

Integrating IS Curriculum Knowledge through a Cluster-Computing Project – A Successful Experiment

Fred L. Kitchens, Sushil K. Sharma, and Thomas Harris
Ball State University, Muncie, IN, USA

fkitchens@bsu.edu ssharma@bsu.edu tharris@bsu.edu

Executive Summary

MIS curricula in business schools are challenged to provide MIS courses that give students a strong practical understanding of the basic technologies, while also providing enough hands-on experience to solve real life problems. As an experimental capstone MIS course, the authors developed a cluster-computing project to expose business students to the technologies of distributed computing, parallel computing, networking, securing networked environments and designing an application for distributed computing. Building such a cluster requires the integration and application of concepts including operating systems, networking, parallel computing, and computer architecture, making it an excellent subject for a capstone project. This paper describes the author's experience in developing a cluster-computing course, the use of a cluster-computing project to integrate the knowledge of various MIS courses and, a blueprint of the authors' recommendations to other faculty interested in developing a cluster-computing course. Building a Beowulf cluster was a rewarding project for a capstone course in which students learned many MIS and project management concepts. Based on the authors' experiences, building a Beowulf cluster as a senior project for a capstone course is highly recommended. Students reported that this course enhanced their understanding of operating systems, networking, security, and parallel computing concepts. The purpose of this paper is to share the experiences of cluster computer implementation in a business curriculum and how cluster computing capstone course could be used for integrating MIS concepts and provide strong hands on experience. The theoretical and technical challenges make it an absorbing and rewarding capstone project for an undergraduate education.

Key words: Cluster Computing, MIS curriculum design, Beowulf, course development, high performance computing

Introduction

Business students majoring in areas such as MIS, IS, and IT in business colleges are arriving at college with ever-increasing technical proficiencies. Many businesses, as a result of changing needs, have begun to appreciate the value in employees who have a combination of business and technological skills. Businesses finally realize that hiring individuals, who possess strong technology skills and come from business programs, is not only desirable but reduce their up-

Material published as part of this journal, either on-line or in print, is copyrighted by the publisher of the Journal of Information Technology Education. Permission to make digital or paper copy of part or all of these works for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage AND that copies 1) bear this notice in full and 2) give the full citation on the first page. It is permissible to abstract these works so long as credit is given. To copy in all other cases or to republish or to post on a server or to redistribute to lists requires specific permission and payment of a fee. Contact Editor@JITE.org to request redistribution permission.

front training costs. "Value-added" is a popular term in businesses. How do the MIS courses add value to the ability and knowledge of the MIS graduates? Many programs are devoted to explaining the concepts of building networks and managing IT infrastructures through case analysis. Broad-based lecture-oriented case-supported education cannot prepare students to handle the tasks and responsibilities that they will be expected to take-on (Wrege, 1982). Of course, it prepares students to understand and simulate various business situations in which they may have to work (Korwin & Jones, 1990). However, organizations demand that MIS graduates coming out of business schools not only have the ability to understand and analyze different business situations, but also are proficient in handling and managing technology tools and environments. Organizations spend millions of dollars to train the MIS graduates and to customize to their role and responsibilities. Practitioners frequently claim that MIS curriculums do not prepare students to handle the roles and responsibilities that graduates have to take-on in organizations. As a result, companies have to bear the high cost of training them. Graduates entering the job market today will find employers seeking recruits well versed in theory, but also with networking experiences in a hands-on environment. Many institutions are trying to adjust their course offerings to address this situation (Harvey, Sirna, & Houlihan, 1998; Lane, 1981). From an educational perspective, teaching cluster computing is an ideal way to expose students to theoretical concepts as a foundation to the learning experience, and then providing them with the opportunity to apply what they have learned in a hands-on environment (Boud, 1995; Scanlon, Tounclu, & Jones, 1998). Often, MIS courses are taught from a theoretical point of view with little emphasis on actually using equipment.

The use of cluster computing in business requires administrators and analysts with knowledge, skills, and experience in networking to effectively build and operate a system. As cluster computing migrates to an increasing number of business applications, personnel with a unique knowledge of both the technical aspects of cluster computing and the business acumen will become necessary. Thus, there is an increasing need to teach cluster computing in colleges of business (Hyde, 2000). The information curriculum being taught in colleges of business - by whatever name, MIS, IS, IT, etc. - should take a close look at cluster-computers. Such a course could help to provide cluster-administrators to the corporate world, and prepare future employees for the tasks ahead of them (Rice, Wilson, & Bagley, 2001).

Cluster Computing

Cluster computers are increasingly important elements in mainstream computing. In recent years, cluster computers have emerged as the leaders in high-performance computing (Sterling, 2000). Today's cluster-computers are more than a match for mainframes in terms of both cost and performance (Baker, Fox, & Yau, 1996). Cluster computing harnesses the combined computing power of multiple microprocessors in a parallel configuration. Cluster computers are a set of commodity PC's dedicated to a network designed to capture their cumulative processing power for running parallel-processing applications. Clustered computers are specifically designed to take large programs and sets of data and subdivide them into component parts, thereby allowing the individual nodes of the cluster to process their own individual "chunks" of the program (Baker, Apon, Buyya, & Jin, 2002; Baker & Buyya, 1999).

