Lois Curfman McInnes, Li Li, Oscar Hernandez, Barbara Chapman “A Component Infrastructure for Performance and Power Modeling of Parallel Scientific Applications Proceedings of the 2008 compFrame/HPC-GECO workshop on Component based high performance Article No. 6 USA, 2008.

- [8] Diego R. Martinez, Vicente Blanco, Marcos Boullon, Jose C. Cabaleiro, Software Tools for Performance Modeling of Parallel Programs, IEEE International Parallel and Distributed Processing Symposium, pp. 1 - 8 Marcos Boullon, Spain, 2009.
- [9] MuratKocaoglu, Derya Malak, Ozgur B. Akan Fundamentals of green communications and computing: modeling and simulation, Computer Vol,45 , Issue: 9, pp. 40-46, 2012.
- [10] Baris Aksanli ,Jagannathan Venkatesh, Tajana Šimunić Rosing B, Using datacenter simulation to evaluate green energy integration, Computer 45 (9), 2012, 56-64.
- [11] Charles Lively, Xingfu Wu, Valerie Taylor, Shirley Moore, Hung-Ching Chang, Chun-Yi Su, Kirk Cameron, Power-Aware Predictive Models of Hybrid MPI/OpenMP) Scientific Applications on multicore Systems, Computer Science - Research and Development November 2012, Volume 27, Issue 4, pp 245-253.
- [12] Jacobo barro, Juan touriño, Ramon doallo and Victor m. gualias, Performance Modeling and Evaluation of MPI-I/O on a Cluster, journal of information science and engineering 18, 825-836, 2002.
- [13] Nanjesh, B. R., K. S. V. Kumar, C. K. Madhu, and G. H. Kumar. MPI based cluster computing for Performance evaluation of parallel applications, Information & Communication Technologies ICT), IEEE Conference on11-12 April 2013, pp. 1123-1128.
- [14] *Molecular Dynamics simulation of Lennard-Jones system Systolic algorithm using MPI*, www.physics.buffalo.edu/phy516/Files/Topic2/md_systolic.cpp.



D B Srinivas received B.E and M.E degrees, in Computer Science and Engineering from Bangalore University in 1998 and 2002 respectively. Pursuing Ph.D in Computer Science and Engineering from Visvesvaraya Technological University, Belagavi, majoring in Grid and parallel computing. Currently, he is working as Assistant Professor in department of Information

Science and Engineering at Nitte Meenakshi Institute of Technology Bangalore.



Puneeth R P obtained his Bachelor's of Engineering and Master's of Technology in computer science and engineering from Visvesvaraya Technological University, he has also obtained EMC Academic Associate. Currently, he is a Assistant Professor in department of Computer Science and engineering in NMAM Institute of Technology, Karkala, Udupi574110. His Areas of interests are

parallel computing, Analysis and Designing of Algorithm, networking.



Rajan MA – Rajan is a Scientist at TCS Research, Bangalore since 2010. He has 16 years of experience in the area of cryptography, computer networks, spacecraft technology, Cross layer design, Number Theory, Graph Theory, Combinatorics, coding theory and Functional Analysis. He has BE, MTech. and PhD in Computer Science and Engineering and MSc, MPhil and PhD in Mathematics. During 2000-2005, he worked at the ISRO Satellite Centre (ISAC), Bangalore, India as a Scientist and was actively involved in realization of several spacecrafts. From September 2005, he is with TCS Bangalore. During 2005-10, his worked in the area of Optical Network Management System (ONMS) and was involved in design and implementation of ONMS software for Lucent, Bell Labs. At CTO Innovation Labs, he is involved in the design of efficient lightweight cryptographic algorithms for Internet of Things/M2M communication based on functional encryption which includes identity based elliptic curve cryptography, attribute based encryption. Currently he is working in the area of Homomorphic encryption, program obfuscation and privacy enhanced techniques. He is actively participating and contributing to standardization activities (GISFI, TSDSI) in the area of security for M2M workgroup. He has published several research papers in national and international conferences (IEEE Globecom, ANTS, AINA) and journals. He served as a reviewer, TPC member for IEEE COMSNETS, NCC. He has filed and obtained several patents at India and outside India as well.



Sanjay H A received the BE degree in electrical and electronics engineering from the University of Kuvempu, India, in 1996 and the M.Tech. degree in computer science and engineering from Visvesvaraya Technological University, India, in 2001. He has done his Ph.D. from Indian Institute of Science, Bangalore. His Research interests include Grid Computing, Cloud computing, Parallel Computing, Distributed Systems, and Performance Modeling of parallel applications. He is having several research publications in reputed international journals & conferences.