BULETINUL Universităţii Petrol – Gaze din Ploieşti	Vol. LXI No. 1/2009	71 - 76	Seria Matematică - Informatică - Fizică

Availability of Computational Resources for Desktop Grid Computing

Monica Vlădoiu[*], Zoran Constantinescu[**], Cătălina Negoiţă[*]

[*] Universitatea Petrol-Gaze din Ploieşti, Bd. Bucureşti 39, Ploieşti, Catedra de Informatică
e-mail: mvladoiu@upg-ploiesti.ro

[**] ZealSoft SRL, str. Tg. Neamţ, nr. 11, Bucureşti
e-mail: zoran@unde.ro

Abstract

The computing power of a single desktop computer is insufficient for running complex algorithms. A very convenient solution is based on the use of non-dedicated desktop PCs in a Desktop Grid Computing environment. Here, the costs are highly distributed: every volunteer supports her resources (hardware, power, internet connections) while the benefited entity provides management infrastructures, (network bandwidth, servers, management services), receiving in exchange an enormous and otherwise unaffordable computing power. There has been little insight into the temporal structure of resource availability within an organizational setup. We present here our observations regarding the availability of computing resources to be used in a desktop grid.

Key words: *distributed computing, desktop grid computing, computational resource monitoring*

The Need for Desktop Grids

In many universities, research organizations, and, lately, enterprises, the following recent trends can be identified: larger amounts of data are being accumulated and manipulated; hardware performance of desktop computers increases dramatically; new technological advancements stimulate use of computing applications with extreme requirements for computational power; use of computing, simulations, visualizations, and optimization in various research fields and practical applications is accelerating and leads to very high demands on computing power; and the pace of development of high-performance servers hardly equals these trends, but for very high financial costs. Increasing hardware performance of desktop computers accounts for a low-cost high-performance computing potential that is waiting to be efficiently put in use [1, 6].

The computing power of a single desktop computer is insufficient for running complex algorithms, and, traditionally, large parallel supercomputers or dedicated clusters were used for this job [6, 10]. However, very high initial investments and maintenance costs limit the availability of such systems. A more convenient solution, which is becoming more and more popular, is based on the use of non-dedicated desktop PCs in a Desktop Grid Computing environment. Harnessing idle CPU cycles, storage space and other resources of networked computers to work together on a particularly computational intensive application provides for such a solution, while increasing power and communication bandwidth of desktop computers endow for this novel environment.

Distributed Computing harnesses the idle processing cycles of the available workstations on the network and makes them available for working on computationally intensive problems that would otherwise require a supercomputer or a dedicated cluster of computers to solve [13, 16]. A distributed computing application is divided to smaller computing tasks, which are then distributed to the workstations to process in parallel. Results are sent back to the server, where they are collected [7, 8]. The more PCs in a network, the more processors available to process applications in parallel, and the faster the results are returned. A network of a few thousand PCs can process applications that otherwise can be run only on fast and expensive supercomputers [17, 18]. This kind of computing can transform a local network of workstations into a virtual supercomputer.

In a desktop grid computing environment, the execution of an application is orchestrated by a central scheduler node, which distributes the tasks amongst the worker nodes and awaits their results [2]. An application only ends when all tasks have been completed. Exploiting desktop grids is further appealing due to the fact that costs are highly distributed: every volunteer supports her resources (hardware, power costs and internet connections) while the benefited entity provides management infrastructures, namely network bandwidth, servers and management services, receiving in exchange an enormous and otherwise unaffordable computing power. The typical and most appropriate application for desktop grid comprises independent tasks (no communication exists amongst tasks) with a high computation to communication ratio [3, 17, 18]. The usefulness of desktop grid computing is not limited to major high throughput public computing projects. Many institutions, ranging from academics to enterprises, hold vast number of desktop machines and could benefit from exploiting the idle cycles of their local machines [19, 20]. In fact, several studies confirm that CPU idleness in desktop machines averages 95% [5, 9].

Computational Resource Monitoring

Several desktop grid systems have been successfully used for many high throughput applications. Yet, in an organization or enterprise setup there has been little insight into the temporal structure of resource availability. We present here some of our observations regarding the availability of computing resources to be used in a desktop grid. The results are from an undergraduate student laboratory of 70 personal computers from the university [3].

It is well known that the highest availability of computers is during night, when few of the students are using lab computers. We are more concerned about resource availability during work hours, with the idea of using desktop grid resources for interactive tasks, like group visualizations or interactive presentations, where a group of researchers and students are working together.

We present our observations between 08:00 in the morning and 20:00 in the evening during several days, when computers from labs are actually used intensively (e.g. project deadlines, homeworks, lab hours). The results from the following plots show the actual number of computers available for computations from the running desktop grid (Figure 1 to Figure 6).

