Evaluating the Use of a Dynamic Cluster as an Educational and Research Infrastructure[1]

D. KEHAGIAS, M. GRIVAS, B. MAMALIS, and G. PANTZIOU
Department of Informatics
Technological Educational Institution (T.E.I.) of Athens
Ag. Spyridona St., 122 10 Egaleo, Athens
GREECE

www.cs.teiath.gr

Abstract: Clusters, built using commodity-off-the-shelf (COTS) hardware components and free, or commonly used, software, provide an inexpensive computing resource to educational institutions. In the Department of Informatics of T.E.I. of Athens, we have built a dynamic clustering system consisting of a Beowulf-class cluster and a NoW. The technical details of the system, as well as the experience gained in constructing the dynamic clustering system have been presented in a previous paper. The goal of constructing this system was to build an non-expensive computing resource and use it as an educational and research infrastructure. In this paper, we evaluate the use of the dynamic clustering system, as a platform for running the laboratory work of various courses (parallel computing, operating systems, distributed computing), as well various parallel applications in the framework of research, which is in progress under the programme "Archimedes".

Key-Words: Beowulf Clusters, NoW, Dynamic Cluster, Computer Science Course, Education, Research.

1 Introduction

Since NASA researchers used PCs [8] to implement a high performance system until today, the High Performance Computing (HPC) community witnessed a diversion from expensive, proprietary parallel systems to complexes of common computing devices, to a gigantic distributed system built on top of the Internet. Such platforms have also enabled smaller organizations build or participate in Clusters or multinational Grids [7]. Clusters share conventional technology, they are easier to understand and administer and they offer unrivalled availability and are largely expandable ([1], [9],[6]).

Education did not follow. Since Clusters and Networks of Workstations (NoWs) are made of common computers, there is a feeling that no new courses or material is required, as happened with transputers or vector computers. This lacks the consideration of idiosyncrasies, stemming from the complexity of such systems. PCs and workstations working together permanently or temporarily, can realize a costless and powerful computing system. Therefore, this paradigm of computing has to be presented to the students of higher education. On such a system, students can work and be trained on issues such as distributed computing, HPC, parallel programming and Grid computing.

In 2001, IEEE and ACM released a joint recommendation for undergraduate curricula [2]. Several courses have been included in their recommendation, related to high-performance computing, including:

– Parallel Algorithms
– Parallel Architectures
– High Performance Computing
– Net-Centric Computing
– Specialized Systems Development

In addition, the number of graduate programmes and research projects, which revolve around high performance computing topics -such as high-availability systems, parallel I/O, grand-challenge problems - has increased significantly during last years.

Our previous work on clusters and NoWs concluded to a dynamic, high-performance, versatile, multi-computer complex, that exhibited dynamically adapted performance, ensured cluster-level minimal performance and availability, as well as an interesting Grid resemblance. The complex was used primarily as an educational platform ([3]). Nevertheless, as a dynamic platform, it provides high-performance, that can be used for research purposes, too. Its potentials are valuable, since the minimum cluster ensures availability, but the dynamic, on-demand or when available expansion of performance over the NoW,

[1] This work is co-funded by 75% from E.U. and 25% from the Greek Government under the framework of the Education and Initial Vocational Training II, programmes "Archimedes" and "Supporting the Informatics Studies in the T.E.I. of Athens".

offers a significant power for several kind of scientific problems.

The use of the Dynamic Cluster as an educational platform for the needs of laboratory work on different courses in our Department, gives to the students the opportunity to gain hands-on experience on a not very common to them but useful platform. They are exposed to Operating Systems, networking, recourse allocation, hardware configuration, problem solving, and trouble shooting issues on a high-performance multi-processor environment. Through the use of the Dynamic cluster, we expect students to be able:

- To build a PC cluster.
- To understand and evaluate high-speed communication techniques and strategies, as well as, architectural design issues in HPC systems.
- To understand network issues for a HPC environment.
- To understand various HPC development tools and their appropriate usage in the HPC environment.

In this paper, we evaluate the dynamic clustering system, as a platform, which is used in the Lab to run various courses, (parallel programming, operating systems, distributed computing), and for running various parallel applications in the framework of research, which is in progress under the programme "Archimedes".

