Title: Cluster Computing Curriculum Revision Project

Abstract of the Professional Development Project:
Using obsolete equipment, the project will build a Linux based Beowulf cluster computer carefully documenting each step in the process. This will be converted into a curriculum project to help networking and computer science students reproduce the machine while studying form factors and machine architecture. Beowulf clusters use obsolete equipment combined together to achieve low end supercomputer performance. This project builds on a previous project that was highly successful, incorporating the lessons learned into the new work.

Statement of problem, purpose, and rationale of the Professional Development Project:
Students in CSC and ITN courses need an extended networking problem to sharpen their skills. This study will help the project developer enhance skill in cluster development with the goal of producing sufficiently developed materials to guide introductory networking (ITN) and Computer Science (CSC) students through the process. This will build on previous VCCS grant funded work by the project developer in the area of cluster computing. The past work has excited students in the introductory courses, and the next step is to do the curriculum work necessary to have students build a cluster computer in a class. This project will achieve that goal. The past work was highly successful, including New Horizons and national computing conference sessions.

To solve a cluster computing problem, students must work in three areas. First, they must properly load an operating system on a personal computer. This project will primarily use Linux which encourages students to become involved in the open source area. The second step necessary is to connect the computers into a network and make them communicate. The final step is to configure the computer to attach to the cluster. The cluster is usually implemented with products like PVM (Parallel Virtual Machine) or MPI (Message Passing Interface) on the Linux platform. This will also involve loading software which requires a parallel system to see the cluster perform.

The overall goal of detailed documentation of the necessary steps is important because much of the material available assumes too much on the part of the person new to clustering. This will allow students to be successful in the assembly and test of the system.

A secondary goal of the work is to update the cluster using newer equipment while incorporating the lessons learned in the previous work. It is anticipated that a comparative study of performance will be done with the previous results.

For research materials, Beowulf details found through http://www.beowulf.org/ will be used. They recommend the work of Robert Brown from Duke (http://www.phy.duke.edu/~rgb/ ) who has an online book on clustering. Also, several schools use books either authored or co-authored by Thomas Sterling (http://www.cacr.caltech.edu/~tron/ ) on cluster computers, so they will be used in the study as well. These materials are ones the project developer has used before, but the second time through the work is more understandable.

Methodology:
Assess available computing equipment to use in the project (May 16 to May 21)
Determine a configuration for the cluster (May 22 to June 4)
Finalize calculations related to power and heat requirements (June 26 to July 2)
Test the cluster and the casing structure, modify as necessary (July 3 to July 23)
Documentation Materials Review and development (July 24 to July 29)
Implementation of a web site to house project materials (July 30)
Classroom materials Ready for Student Use (August 1)
Final Report (August 1)

Step 1 will be to research web sites on cluster computing including single box solutions. One such project is the LittleFE project. In the context of what is found, it will be necessary to examine the equipment that is available and select appropriate machines for the project.

Step 2 will be to determine a configuration for the cluster.
Step 3 will be to calculate power and heat requirements for the new cluster.

Step 4 will be to setup and test the cluster and its software.

Step 5 will be to formalize and document the development process.

Step 6 will be to review, assemble, develop and edit the documentation of the project. In addition, classroom materials will be developed to help students build a cluster using Linux.

The final step will be to implement a web site for the materials (which the project developer does routinely) and write the final report. It is anticipated that the project will be complete by August 1 to allow for use in Fall 2009 classes. The results of this work will be developed sufficiently for submission to conference for presentation.

**Anticipated outcomes and benefits of the Professional Development Project:**
The researcher’s professional development requires continual World Wide Web, hardware and programming study to enhance those areas of teaching. This project will advance my skill with technology by providing a project development experience that is directly applicable to classroom content. In addition, the use of Microsoft Project to manage the project will advance my skill in that area.

The benefit to the students would be new hands on projects in cluster computer implementation. In addition, it is likely that new course segments in programming may result, allowing students to advance in the age of clustering computing. In addition, the work with Linux will help students that transfer because that is the operating system used at most universities in the computer science department. The researcher’s professional development plan includes continual hardware and architecture study to enhance those areas of teaching. The benefit of this project to the discipline in general is the availability of online materials on clustering, documented and available for instructional use. The benefit of this project to the college would be the involvement of its faculty in a significant, sponsored research effort that has a direct influence on course content and the understanding of technology of the college. Finally, the benefit to the VCCS will be the availability of the project materials to faculty within the system to enhance teaching and learning throughout the system. A primary tool for dissemination of these results will be the College Web Site.

**Collaboration:**
This project requires the use of Internet resources, which by definition means that the effort is collaborative in some sense. Collaboration will also occur in the discussions on clustering with colleagues at other colleges. Much of the work of researchers and practitioners in this area are available as Internet resources, providing a rich way to explore information. Where necessary the developer will seek the opinions of others familiar with the clustering and its use in academic programs. Comments of students will be solicited, reviewed and incorporated into the materials where appropriate.

**Evaluation of the Professional Development Project:**
The evaluation of the project will be based on the ease with which students can implement the cluster given the plan developed. Success will be indicated by a working cluster and enhanced if the machine in fact achieves low end supercomputer performance as predicted in the research materials at http://www.beowulf.org/. It is also anticipated that the current example would perform better than the machine of 2006. The evaluation mechanism for the student materials will be questionnaires developed to assess student and instructor reaction to the work. A project guide for student use will indicate what a working cluster looks like and allow detailed evaluation of the student work. The experiences of the author and others in the classroom and computer laboratory will be documented.

**TECHNICAL INFORMATION: BUDGET, IN-KIND CONTRIBUTIONS AND DISSEMINATION OF RESULTS**

Many of these items are not repeated here but are on file with the VCCS. This project extends the 2006 funded cluster computing project. It recycles old computing resources, and will involve students in developing a cluster. The College donates the used computers. The instructor donates and/or makes arrangements on all other materials. The funding requested is for release time for the project director.
**College Commitment:**
The College commitment to the project includes equipment which is obsolete for other purposes and assigned from other areas of former use. The approximate market value of $13000.00 includes a server and workstations. The equipment is available to the project director in the networking lab.

**Personal Commitment:**
The developer will use personal materials including books, CD-ROMs, tools and software to assist in the development. The developer will use images collected for the project, most through digital photography. In addition, many of the programming resources used will be those of the developer, especially those that will help plan the complex program the cluster will run. The result of the project will make resources more widely available.

**Dissemination of Results:**
It is anticipated the results will be presented at the Peer Group and the New Horizons conference. In addition, a proposal will be submitted to the Association for Computing Machinery (ACM) Special Interest Group on Computer Science Education (SIGCSE) Conference.