Curriculum in High-performance Scientific Computing

Lloyd Fosdick

CU-CS-621-92 October 1992

University of Colorado at Boulder
DEPARTMENT OF COMPUTER SCIENCE

Curriculum in High-performance Scientific Computing

Lloyd Fosdick

CU-CS-621-92 October 1992

University of Colorado at Boulder

Technical Report CU-CS-621-92
Department of Computer Science
Campus Box 430
University of Colorado
Boulder, Colorado 80309

Abstract

This report summarizes the goals and plans for a project to improve undergraduate education in computational science. It is taken from the executive summary of a proposal submitted to the National Science Foundation and funded for three years in September 1990.

The project includes development of a one-year course in high-performance scientific computing, and laboratory facilities, a laboratory manual, tutorials, and software. The course uses scientific workstations, a Cray YMP, a Thinking Machines CM2, and an Intel iPSC/2. Problems solved on these systems are drawn from molecular dynamics, climatology, and tomography. Numerical methods, performance measurement, and scientific visualization are an important component of this course. Development of a course in scientific visualization is a collateral part of this project.

Students coming into this course are expected to have had an undergraduate course in numerical analysis, and be familiar with the Unix environment.

The course materials will be exported to other colleges and universities through summer workshops, the Internet, and interuniversity collaboration.

The following pages contain the Executive Summary of a proposal submitted to NSF, under their Educational Infrastructure Program, by L. Fosdick, Gitta Domik, and Elizabeth Jessup, 30 May 1990

I. Executive Summary

This proposal is concerned with education in computational science for undergraduate students of computer science and computer engineering. We propose the development of a one-year course and supporting laboratory facilities that will give students an understanding of high-performance computing systems and training in their use. These systems include vector and array supercomputers such as an Alliant, Cray Y-MP, and Connection Machine, and high-performance workstations. The laboratory facilities we propose, and our course materials, will support courses in visualization, structural analysis, computational fluid dynamics, and others.

Students receiving the type of instruction we propose here will be much better suited for participating in scientific computational work like that conducted at Los Alamos and Lawrence Livermore National Laboratories, NASA laboratories, and the National Center for Atmospheric Research. We also believe that this training will give them a better understanding of how computers are used by scientists and engineers, thus giving them a better perspective for the design of new computing systems. Finally, the course facilities will encourage interdisciplinary work between computer science students and students from other fields of science and engineering.

Thus the objectives of this proposal are:

- To teach undergraduate students how to use high-performance computing systems in scientific and engineering applications.
- To provide an advanced computing environment for undergraduate instruction in particular applications.
- To develop instructional materials: course notes, problem examples and exercises, manuals, and software.
- To export our instructional materials and experience to support undergraduate studies at other institutions.
- To encourage interdisciplinary studies for computer scientists and engineers.

- To encourage research scientists to become involved in undergraduate education.
- To prepare students for work in the field of computational science.

Course in Scientific Computation. This one-year course is intended for upper division undergraduate students in computer science or computer engineering. As prerequisites, students will need a basic understanding of numerical analysis, computer architecture, Unix, C and Fortran.

The major theme of the course is how to solve scientific problems on high-performance computing systems. This incorporates the study of numerical methods, matching algorithms to machines, measurement and modelling of performance, and scientific visualization. The course will use a small set of *prototype* scientific problems, such as heat flow in a rod or the motion of a classical many-particle system, and present the numerical and computational tools used to solve such problems. Students will use high-performance workstations, an Alliant, a Cray Y-MP, an Intel iPSC, and a Connection Machine, and they will study the performance of these systems in solving the prototype problems.

A major part of the course development will be the creation of texts and computer software. A minimanual and tutorial software for each computing system will be developed. A text for each prototype problem and associated software will be developed. This text will provide a description of the physical background of the problem and the numerical techniques used to solve it. It will also discuss the problems associated with developing good algorithms for solving the problem on the various computing systems. Three additional texts will be provided: a review of the basic concepts from numerical analysis that are prerequisite for the course; performance measurement and modelling; graphics and visualization. The texts will be modularized to make it easy to use only selected parts of the course.

Laboratory facilities. The laboratory is to be equipped with high-performance workstations, such as IBM RS/6000 series machines and X-terminals; a small number of graphics workstations, such as Silicon Graphics, equipped with hardware to display output on a TV and make video tapes; and an Intel iPSC. In addition to these facilities, remote access to a Connection Machine and a Cray Y-MP is provided. The iPSC, CM time, and Cray Y-MP time are all provided as part of the matching, and thus are not part of the funds requested from NSF. Furthermore, an Alliant FX/8 is already available for instructional use through the university's computing center.

Collaborators. Faculty members at the following Colleges and Universities have agreed to collaborate with us on this project:

- University of Texas at San Antonio
- University of New Mexico
- New Mexico Institute of Mining and Technology
- Colorado School of Mines
- Fort Lewis College

The collaborators will advise us on course development, critique and contribute to course materials, and participate in a "shake down" workshop after the course materials are completed. Collaborators have agreed to use our course, or a closely related course using our materials, at their institution. (Statements from the collaborating institutions are in the Appendix.) Each collaborating institution will be provided with a workstation and X-terminal for course development. The PI will work with each collaborator to secure additional computing facilities and supercomputer time.

Review and evaluation. An advisory committee consisting of three scientists from major laboratories such as Los Alamos, Sandia, and NCAR will be appointed by the PI. This committee will meet at least four times during the course of the grant to advise on course development and review outcomes.

Students in the course will be asked to keep a log of their activities. This log is to include problems in using or understanding course materials, and difficulties encountered with laboratory facilities. They also will be asked to provide a summary evaluation at the end of the course.

Management plan. The PI will hold overall responsibility for the development of the course materials, and will do the major part of the work in writing the texts. Other text development will be done by two coinvestigators. Software exercises and examples will be written by graduate student assistants. The PI will also be responsible for acquisition of the laboratory facilities.

The schedule of development is:

- Year 1: Write texts and software, acquire laboratory equipment, summer shake-down workshop.
- Year 2: First offering of course, two summer workshops for instructors and students, critique course.
- Year 3: Second offering of course, run course at collaborating schools, two summer workshops for instructors and students, summary review of course, write final report.

Budget. The total proposed budget is $756K in direct costs, of which the University of Colorado will provide $139K. Including indirect costs of $117K, a total of $734K is being requested form NSF. In addition, the National Center of Atmospheric Research (NCAR) and the Center for Applied Parallel Processing (CAPP) will provide computer time on their Cray Y-MP and CM2, and the university will provide access to an Alliant.

Almost half of the senior personnel cost ($81K of the three-year total of $175K) is budgeted for the first year when much of the initial course development will take place. The university will provide $41K of that $81K, by granting the PI a sabbatical leave to work on the proposed project.

The equipment portion of the budget is $242K, to provide a computing lab with high-performance workstations such as IBM RS/6000s and two graphics workstations such as Silicon Graphics Irises. Each collaborating institution is to get one machine such as an IBM RS/6000 and one X-terminal. The equipment is to be purchased during the first year in order to be in place for the first workshop in the summer of 1991.

A sizable portion of the budget (most of the $128K in participant support costs) is for the dissemination of the course materials. This will support five summer workshops. The first of these workshops, in the summer of 1991, will be an "evaluation workshop" attended by faculty from the collaborating institutions. The other four workshops, 2 each in the summer of 1992 and 1993, will be attended by twenty participants each, including faculty members and students from other institutions. Each of these two-week workshops will be based on the newly developed course materials (minimanuals, notes, software).