The following section describes the platform as it is used. Next, we evaluate the Dynamic Cluster as an educational and research infrastructure. In the sequel we comment on the future use and expansion of the platform. The conclusion summarizes the paper.

2 Configuration

The setup used in the original form of the dynamic cluster is depicted in Fig. 1. It includes a small Beowulf-class cluster and a NoW. The cluster consists of 8 PCs, employing Pentium 4 and 512 MB RAM, and a dedicated 100 Mbps Ethernet. No swap is used. The NoW configuration is based on the PC equipment of the Microcomputers laboratory, consisting of 18 PCs connected to a 10/100 Mbps Ethernet. Through a gateway, the laboratory LAN connects to the Institution (TEI) backbone and to the Internet.

The cluster PCs are dedicated to the parallel processing. On the other side, workstations, that are members of the NoW, may be used by students. Such use may consist of laboratory operation that can be light, moderate or heavily burdening of the processor nodes. In addition, students may overload the network (downloading large files) or the workstation (complex software, games etc). Last but not least, workstations may stack or reboot any time, without prior notice. Hence, NoW nodes cannot be considered totally available. Instead, they can assist dynamically to increase computing power, especially, when no workshop sessions are taking place.

The parallel processing environment used is the message passing interface (MPI) and specifically LAM, an open-source implementation of MPI. One of the cluster PCs plays the role of the central controller for both the cluster and the NoW. This PC carries two NIC cards, one for the cluster's LAN and one that connects to the laboratory LAN. As explained in [4], this configuration ensures high security, availability and dynamic high performance.

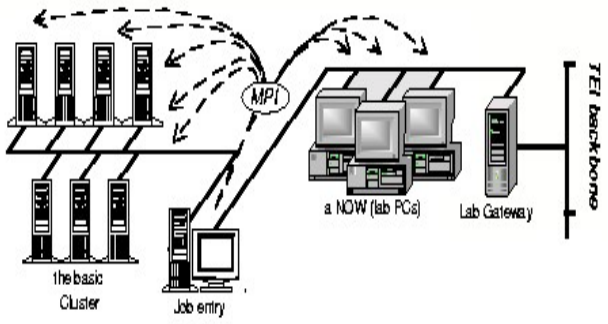

Fig.1. The dynamic cluster.

The cluster LAN is physically isolated from the rest; while only the central controller may communicate with the cluster nodes. Hence, their communication happens without any security measurements. Packets encoding or cryptography would impose an impediment to performance, introducing higher latencies and limited bandwidth, as well as wasting CPU cycles. Thus, there is no secure shell or SSL installed in cluster nodes. On the other side, security is of prime importance on the NoW, where all systems are exposed to Internet communications, that imply external and internal attacks. A multi-layered protection scheme in the Institution enforces protection and therefore, external attacks are hard to happen. However, students that work on the NoW nodes may, intentionally or unintentionally, expose the system to attacks, over their workstations. Therefore, security on this side has to be enforced. Using ssh has been proved a feasible and adequate solution that does not degrade severely nodes and network performance.

3 Benefits of the Dynamic Cluster in the curriculum

Three distinct groups from the academic community of the Technological Educational Institution (T.E.I.) of Athens can benefit directly from the Dynamic Cluster:

the students of the Department of Informatics, the faculty members and researchers of the Department, and researchers from other Departments of the Institution.

3.1 Students

The students of the Department of Informatics can benefit from the Dynamic cluster, by gaining hands-on experience working on and running various projects. These projects may be part of the educational process, mainly for the courses of "Parallel Computing" and "Operating Systems".

Having used the Dynamic Cluster as a teaching tool for the aforementioned courses [5], we end up with the following findings:

Parallel Programming

Students, through the use of MPI libraries for their programs, had the opportunity to apprehend the fundamental principles of parallel computation. Several workshop exercises demonstrated the differences between sequential code and parallel code execution. Those exercises assisted the students in gaining deep understanding of the way that a parallel algorithm offers better performance.

