Performance of Scientific Processing in Networks of Workstations

Fernando G. Tinetti

Centro de Técnicas Analógico-Digitales (CeTAD)[1]
Laboratorio de Investigación y Desarrollo en Informática (LIDI)[2]
Grupo de Investigación en Procesamiento Paralelo (GIPP)[3]
fernando@lidi.info.unlp.edu.ar

1. Introduction

The growing processing power of standard workstations, along with the relatively easy way in which they can be available for parallel processing, have both contributed to their increasing use in computation intensive application areas. Usually, computation intensive areas have been referred to as scientific processing; one of them being linear algebra, where a great effort has been made to optimize solution methods for serial as well as for parallel computing [1] [3].

Since the appearance of software libraries for parallel environments such as PVM (Parallel Virtual Machine) [4] and implementations of MPI (Message Passing Interface) [5], the distributed processing power of networks of workstations has been available for parallel processing as well. Also, a strong emphasis has been made on the heterogeneous computing facility provided by these libraries over networks of workstations. However, there is a lack of published results on the performance obtained on this kind of parallel (more specifically distributed) processing architectures.

From the whole area of linear algebra applications, the most challenging (in terms of performance) operations to be solved are the so called Level 3 BLAS (Basic Linear Algebra Subprograms). In Level 3 BLAS, all of the processing can be expressed (and solved) in terms of matrix-matrix operations. Even more specifically, the most studied operation has been matrix multiplication, which is in fact a benchmark in this application area.

2. Characterization of Heterogeneous Computing

There are a number of distinguishing factors that characterize the heterogeneous computing hardware of a network of workstations such as processor, clock cycle, memory hierarchy, main memory size, etc. All of these factors affect the relative processing power of each workstation.

It is expected that intercommunication times between workstations are almost the same, given that the usual interconnection topology in a network of workstations is a 10 Mbits or 100 Mbits Ethernet bus. The network workload along with the different kinds of communication subsystems of workstations hardware make communication times not as similar as expected. The communication pattern of a parallel (distributed) computing may be affected by this kind of communication heterogeneity. Communication times are harder to characterize when more than a local area network of workstations is used. In this case, the communication times between two workstations are dependent on the physical location (i. e., LAN) where each of them resides.

---

1 Facultad de Ingeniería, Universidad Nacional de La Plata
2 Facultad de Informática, Universidad Nacional de La Plata
3 Dto. de Coordinación de Investigación y Desarrollo, Instituto Tecnológico de Bs. As.

Workstations heterogeneity usually implies software heterogeneity, basically at the operating system and development tools levels. Software heterogeneity also produces different ways in which the overheads (e.g. system calls) affect processing performance. Thus, the relative processing power is affected not only by the underlying hardware, but also by the software running along with the computing processes of the application. Development tools heterogeneity usually introduces some problems in the software development phase, but the influence on computing performance is not significant for scientific applications.

3. Parallel Scientific Applications on Heterogeneous Hardware

Parallel applications in general, and parallel scientific applications in particular, face some specific problems when the underlying computing hardware is heterogeneous in order to obtain a near optimal performance. Processing workload and communication workload are two of the most important factors affecting performance.

Traditionally, parallel applications have had homogenous hardware target machines (i.e., the same processors and a similar message communication time). Thus, algorithms have been designed assuming homogeneous hardware, and when used on heterogeneous hardware their performance is far from optimal. However, it should be pointed out that these traditional algorithms on heterogeneous hardware solve the same applications from the numerical point of view. This is one of the reasons for claiming the goodness of making heterogeneous hardware work as a parallel machine: it can solve the same problems as the traditional (and more expensive) parallel computers with minor adaptations of the algorithms.

Processing workload seems to be easily solved in the field of parallel scientific computing given that most of the programs fit the SPMD (Single Program Multiple Data) model. The key idea is based on sequential relative processing power amongst workstations. If a workstation ws_i is twice as fast as ws_j, then it should receive twice the workload of ws_j, which most of the times implies twice the data to process. Even if this seems to be reasonable, it has to be experimentally justified for the optimized numerical algorithms where, for example, memory hierarchy is strongly used to achieve near peak processor performance. When computing processes have to share the workstation with communication processes, the sequential relative processing power could change, and it could be necessary to define the parallel relative processing power.

Communication workload is particularly changed in networks of workstations. First, because Ethernet bus is the most common LAN architecture to which workstations are connected. Traditional scientific parallel algorithms are based mostly on static meshes or dynamic networks where physical communication is solved point to point without any interference from other communicating processors. Second, the heterogeneity given by more than one interconnected LAN (e.g. by means of Reuters for Internet traffic) has not been studied from the scientific processing performance point of view.

4. Areas for Experimentation

In order to establish an incremental research project, it is necessary to define a number of experimentation areas to make the effect of heterogeneous networks of workstations on the performance scientific applications clear:

- Selection of a specific application. Matrix multiply has been selected for its many representative characteristics of the scientific processing area [7].
- Analysis of the performance of traditional parallel algorithms without taking heterogeneity into account, or only taking into account balanced processing workload based on sequential relative processing power [6].
- Optimizations of scientific code for maximum sequential performance [2] [7].
- Relationship between sequential and parallel relative processing power.
- Impact of bus interconnection network topology on the parellelization of scientific algorithms.
- Impact of the interconnection network topology (including more than one LAN) on scientific processing performance.

5. References

[1] Anderson E., Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. DuCroz, A. Greenbaum, S. Hammarling, A. McKenney, D. Sorensen, LAPACK: A Portable Linear Algebra Library for High-Performance Computers, Proceedings of Supercomputing '90, pages 1-10, IEEE Press, 1990.

[2] Bilmes J., K. Asanovi*ƒ*, C. Chin, J. Demmel, Optimizing matrix multiply using phipac: a portable, high-performance, ansi c coding methodology, Proceedings of the International Conference on Supercomputing, Vienna, Austria, July 1997, ACM SIGARC.

[3 Blackford L., J. Choi, A. Cleary, E. D'Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, R. Whaley, ScaLAPACK Users' Guide, SIAM, Philadelphia, 1997.

[4] Dongarra J., A. Geist, R. Manchek, V. Sunderam, Integrated pvm framework supports heterogeneous network computing, Computers in Physics, (7)2, pp. 166-175, April 1993.

[5] Message Passing Interface Forum, MPI: A Message Passing Interface standard, International Journal of Supercomputer Applications, Volume 8 (3/4), 1994.

[6] Tinetti F., A. Quijano, A. De Giusti, Heterogeneous Networks of Workstations and SPMD Scientific Computing, 1999 International Conference on Parallel Processing, The University of Aizu, Aizu-Wakamatsu, Fukushima, Japan, September 21 - 24, 1999.

[7] Whaley R., J. Dongarra, Automatically Tuned Linear Algebra Software, Proceedings of the SC98 Conference, Orlando, FL, IEEE Publications, November, 1998.