There are many cluster configurations, but a simple architecture such as the one shown in Figure 1, is used to visualize the basic concept (Baker et al., 2002). In a typical cluster, the application is run on a Master node. However, the computational work is split-up and parsed out to be done by the multiple nodes in the cluster. This way, the cluster is better equipped to handle larger amounts of data and complex problems than otherwise possible on a stand-alone machine. Cluster-computers are a realistic alternative for a variety of applications similar to that of the more expensive traditional supercomputers (Baker, Fox & Yau, 1996). Many educational institutions and

companies are beginning to explore this option as a cost-effective alternative for supercomputing applications, and retailers are looking to sell them these systems (Read, 2002).

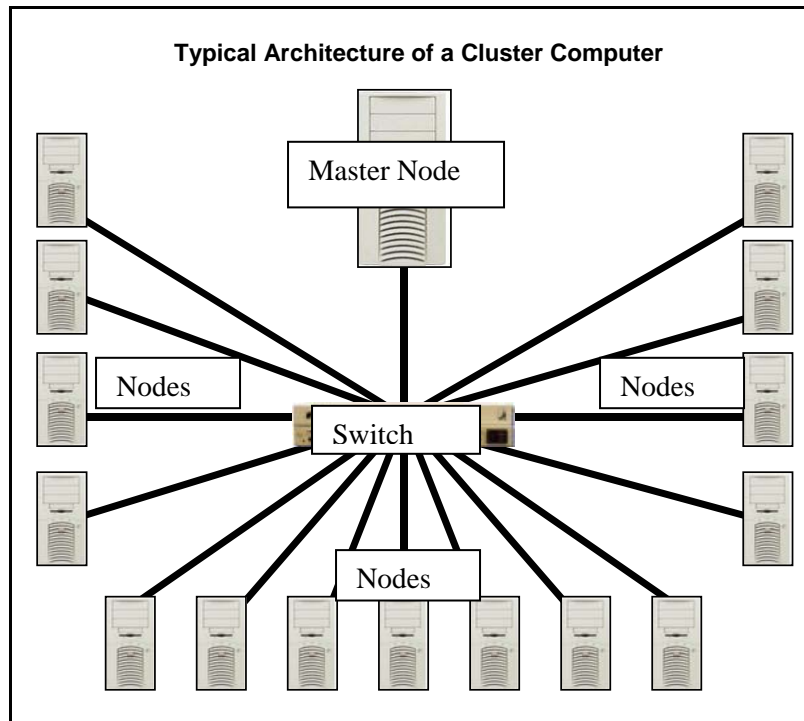


Figure 1: Typical Architecture of a Cluster Computer

Beowulf clusters

Supercomputers were once a resource limited to major universities and government agencies with huge funds at their disposal. But, the availability of low-cost commodity hardware, free operating systems such as Linux, and faster, lower cost networking have made it possible for institutions at all levels to provide supercomputing resources to their researchers. Beowulf is perhaps the most well-known type of parallel processing cluster today (Sterling, Salmon, Becker & Savarese, 1999). The Beowulf architecture was first introduced by scientists at NASA who were trying to build a system that would provide increased computational power using inexpensive off-the-shelf computer technology. Beowulf class cluster computers range from several nodes to several thousand nodes and are being applied in many areas. Unlike Internet server clustering, which typically distributes multiple requests/tasks to different nodes (load-sharing), the Beowulf environment directs the processing power of all nodes to a single task (versus distributing multiple tasks to different nodes). This makes Beowulf a good choice for processing-intensive, analytical applications, such as mathematical computation, scientific analysis (weather forecasting, seismic analysis, etc.) and financial data analysis. The processing power of Beowulf clusters has been applied to various complex computation problems such as mapping the human genome, weather forecasting, cryptanalysis, fractal simulation, simulation of nuclear explosions, and more (Hoffman, Hargrove and Schultz, 1999).

Evidence of cluster computing courses is found in the areas of computer science and engineering, generally at the graduate level (Apon, Buyya, Jin and Mache, 2001). As cluster computing becomes more popular in business, the need for college graduates with both business acumen and cluster computing knowledge increases. One way to meet this emerging need is to offer cluster-

computing courses within the MIS curriculum in business programs. Several universities have taken initiative and started courses and projects in cluster computing, but until the authors' attempt, it had not been offered in business colleges, nor had it been offered at the undergraduate level (Apon, Buyya, Jin and Mache, 2001; Wilkinson and Michael, 1999). The authors have developed a cutting-edge educational endeavor for undergraduate students in Information Systems. They are developing and refining a business-oriented cluster-computing course. It is anticipated that this work will serve as a model for other academic institutions to incorporate cluster-computing courses into their business curriculum. Such large scale projects are generally out of the realm of a typical college undergraduate experience. Some schools have implemented small clusters with success, targeted to teach an entire course on parallel computing. The author's goal in the cluster computing project was to integrate the IS concepts that students had received through various courses through a capstone project-based course. Students were asked to build a small scale cluster, to learn both implementation and application levels of clustered networking, and to create a reference for undergraduate students interested in this topic. A secondary motivation was to develop modular course content and reference materials which could be incorporated into an undergraduate curriculum in courses such as computer architecture, operating systems, network management, and security courses (Childers, Kitchens and Sharma, 2003a, 2003b).