Scalability

Users' needs in computational resources increase continuously. Expansion of a system without the need of reassembling it from scratch, is crucial to most uses. The students, by using the Dynamic Cluster, did study and appreciate the concept of system scalability. Technically, scalability of resources took place in cases of severe workload increase, when NoW nodes participated to the whole system, over the Clusters and its server.

Upgrade ability

One of the most significant properties of NoWs is that each node works with its own, independent Operating System, on its own processor and using its own memory. Such elements should not be statically described from the beginning. Instead, they may adapt to newer needs or availability of equipment, such as a newer processor or more memory. With this property of NoWs, students had the opportunity to investigate and comprehend the notion of upgradability, to its smallest level, i.e. the improvement of a node.

Administration

The practical experience on LAN administration is only a part of the valuable experience of administering a complex, hybrid system. Its constituents have similar, but also particular properties. The students understand, that such a platform is a proper tool to learn the responsibilities and procedures of system administration.

General findings

- The students using the Dynamic Cluster may clarify issues, such as "doubling the number of processors does not mean doubling execution speed" [10].
- They can learn how to build and configure a Cluster without going deeply into the complexity of the software setup.
- The students need introductory courses on computer architecture and operating systems to benefit the most from a course on parallel computing.
- The use of the Dynamic Cluster offers to the students new abilities regarding HPC, which will improve their potential for professional excellence.

3.2 Faculty members

The faculty members benefit from the Dynamic Cluster, as they get more computing power for their research. While many "large" academic institutions have a supercomputer available for faculty research, "smaller" institutions do not have the resources for such a system. The Dynamic Cluster provides a research platform for many faculty members, from a variety of disciplines, allowing them to conduct research requiring intensive computational power.

Current use of the Dynamic Cluster by faculty members takes place in three directions:

Reducing overhead

As the bandwidth gap between Ethernet and non-IP SANS continues to decline, commodity Ethernet/IP networks offer a cost-effective and standards-based alternative for demanding applications. A frequent cited drawback of IP networks is that data copying and TCP/IP protocol overhead can consume a significant share of host CPU cycles and memory system bandwidth, siphoning off end-system recourses needed for application processing. This overhead can increase latency and limit application throughput relative to a non-IP SAN with equivalent wire speed.

With respect to the Dynamic Cluster, our team is interested in reducing overhead for efficient communication. Towards this goal, our research is organized into two directions. The first direction is towards removing some context-switching and copying of data by removing part of the kernel from

the communication critical path. The second direction investigates the possibility of using new I/O approaches within the kernel to remove all unnecessary data copies. These two approaches not only improve the network communication overhead, but also the I/O within the kernel.

Scheduling algorithms and prototypes
Within the frame of the European funded research programme "Archimedes", the "Parallel and Distributed Processing" research team designs, analyses and simulates on-line and off-line scheduling algorithms, especially for tree-structured networks.

The Dynamic Cluster is used as a real, practical platform for tests, evaluations and performance measurements and comparisons.

Parallel text retrieval
Again, in the context of the research programme "Archimedes", the "Information Retrieval" research group, has built an efficient parallel text retrieval system aiming at the achievement of high speed-up values and response times over PC-Cluster environments. The indexing scheme of this system is based on classical and well studied IR automatic indexing features (statistical extraction of word-stems, phrases and thesaurus classes) according to the Vector Space Model (VSM – representing each document by an appropriately weighted document-vector), whereas, the parallel approach aims at the efficient parallelization of the basic VSM-implied retrieval tasks: scoring (computing the similarity between each document vector and the user query-vector via the well known cosine similarity function) and ranking (extraction and presentation to the user of the top-score document-vectors). Furthermore, they try to efficiently handle the two main problems (with regard to parallelization over PC-Clusters) met in this case: (a) even distribution of the computational load (with regard to local scoring and local ranking sub-tasks) over the multiple processors (by applying a well designed 'data sharing and balancing scheme' based on document clustering) and (b) efficient parallelization of the necessary global ranking sub-task required at the end of parallel processing (by applying well designed tree-based merge-sort algorithms). They also incorporate an efficiently implemented 'clustering-based search mechanism' (via the extraction of cluster/superclusters and corresponding centroid/supercentroid vectors) in order to gain high response times over very large document collections.

The above techniques have been appropriately implemented and evaluated with the use of MPI on the Dynamic Cluster.