From a rough estimate, we can say that computers are available for computations about 50-60% of the time during weekdays, between 08:00 and 20:00. During the night, the availability is close to 95-100%. This amounts to approximately 75-80% availability of computers during a 24 hours interval of a working day, growing to 90-95% during weekends [3].

As a conclusion, we can say that, based on our available measurements, we can say that there is a lot of computing power available in such laboratories, which can easily be used for scientific experiments, provided that an appropriate resource-harvesting framework is available.

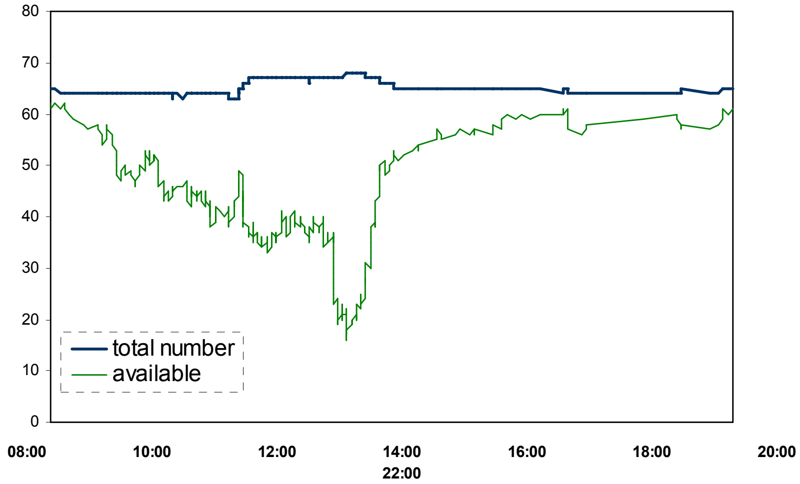

Fig. 1. Available desktop computers in laboratory (day 1)

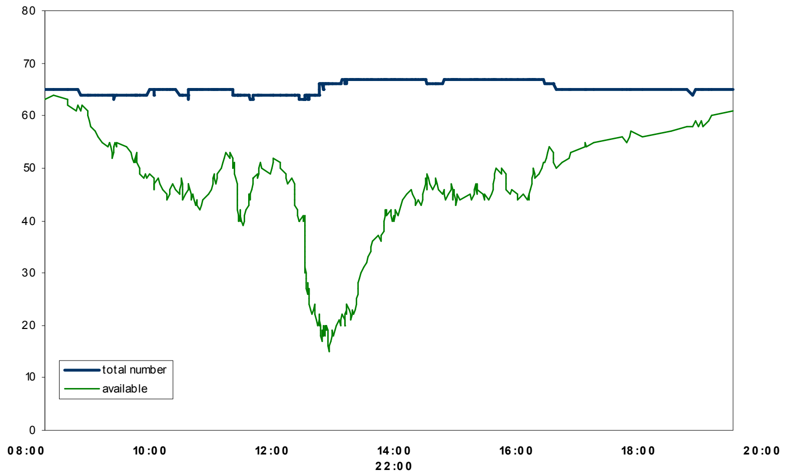

Fig. 2. Available desktop computers in laboratory (day 2)

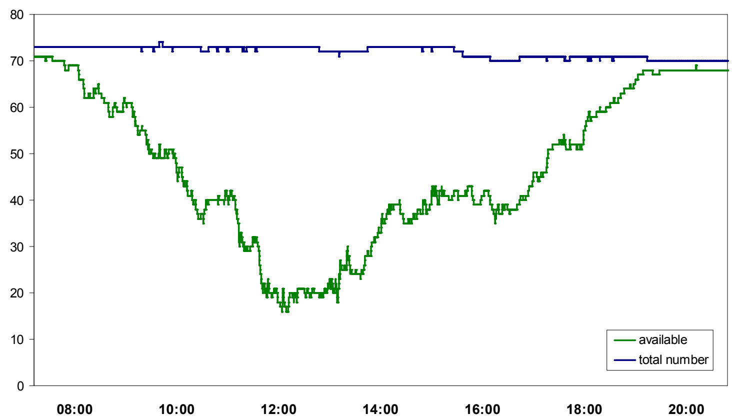

Fig. 3. Available desktop computers in laboratory (day 3)