The paper is divided into three main sections. The first section describes the emerging need for cluster computing environments and why it would be beneficial to have a cluster-computing project in business schools. The second section describes how Beowulf Cluster-Computing Project experiment was conducted to integrate the knowledge of various MIS courses. The third and concluding section describes the author's experience in developing a cluster-computing course, and lay out a proposed blueprint as to the author's recommendations to other faculty interested in developing a cluster computing course.

The Need for Cluster Computing in the Curriculum

In recent years, it has become clear that the future of the high performance computing industry is in cluster computing. Cluster computers are found in the Top 500 List of the highest performance computers in the world. Cluster computer systems can be found anywhere from undergraduate dormitories to national laboratories and everywhere in between (Sterling, Salmon, Becker and Savarese, 1999). With this in mind, universities need to begin including clustering topics in their curriculum (Hyde, 2000).

Beneficiaries of a cluster in a business curriculum

Three segments from the academic and business community can benefit directly from a cluster-computing project: students, faculty, and local businesses.

Students. University students in the colleges of business can benefit by gaining hands-on experience working on and running a technology-based project. Students learn problem-solving skills as they build systems out of miss-matched and old equipment (Apon, 2000). They gain communication skills by making professional presentations to local organizations in attempt to acquire donations of equipment, and by interacting with industry experts requesting information and advice. They also gain organizational and managerial skills by working on a real-life on-going project. These skills will serve them well in their future careers.

Faculty. Faculty members benefit from cluster project as they get increased computing power of a cluster-computer system for their research. Where many large universities have a supercomputer available for faculty research, smaller schools do not have the resources for such a system. A cluster computer can provide a research platform for many faculty members, from a variety of

disciplines including business and non-business applications, allowing them to conduct research requiring intensive computational power.

Local Businesses. Businesses will also benefit from the development of a cluster computer in a business school (Hyde, 2000). As the project evolves and business application software is developed (possibly in conjunction with a computer science department) for such systems, the project will be in a position to provide instruction and direction for businesses to build a similar system using their own old computers. There is also the possibility of faculty consulting opportunities using the power of the cluster for research in data mining, marketing, forecasting, financial analysis, and other applications.

The following section of this paper provides a brief description of the authors' experience, followed by a semester-by-semester blueprint on developing a cluster-computing course for an MIS curriculum based on the authors' experiences and hindsight.

Beowulf Cluster Computing Project – A Successful Experiment

Cluster computing has emerged as an opportunity for businesses and academic institutions to garner supercomputing-processing power in their respective practices (Baker, Fox & Yau, 1996). The competitive nature of the business world today fuels the need for increasing levels of computing power. While additional computing power is needed, it can be expensive. Cluster computing can provide extensive processing power to satisfy many business applications, and can do so at far less expense than a traditional supercomputer (Sterling, 2000). Effectively, it offers users two primary benefits. First, it enables supercomputer-processing power to be developed at a fraction of the cost of a traditional supercomputer (Los Alamos News Letter, 2002). Businesses and IT professionals today are faced with difficult challenges when designing systems. Users and applications now have high expectations for the availability, scalability, and performance of computing platforms. Building a cluster computer is an alternative to purchasing an expensive supercomputer. Second, cluster computing is an effective way to make use of older computer hardware. Computer obsolescence rates are growing at a feverish pace (Rochester Computer Recycling, 2002). At the rate businesses and universities update and replace their technological infrastructure, they face a dilemma: what to do with out-dated equipment. Building a cluster computer makes use of old equipment and offers the opportunity to harness greater computing power. A recent surge in the demand for electronic equipment in the last decade has driven manufacturers to develop new and improved products at a feverish pace. In the United States alone, more than half of all households own a computer. Businesses are all but forced to upgrade equipment every few years just to remain competitive technologically. At the same time, product innovation has led to the rapid increase in obsolescence. The useful life of computers and monitors decreases with each successive generation. Experts estimate that by the year 2005 one computer will become obsolete for every new one released on the market (National Safety Council, 1999). As new and improved products enter the market, businesses and households alike face a dilemma deciding what to do with their old computers.

Breaking from similar projects at other institutions where cluster computing systems have been built by purchasing brand new equipment in attempt to construct the most powerful system possible, the Cluster Computing project at Ball State University is aimed at building a system with the highest performance possible for the least investment possible. Nearly all equipment used in the project was acquired from the university's excess inventory or donated from local businesses and private citizens from the surrounding community.

The Cluster Computing Research Project at the authors' institution began as a project with little background or direction as far as previous course development projects to model from other aca-

demographic institutions. Other universities have built cluster computers, but surprisingly there exists little documentation on “how to assemble a cluster,” especially when using mismatched and used equipment. What literature is available tends to be too focused for general application, or too general for hands-on work. Project members essentially had to start their research with only bits and pieces of information from a variety of sources. This required setting goals and objectives, and establishing a fundamental methodology. The project has become successful, as the group learned a great deal through trial-and-error. While frustrating at times, this has been a great learning experience for the students involved (and for the professors as well). The cluster-computing project began in the fall of 2001 as a group project in a “Systems Analysis and Design” course in Management Information Systems. As a team project, undergraduate students broke from the traditional business-application project and proposed to build a cluster computer. Upon the professor’s approval, the group performed feasibility and needed analyses, which included determining equipment, space, and resource needs for the proposed project. The result of the project at the semester’s end was a working Beowulf-type cluster computer. Following the conclusion of the initial semester, the project took on a life of its own and continued the following semester as an experimental course. As the project evolved, the group established three primary objectives:

- 1) Establish a teaching methodology for cluster computing in an MIS curriculum,
- 2) Show how mid-size businesses can take advantage of their old computers, and
- 3) Provide the university with a high-performance computing system to support faculty research

Cluster computing has been taught as a formal class in a few instances either as a computer science class or engineering class – at an honors or graduate level (Apon, Buyya, Jin and Mache, 2001). As part of the experimental class, the undergraduate students took part in what is believed to be the first instance of cluster computing being formally taught in a business (MIS) curriculum. With the second semester of the project, interest in the system and course quickly developed. The class launched a public relations campaign, seeking to inform the local community about their work, as well as gain contributions for the project. In its third semester, the course was offered as an elective course in MIS. A curriculum was established which would provide hands-on experience in hardware, Linux operating system, professional presentations, problem solving, technical report writing, and research methods.

Student Activity and Responsibility

The students were given responsibility for all phases of the project. The students met twice a week for lab time to work together on the project, and scheduled weekly individual research time. The instructor met with the students twice in a week. A work log was maintained and students were asked to submit bi-weekly progress reports.

The students were responsible for the physical implementation of the project, building the Beowulf cluster and installing Linux. They installed and configured parallel processing software. The next phase involved monitoring the performance of the cluster with a viewable application. This assured the success of the project and allowed students to trouble shoot and increase performance. Finally, the last phase after successful implementation involved documentation. The students were responsible for creating a reference manual, or a “how-to” geared toward the undergraduate level.

A lab manual for use in future classes involving parallel computing was planned. Finally, students made a number of presentations to various civic organizations, sharing their cluster computing knowledge. Students also made presentations at conferences, seminars, university committees, and alumni board meetings.

Faculty Activity and Responsibility

The faculty member as an advisor was responsible for directing the project. The faculty advisor arranged access to available hardware, computing resources, and facility space; and was responsible for monitoring the progress of the project. The faculty advisor not only conducted classes for imparting basic training of how to build the cluster in learning modules, but also provided technical support for the hardware and software. The authors found that the greatest learning was achieved when they provided more ‘resources and direction,’ and less ‘how-to instruction’.

Classroom Learning Experience

Instructors’ perspective towards capstone cluster computing course has been that Beowulf clusters proved to be a very valuable classroom resource for teaching computer architecture, networking, operating system, security and parallel programming concepts. A parallel computing architecture provided students with an environment in which they understood more effectively network operating systems, parallel programming techniques, and distributed computing. A few examples of projects and learning opportunities that were created through the cluster computing project are listed here.

Parallel Computing

Using a parallel algorithm to implement a solution to a software problem can be both new and interesting. Laboratory exercises for cluster computers provided students the challenge of comparing the execution of serial code to the execution of parallelized code. This project was very helpful to help students understand the basics of parallel programming.

Students performed a few exercises to understand how a parallelized algorithm offers increased speed. They also discovered how changes to the algorithm speed up or slow down execution time. A cluster illustrates important concepts such as "doubling the number of processors does not mean doubling execution speed".

Building a Cluster

After providing basic instruction about cluster computing, students were asked to build a cluster on their own. With a little hand-holding, students learned the issues involved in constructing clusters without having to delve too deeply into the complexities of software setup and configuration.

System Administration

After the cluster is constructed, it requires an effective system administration to remain useful. Students found the cluster environment a very valuable tool for learning system administration duties and responsibilities. Maintenance and administration of a cluster are similar to those of a LAN, and gave students valuable experience. Four major domains of work were explored in this area: hardware management, software platform maintenance, automation, and security.

Hardware Management

Students were asked not only to take care of system administration but also to manage hardware. The jobs that students handled were: hardware acquisition and replacement; space, heat, and power management; and most importantly, network management.

Dealing with this process helped future system administrators develop system design skills. Students also managed the location of shelving, machines, wires, and peripherals, all within the constraints of the room available. Because of space constraints, cluster nodes were packed tightly

together and students even went as far as developing a plan for proper cooling of all of the machines.

Networking is a unique problem with Beowulf clusters. It can be divided into two major, related areas: cabling and topology. Cabling is a problem because the large number of wires and their length constrains the space in which the cluster can be placed and limits signal quality. Topology is a consideration because of its effect on performance. Each computer will use a power, mouse, keyboard, video and potentially multiple network cables. It is difficult to service a cluster when there is a cascading mess of wires covering everything on the cluster's backside. The cables quickly become unmanageable if a coherent labeling and bundling plan is not thought of ahead of time. Another factor is cable cost. It is cheap and educational for students to create custom cables from a box of cable parts and crimping tools.

Software Platform Maintenance

Students received good experiences in basic configuration of servers in system administration classes, but a Beowulf cluster illustrated the demands of managing a LAN. Software configurations change over time by upgrading software packages, introducing new ones, removing old ones, and tweaking the existing packages. Every machine in a cluster must be able to work with the other machines. Maintaining the software on a cluster consists of administrative work multiplied by 'n' nodes - each of which is potentially dependent on other nodes. Students learned the skills necessary to identify the needs of the users of the cluster (and the needs of the cluster itself) and successfully fulfilled those needs.