3.3 The Future
Following numerous conversations and exchanging of ideas with colleagues of other Departments of the Institution, it comes out that the research community is interested in using HPC clusters as a tool for their research. In the light of that fact, we intend to install the Dynamic Cluster as the basis for HPC, that several research groups from other Departments will be able to use.

Our further objective, though, is to use the Dynamic Cluster as a compass and forerunner for the implementation of a campus Grid. That Grid will enable the collaboration of multiple departmental labs and centers within the Institution. A campus Grid is an important research tool, which brings together computational scientists, students, and researchers across the campus, over and within a common, shared computer platform. Its implementation is easier than that of a larger multi-institutional Grid since there are common infrastructure services, such as high speed networking and naming services (DNS).

The Department of Informatics soon will run a graduate program where the Dynamic Cluster will be the appropriate tool in order to run workshops of courses dealing with topics such as: advanced communication and networking in a HPC environment, advanced hardware architecture for a HPC system, and grid technologies and distributed computing architectures.

4 Conclusions
Cluster computing is an emerging answer to several technological dilemmas. The Dynamic Cluster –an architecture that combines NoWs and clusters- can support HPC in a dynamic way, where any available computer can participate as an additional node. Its versatility and flexibility with respect to configuration and functionality, together with its dynamic, strong computational power, renders it a very helpful tool for educational and research purposes.

The platform has already become a useful tool for several courses. The results of its use by the students, especially in the context of the course of parallel computing, are very promising. It comes out that the students, by exploring and working with the Dynamic Cluster, can apply the parallel programming techniques and build a stronger knowledge background beyond the basic course material being taught. Also, its use by faculty members renders it a very helpful tool for research purposes.

References

[1] C. Ernemann, V. Hamscher, U. Schwiegelshohn, A. Streit, and R. Yahyapour, "On advantages of grid computing for parallel job scheduling," in *2nd IEEE International Symposium on Cluster Computing and the Grid (CC-GRID 2002)*, 2002.

[2] CARL CHANG ET.AL. Computing curricula 2001; computer science, final report. Curricular Standards, available for download at http://www.sigcse.org/cc2001/, 2001.

[3] D. Kehagias, M. Grivas, G. Meletiou, G. Pantziou, B. Sakellarios, D. Sterpis, and D. Ximerakis, " low-cost dynamic clustering system for education and research", in *Proceedings of TEMPUS Workshop*, Jan. 2004.

[4] D. Kehagias, M. Grivas, G. Meletiou, G. Pantziou, B. Sakellarios, D. Sterpis, and D. Ximerakis, "Building a low-cost high-performance dynamic clustering system", in *6th International Conference on Telecommunications in Modern Satelite, Cable and Broadcasting Services (TELSIKS 2003)* (B. D. Milovanovic, ed.), vol. 2, pp. 608 613, IEEE and Faculty of Electronic Engineering, Nis, Serbia, October 1-3 2003.

[5] D. Kehagias, M. Grivas, G. Pantziou, "Using a Hybrid platform for Cluster, NoW and GRID computing" to be published in the International journal of FACTA UNIVERSITATIS, 2005.

[6] I. Baird, "Understanding GRID computing," *DAILY NEWS AND INFORMATION FOR THE GLOBAL GRID COMMUNITY*, vol. 1, July, 1 2002.

[7] I. Foster and C. Kesselman, *The Grid: Blueprint for a New Computing Infrastructure*, ch. 2. Computational Grids. Morgan-Kaufman, 1999.

[8] J. A. Kaplan and M. L. Nelson, "A comparison of queueing, cluster and distributed computing systems," Tech. Rep. TM 109025 (Revision 1), NASA Langley Research Center, June 1994.

[9] M. Baker, R. Buyya, and D. Laforenza, "Grids and grid technologies for widearea distributed computing," *International Journal of Software: Practice and Experience (SPE)*, vol. 32, pp. 1437–1466, December 2002.

[10] R. Benner, J. Gustafson, and G. Montry, "Development and analysis of scientific application programs on a 1024-processor hypercube", Tech. Rep. SAND 88-0317, Sandia National Laboratories, Feb. 1988.