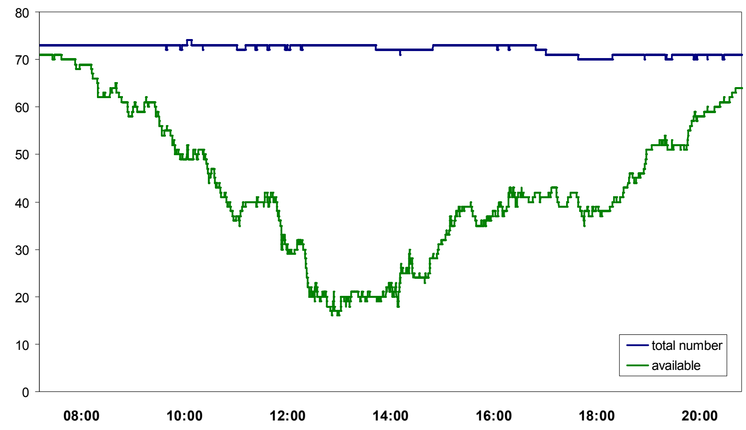

Fig. 4. Available desktop computers in laboratory (day 4)

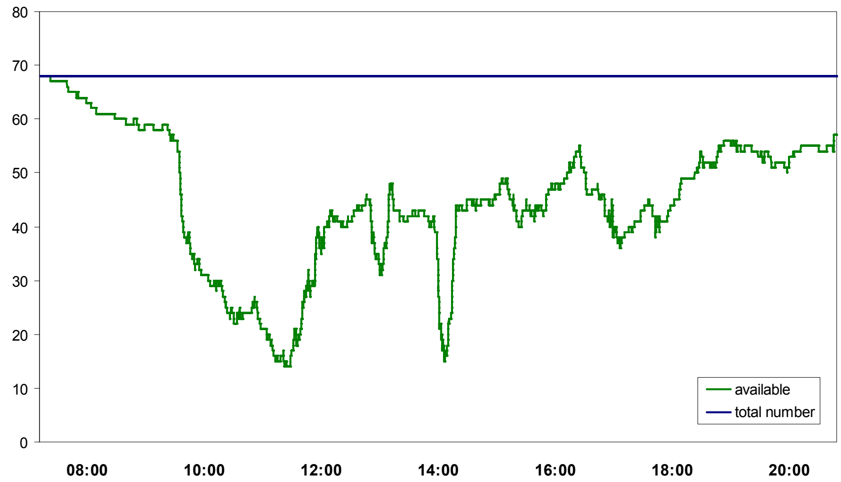

Fig. 5. Available desktop computers in laboratory (day 5)

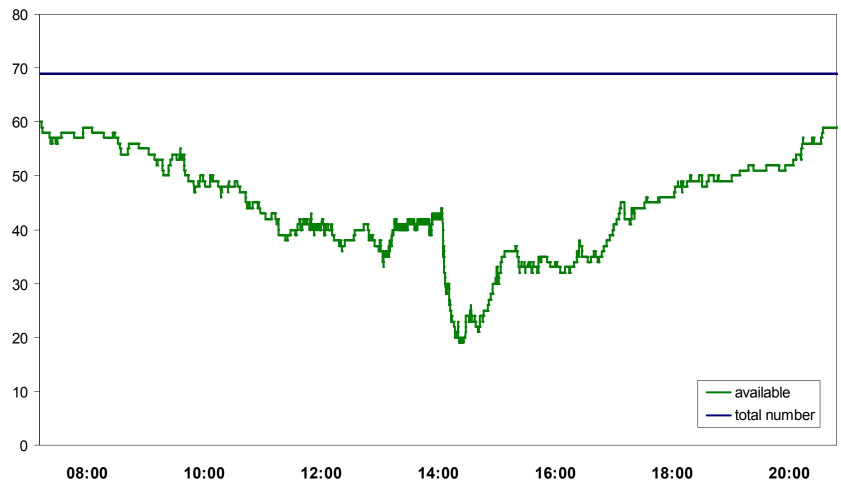

Fig. 6. Available desktop computers in laboratory (day 6)

Conclusion

This paper deals with processing power as the vital resource for any computer application. The motivation for harnessing the available processing power on the network is simple: to increase the size of problems that can be solved, to increase performance and obtain results faster [4]. Consider a typical local area network, where many low-price machines on the network will be idle for significant periods of time. If these wasted processor cycles could be utilized, they could represent a significant processing resource. This approach provides a more flexible and cost effective processing system. Normally, workstations may be in use as desktop machines, but become part of a distributed computation resource when they are not in use, for example, at night or during weekends.