Security

A cluster connected to the Internet or a campus LAN presents a good opportunity for students to study and implement computer security. Networked computers are vulnerable by virtue of being connected, and a cluster may attract more than its fair share of curious joy-riders and system crackers looking for some serious computing power. Students created, implemented, and maintained a security plan. The implementation of the security plan proved to be both complex and time-consuming but a great learning experience for them. Aside from understanding the security features of various services, the students learned a great deal in the security rule setup, integrity checking, and parsing of log files etc. Typical exercises included the following: examining the log files of each node in the cluster (especially nodes connected to the Internet), using an integrity checking system such as tripwire to see if any configurations have been changed without authorization, searching for strange files (e.g. SUID/GUID executables), investigating any unusual activity, checking BUGTRAQ or CERT advisories for new exploits found in services being run on the cluster and downloading patches that correct problems. These exercises made students more confident in handling security and system administration on a live project.

In summary, a Beowulf cluster project provides students and faculty both an opportunity in the administration and growth of a production system, opportunities for learning in the theory and practice of parallel programming, and opportunities in major research and development. Lessons learned in building and fine tuning the authors' Beowulf cluster, and later some research projects conducted using this system, led to rich experiences for both faculty and students. A number of courses were positively impacted as the direct result of the Beowulf cluster. These courses include operating systems, computer networks, database management systems, algorithms design and analysis, and security.

Academic/Pedagogical Value of Clusters

To be effective as an MIS specialist, students need the "learn by doing" approach rather than trying to grasp knowledge through various simulated case studies and exercises. A Beowulf super-

computing cluster is an invaluable educational resource for teaching various concepts of information systems program, application development and research.

Students in undergraduate business schools studying information systems are not computer scientists or engineers by training, and will require help in understanding complex subjects like parallel programming, and distributed environments. Conversely, computer science students generally lack good project management skills. Teaching a Capstone course through cluster building is a very exciting opportunity for both business school students and computer science students to learn these domains of knowledge and apply that knowledge to create usable applications. Cluster computer project had few students from computer science program. This helped to established relationships between departments for mutually beneficial collaborations.

The Beowulf Cluster-Computing Project was built, maintained and managed by students; this exercise integrated various concepts of MIS that are taught in different courses (Childers, 2003a). The goal was to provide students with a practical understanding of the technologies that will support them in their careers. The educational aspect of the course emphasizes the practical application of various IS theories, concepts and technical tools to a real life information systems problem. Students gain an appreciation of the complexity of challenges associated with the management of information technology through group interaction, comprehensive projects, research and active participation in the building of a cluster-computer. The student's exposure to handling the 'nuts and bolts' and building a cluster-computer is a key element in each student's development and understanding of the integration of various aspects of information technology. These skills will be needed for developing and managing IT applications in organizations. The course's unique philosophy, integrative methods, and "learning by doing" are the basis for the professional achievements of its graduates. This approach gives students the unique opportunity to learn the basic computer components that make an IT infrastructure in terms of servers, nodes, network and parallel distributed environment for running and managing multi user application in organizations. By effectively combining various areas of study, students become technically proficient and able to address issues such as building a network, securing it, running multi task applications, and the effective management of information technology environments.

A Blueprint for Developing a Cluster Computer

Given the experience of developing this cluster-computing course along with the advantage of hindsight, the authors propose a blueprint for developing a cluster computer to be used for MIS instruction and faculty research. To develop a working cluster requires time - time to gather equipment donations, support, and to develop a working hands-on knowledge of cluster computing as the project evolves. Rather than suggest that a faculty member gain a working hands-on knowledge before attempting to teach cluster computing for the first time, the authors suggest a developmental course of action. By following a developmental blueprint, faculty can begin by teaching project management or systems analysis and design at first – slowly converting course curriculum to clustering topics as the project and the cluster are developed over several semesters.

The authors' recommendation is to plan on spending two semesters developing a stable cluster computer as a student team project. Then, use the developed cluster to teach cluster-computing beginning in the third semester. This gives students experience in project management and system analysis & design, while allowing them to develop a sense of ownership in the project. This plan also removes the pressure from the faculty members who might otherwise feel compelled to quickly learn the material and develop a cluster on their own time, before attempting to teach a course.

Semester One

The authors' proposed method for launching a cluster computer project is to let the project begin in the hands of a group of students. This can be as part of a class project or as an independent study with a faculty member. Whatever the case, student interest and participation is the driving force behind the success of such an endeavor. Following this approach, it is initially important to have more technologically advanced students on the project to help foster its development and overcome technical difficulties. For this purpose, students with extensive technology backgrounds or technology-related minors are recommended. The primary goal in the opening semester of the project is to simply get a working system up and running. Students in semester-one benefit by receiving hands-on experience on an IT-project. They are exposed to operating systems, networking, resource allocation, hardware configuration, problem solving, and trouble shooting in a real-life environment. Aside from the technological benefits students gain in this semester, they also gain experience working with systems analysis & design methods, proposal development and project management. This is an opportunity seldom available to undergraduate students (Baker, 2000). Semester-one of the project requires certain resources to be gathered. Resource requirements for this semester include:

- Students with interest in the project
- A room or lab to house the project
- A 100 Mb/second switch (A hub may be initially substituted until the cluster grows to approximately 4-6 clusters in size.)
- Personal Computers as described below (initially, these may come from excess inventory or from computing labs as they are updated with newer equipment)

The authors' experience is that due to the age of many of the computers, component parts from two to three PCs may be required to get a node up to acceptable minimum standards. Although it is possible to use a node with less hardware, the authors' current minimum requirements are:

- Pentium-based processor (any Mhz as long as it is Pentium). Other manufacturers such as Sun or Apple may be used, but are less frequently available on the author's campus.
- 64 MB of RAM (sometimes difficult to obtain due to the low levels of RAM in older PCs and the number of available slots to insert additional RAM)
- 100 Mbps Network Interface Cards
- A floppy drive or a hard drive
- Category-5 cable

In this semester, students should be encouraged to follow standard system analysis and design methods. This includes following all the regular steps (Dennis, 2000):

- Planning
 - Identifying value
 - Analyze Feasibility
 - Develop a work plan
 - Staff the project
 - Control and direct the project
- Analysis
 - Analysis
 - Information gathering

- Process modeling
 - Data modeling
- Design
 - Physical design
 - Architecture design
 - Interface design
 - Database and file design
 - Program design
- Implementation
 - Construction
 - Installation

Certain aspects of the system analysis and design phase of the project were key issues in getting the project going. For example, the report generated in the planning stage was used to garner lab space from the Dean. As a direct result of the design stage, the authors' cluster currently uses Scyld software, which is a commercial application, built on Linux Red Hat. A free version is downloadable from the Scyld web site (www.scyld.com).

Semester Two

Once the working system is firmly established, the focus of the project turns from system analysis and design to a systems-management/development mode. In the second semester, the focus should be to move from a "working system" to a "research-ready system." Items to work on in the second semester include public relations, equipment acquisition, system growth, and security. The public relations effort includes two main functions, media and presentations at civic organizations. The authors' University Relations office was eager to help with printed media. Additionally, after the first newspaper article was published, many news reporters were eager for an interview. Other media outlets include a web site that students can develop as a sub-project of its own. Many civic organizations, such as Rotary International, schedule weekly meetings and look for new and interesting guest speakers each week. The authors found presentations at these organizations easy to arrange through faculty members who belong to such organizations (but in some cases they required as much as 6-months advance scheduling). Equipment acquisition is as simple as asking. Every printed article and every civic presentation should discuss ways of donating to the project. Everybody who has a computer is likely to someday buy a new one. Making presentations at civic organizations and handing out fliers with contact information may prove to be one of the best methods of acquiring used equipment. Results may not be immediate, but they are lasting. Business leaders and private citizens alike have called the authors up to 9 months after a presentation to say that they are finally upgrading their system(s) and would like to donate their used equipment. It appears that people don't like to throw away "working equipment" – no matter how old it is. They much prefer to give it to a good cause. Another great source of used equipment is excess inventory. This includes not only the university's excess inventory from computer labs and faculty offices, but also the excess inventory department of major corporations.

System growth is important in that the system should be fast enough to be used as a research tool by faculty and others. This gives the system purpose and credibility. Taking into account the age of used equipment, a possible goal of the semester might be to develop a system that benchmarks at double or triple the speed of a current top-end PC. Goals such as this will help to motivate students, and give purpose to equipment acquisition activities.

Student members of the second-semester team do not necessarily need the high technical skill level that was recommended for the first-semester team members. Additionally, it is no longer advisable to use a student team that is part of another class such as *system analysis and design*.

An experimental class or a team of Independent Study students is more appropriate. It is important to assemble a group that is interested and dedicated to advancing the project. It is generally not overly difficult for a professor to hand-select a team of good students to work on a project of this type (especially if there has been some good press-coverage). If there are any first-semester students who want to continue the project in the second semester, it might be advisable to have one or two carry-over to lend continuity to the team. Students involved in the second semester will gain experience in public speaking, writing press releases, hardware, Linux OS, web-site development, equipment management and identification, trouble-shooting, and problem-solving. Building off of the previous resource requirements from semester one, semester two establishes the need for additional resources. Semester two resource requirements include:

- Students interested in the project
- Project mentor (faculty member)
- Small budget for flyers and other Public Relations materials
- Web site—hosted by the university or by a third party
- PR contacts—these can be members of the university faculty/staff or local community. They are individuals with connections to civic organizations with the ability to get the team on their presentation calendar. Also included in this group are media members of campus and local area news organizations.
- Computer and projector—these are needed for presentations to university groups and civic organizations. It is preferable to use a laptop for mobility. At many academic institutions these can be borrowed from a department, administration, or the university's library.

Optional Summer Semester

If the previous two semesters were conducted in fall and spring semesters, the summer semester provides an opportunity for the faculty member to assess the situation and make any preparations for the upcoming fall semester. If all has gone well, this break time might not be necessary. The authors do suggest a few activities to prepare for the next stage – which will be to offer a full elective course in cluster computing, as opposed to an independent study or experimental course that has been recommended up to this point. A graduate student earning independent study credit could conduct some of these activities.

Physical resources

If computers have been difficult to obtain, summer is a great time to look around campus for computer labs being upgraded. The old computers might already be spoken for, but there is usually a trickle-down effect – somewhere someone is using an old computer that will be replaced and sent to excess inventory. Other considerations include lab space. Summer is a good time to reallocate room space.

Course development

Having worked with the project for two semesters, the faculty member in charge has probably read plenty about cluster computers without exerting too much effort. Summer is a time to put together a reading list for students and to develop assignments. A masters-level student earning independent study credit might help by developing lab exercises.

Skill development

Faculty members who feel their Linux or networking skills could stand some improvement have an opportunity in the summer to brush-up on their basic skills. Plenty of Linux tutorials exist on the Internet and in print. Those who feel more comfortable with a GUI interface will be relieved to know that the Scyld software used for a Beowulf cluster comes with both Gnome and FVWM interfaces (the choice can be made at the log-in screen) as well as a text-based option.

Semester Three

By the third semester (the fourth semester if you count summer), the project should have progressed to the point that a full course may be offered in cluster computing. Where enrollment has been restricted in earlier semesters, the course is now ready to open to any interested students. It is recommended, however, to limit the number of students enrolled in the course to a manageable number.

The course content should consist of both theoretical perspectives on cluster computing and hands-on laboratory experience. Suggested material to be covered in the course includes:

- General clustering concepts and techniques
- Beowulf overview and fundamentals
- Linux operating system
- Building a Beowulf cluster computer
- Benchmarking and Research methods for verification and system analysis

The laboratory, or “hands-on,” portion of the course may be in the form of team or individual projects that involve building mini-clusters of 4 or 5 nodes each, from scratch. Each system built can be benchmarked and used for comparison to others. A well-organized class might include competitions for the fastest system given certain a specific list of hardware components to be used in each system. Students might experiment with system resources; for example, reallocating RAM between the nodes and the master node, looking for increased performance. Depending on resource availability, the student’s mini-clusters might later be added to the main cluster, enhancing its performance.

As an added incentive to the students, the instructor may consider allowing the best student from one semester to be the system administrator for the next semester. This will enhance the real-life working experience for the student. The focus of the project during this semester should be on increasing overall system performance. If the project has matured sufficiently, it might be possible to work toward developing software applications in conjunction with an advanced programming class. Students will benefit in this semester from gaining experience in systems development, problem solving, operating systems, networking, resource allocation, hardware configuration, troubleshooting, and project management in a real-life environment. One area in particular where students will gain valuable experience is in asking for help. Beowulf systems run on the Linux platform. Linux software is available free of charge; however, user manuals are expensive. As the students begin constructing their systems, they will have questions. This is a great opportunity for the students to gain experience using on-line resources such as bulletin boards and industry contacts as resources of information.

Summary

Cluster computing is an emerging answer to several technological dilemmas. The competitive nature of the business world has established a need for scalable, flexible, and reliable computing systems. Advanced applications are now requiring computing power not available in a stand-alone PC. Due to the rapid influx of new PCs on the market and ever increasing obsolescence rates businesses are left with excess equipment gathering dust or filling landfills. Cluster computing adds value to businesses while helping alleviate these problems. The future of high performance computing will be found in cluster computing as a substitute for traditional supercomputers, and at a reasonable cost. As the future draws near, preparing students becomes an important issue in an ever-changing environment. Today's students will face cluster-computing applications tomorrow; they need to be prepared for their careers after graduation. Cluster computing has already begun appearing as a part of university curriculums in the areas of computer science and engineering. The challenge remains, to introduce cluster computing into MIS programs in business schools. Business technology students need to be prepared with hands-on experience when they enter the job market. Cluster computing is a great opportunity to provide students with a unique learning experience.

As an experimental design, the authors developed a cluster-computing project at their institution, to expose business students to the process of building and operating a cluster computer for use in supercomputing applications. The project helped students to learn the technologies of distributed networking, securing networked environments and designing an application for distributed computing. A cluster is a resource for teaching various IS concepts through hands-on experience, application development and research, but it is a resource not limited to the computer science department or business school students. Anyone who needs serious computing power to conduct research or solve complex problems can use the services of a cluster. This would include, but not be limited to: chemists, physicists, biomedical scientists, architects, and engineers.

References

- Apon A. & Baker M. (2000) Network technologies. In M. Baker (Ed.), *Cluster computing white paper* (pp. 13-26). United Kingdom.
- Apon, A., Buyya, R., Jin, H., & Mache, J. (2001). Cluster computing in the classroom: Topics, guidelines, and experiences. *First IEEE/ACM International Symposium on Cluster Computing and the Grid* (CCGrid 2001), Brisbane, Australia.
- Baker, M. (2000). *Cluster computing white paper*. United Kingdom: University of Portsmouth.
- Baker, M., Apon, A., Buyya, R., & Jin, H. (2002). Cluster computing and applications. In A. Kent & J. Williams (Eds.), *Encyclopedia of Computer Science and Technology* (pp. 87-125). New York: Marcel Dekker. Retrieved from <http://www.gridbus.org/~raj/papers/encyclopedia.pdf>
- Baker, M. & Buyya, R. (1999). Cluster computing at a glance. In R. Buyya (Ed.), *High performance cluster computing: Architecture and systems* (pp. 3-47). Prentice Hall.
- Baker, M., Fox, G. C. & Yau, H. W. (1996). Review of cluster management software. *NHSE Review*, 1 (1).
- Boud, D. (1995). *Enhancing learning through self-assessment*. London: Kogan.
- Childers, R. C., Kitchens, F. L. & Sharma, S. K. (2003a). Cluster-computing as a multidisciplinary classroom tool. *Proceedings of Southwest Decision Science Institute (SWDSI) 34th Annual Conference*, Houston, Texas: pp. 440-445.
- Childers, R. C., Kitchens, F. L. & Sharma, S. K. (2003b). Integrating IS curriculum knowledge through a cluster-computing project - A successful experiment. *Hawaii International Conference on Education*, Honolulu, Hawaii.