It is worth to mention here that desktop grid and volunteer computing are not to evolve outside the Grid, but connected intimately with it, inside it [12, 14, 15]. Though there are some notable differences. Firstly, within the Grid, each organization can act as either producer or consumer of resources (hence the analogy with the electrical power grid, in which electric companies can buy and sell power to/from other companies, according to fluctuating demand). Secondly, the organizations are mutually accountable. If one organization misbehaves, the others can respond by suing them or refusing to share resources with them. This is different from volunteer computing or desktop grid computing in some sort of institutions, like universities, where it is practically impossible to track down each user of a resource at some point in time. On the other hand, desktop grid computing, which uses desktop PCs within a more formal organization, is superficially similar to volunteer computing, but because it has accountability and lacks anonymity, it is significantly different.

References

1. Berman, F., Fox, G., Hey, A.J.G. - *Grid computing: making the global infrastructure a reality*, New York, J. Wiley, 2003
2. Browne, J.C. et al. - General parallel computations on desktop grid and P2P systems, *Proceedings of the 7th workshop on Workshop on languages, compilers, and run-time support for scalable systems*, Houston, Texas, ACM Int'l Conference Proceeding Series, Vol. 81, 2004
3. Constantinescu Z. - *A Desktop Grid Computing Approach for Scientific Computing and Visualization*, PhD Thesis, Norwegian Univ. of Science and Technology, Trondheim, Norway, 2008
4. Cummings, M.P. - *Grid Computing*, http://serine.umiacs.umd.edu/research/ grid.php, accessed 2007
5. Domingues, P., Marques, P., Silva, L. - Resource usage of Windows computer laboratories, in MARQUES, P., Ed., *Int. Conf. on Parallel Processing Workshops* (ICPP 2005), Leiria, Portugal, 2005
6. Foster, I., Kesselman, C. - *The grid: blueprint for a new computing infrastructure*, Boston, Morgan Kaufmann Publishers, 2004
7. Garg, V.K. - *Principles of distributed systems*, Boston, Kluwer Academic Pub., 1996
8. Garg, V.K. - *Elements of distributed computing*, New York, Wiley-Interscience, 2002
9. Heap, D.G. - *Taurus - A Taxonomy of Actual Utilization of Real UNIX and Windows Servers*, IBM White Paper, 2003
10. Juhasz, Z., Kacsuk, P., Kranzlmuller, D. - *Distributed and Parallel Systems: Cluster and Grid Computing*, New York, Springer, 2004
11. Krauter, K., Buyya, R., Maheswaran, M. - A taxonomy and survey of grid resource management systems for distributed computing, in *Software – Practice and Experience*, Vol. 32, pp. 135–164, 2002
12. Kondo, D. et al. - Characterizing and evaluating desktop grids - An empirical study, in *Proc. of 18th IEEE/ACM International Parallel & Distributed Processing Symposium* (IPDPS2004), SanteFe, New Mexico, 2004
13. Leopold, C. - *Parallel and distributed computing: a survey of models, paradigms, and approaches*, New York, Wiley, 2001

14. Mustafee, N., Taylor, S.J.E. - Using a desktop grid to support simulation modelling, in TAYLOR, S.J.E., Ed., in *Proc. of 28th International Conference on Information Technology Interfaces* (ITI 2006), Dubrovnik, Croatia, 2006
15. Sarmenta, L.F.G. - *Volunteer computing*, Ph.D. thesis, MIT, Cambridge, USA, 2001
16. Zomaya, A.Y. - *Parallel and Distributed Computing Handbook*, New York, McGraw-Hill, 1996
17. * * * - *distributedcomputing.info*, http://distributedcomputing.info, accessed 2009
18. * * * - *distributed.net*, http://distributed.net/, accessed 2008
19. * * * - EGEE *Enabling Grids for E-Scienc*E http://www.eu-egee.org/, accesed 2008
20. * * * - *NorduGrid Grid Solution for Wide Area Computing and Data Handling* http://www.nordugrid.org/, accessed 2009

Disponibilitatea resurselor computaționale pentru sistemele desktop grid

Rezumat

Puterea de calcul oferită de un singur sistem de calcul de tip desktop nu este suficientă pentru execuția unor algoritmi complecşi. O soluție foarte potrivită este folosirea unor PC-uri desktop ne-dedicate în cadrul unui mediu de calcul de tip desktop grid. Aici costurile sunt puternic distribuite – astfel, fiecare voluntar contribuie cu resursele proprii (hardware, energie, conexiuni la Internet), în timp ce organizația beneficiară oferă infrastructura de management (lățime de bandă, server-e şi servicii de management), beneficiind în schimb de o putere de calcul uriaşă, care altfel nu ar putea fi obținută. Există puține experimente în structura temporală a disponibilității resurselor într-un mediu organizațional. Prezentăm aici observațiile noastre privind disponibilitatea resurselor de calcul care pot fi folosite într-un sistem de tip desktop grid.