- Dennis, A. & Wixom, B. H. (2000). *System analysis and design: An applied approach*. New York: John Wiley & Sons.
- Harvey, B. Z., Sirna, R. T. & Houlihan, M. B. (1998). Learning by design: Hands-on learning. *The American School Board Journal*, 182 (2), 22-25.
- Hoffman, F., Hargrove, W. & Schultz, A. (1999). The Stone Soupercomputer, Ornl's First Beowulf. Retrieved 2004 from <http://www.esd.ornl.gov/facilities/beowulf>
- Hyde, D. C. (2000). Education. In M. Baker (Ed.), *Cluster computing white paper* (pp. 110-119). United Kingdom.
- Korwin, A. R. & Jones, R. E. (1990). Do hands-on technology based activities enhance learning by reinforcing cognitive knowledge and retention? *Journal of Technology Education*, 1 (2), 26-33.
- Lane, M. G. (1981). Teaching operating systems and machine architecture - More on the hands-on laboratory approach. *Technical Symposium on Computer Science Education, Proceedings of the Twelfth SIGSE Technical Symposium on Computer Science Education*, AMC Press.
- Los Alamos News Letter (2002). *Smaller slower, supercomputers someday may win the race*. Retrieved from <http://www.lanl.gov/worldview/news/releases/archive/02-058.shtml>
- National Safety Council. (1999). *Electronic product recovery and recycling baseline report*. Retrieved July 1, 2002 from <http://www.nationalsafetycouncil.org>
- Read, B. (2002, August 7). Cornell U. bolsters its cluster-supercomputing program with \$60 million deal. *The Chronicle of Higher Education*. Retrieved from <http://chronicle.com/free/2002/08/2002080701t.htm>
- Rice, M. L., Wilson, E. K. & Bagley, W. (2001). Transforming learning with technology: Lessons from the field. *Journal of Technology and Teacher Education*, 9 (2), 211-230.
- Rochester Computer Recycling. (2002) *Benchmarks for computer obsolescence*. Regional Computer Recycling & Recovery. Retrieved July 1, 2002 from <http://www.rochestercomputer.com/regionalcomputerrecycling/benchmarks.htm>
- Scanlon, E., Tounглу, C. & Jones, A. (1998). Learning with computers: Experiences of evaluation. *Computers & Education*, 30 (1), 9-14.
- Sterling, T. (2000). An Introduction to PC Clusters for High Performance Computing In M. Baker (Ed.), *Cluster computing white paper* (pp. 3-12). United Kingdom.
- Sterling, T. L., Salmon, J., Becker, D. & Savarese, D. F. (1999). *How to build a Beowulf: A guide to the implementation and application of PC clusters*. Cambridge, MA: The MIT Press.
- Wilkinson, B. & Michael, A. (1999). A state-wide senior parallel programming course. *IEEE Transactions on Education*, 43 (3), 167-173.
- Wrege, R. (1982). Hands on computing. *Popular Computing*, 1, 110-123.

Biographies



Dr. **Fred L. Kitchens** has been an Assistant Professor of Management at Ball State University since September of 2000. He holds a Ph.D. from the University of Mississippi in Management Information Systems with a specialization in Computational Intelligence.

Recent interests involve cluster computing. As director of the Ball State Cluster Computing Research Project, his interests are in business applications and curriculum development. Dr. Kitchens sees great potential for cluster computing in business; and believes a principal component of that vision is in educating tomorrow's technology professionals today.



Dr. **Sushil K. Sharma** is an Associate Professor of Information Systems and Operations Management, at Ball State University, Muncie, Indiana, USA. Co-author of two textbooks and Co-editor of four books, Dr. Sharma's research contributions have appeared in many peer-reviewed national and international journals, conferences and seminars' proceedings. Dr. Sharma's primary teaching and research interests are in e-commerce, information systems security, ERP systems, database management systems, and knowledge management. Dr. Sharma has wide consulting experience in information systems and e-commerce and has served as an advisor and consultant to several government and private organizations including World Bank funded projects.



Dr. **Tom Harris** has been the Chair of the Information Systems and Operations Management Department since its inception in the Miller College of Business. Dr. Harris is also an Adjunct Professor in the Center of Information and Communication Sciences at Ball State University. The author of several articles in information systems, Dr. Harris also served as a research design consultant during his 30+ years at Ball State University. Dr. Harris has worked as an independent information systems and management consultant for many other organizations.