LittleFe + BCCD + CSERD = Acme; Computational Science Education on the *Move*

Charlie Peck[1], Tom Murphy[2], Dr. Paul Gray[3], Dr. David Joiner[4], Alex Lemann[1], Kevin Hunter[1], Kristina Wanous[3]

1: Earlham College - [charliep lemanal kevin] @ cs.earlham.edu, 2: Contra Costa College - tmurphy @ contracosta.edu, 3: Kean University - djoiner @ kean.edu, 4: University of Northern Iowa - [gray kwanous] @ uni.edu

The overwhelming majority of the High Performance Computing (HPC) resources currently deployed are dedicated to research rather than education – yet the nation faces a shortage of HPC expertise, due largely to the lack of faculty trained in HPC pedagogy. To address a part of this situation our group designs and implements hardware, software, and curriculum to support teaching parallel and distributed computing, computational science, and Grid technologies to a range of 9-16 STEM students and faculty. Acme is the sum of LittleFe, a complete multi-node Beowulf-style portable computational cluster, the Bootable Cluster CD (BCCD), a complete Linux distribution designed for HPC and computational science education, and the Computational Science Education Reference Desk (CSERD), a source of verified and validated curriculum modules.

This education platform is visceral: students participate with the science. An amino acid necklace folding into a protein and visualized in simulated-time, provides an engaging presentation to the *Xbox* generation, attracting more students to STEM disciplines. We must reach students with multidisciplinary, interactive, visual, computational science curriculum, allowing them to envision themselves in an exciting science based career.

Hardware

Overview The hardware component of Acme is LittleFe, a portable computational cluster. LittleFe is a complete multi-node Beowulf-style portable computational cluster. The entire package weighs less than 50 pounds, easily travels via checked baggage, and sets-up in 5 minutes.

Design LittleFe's design grew out of our work building stationary clusters and our experience teaching workshops in a variety of places that lacked parallel computational facilities. The principle design constraints for LittleFe are:

- $3,000 (USD) total cost
- Less than 50lb (including the Pelican travel case)
- Less than 5 minutes of setup time
- Minimal power consumption

The current production LittleFe design is composed of the following parts:

- 6 motherboards (mini-ITX or similar, 1GHz+ CPU, 512MB RAM, 100MB ethernet)
- 6 12VDC-ATX power supplies
- 1 320 Watt 110VAC-12VDC switching power supply
- 1 40GB 7200RPM ATA disk drive (2.5" form factor)
- 1 DVD/CD optical drive (slim-line form factor)
- 1 8 port 100MB ethernet switch
- 1 rack assembly
- 1 1610 Pelican travel case
- Fasteners, cabling, and mounting hardware

The $3,000 (USD) cost per unit includes about 10 hours of student labor to assemble and test each unit. This includes liberating the BCCD image onto the disk drive and configuring the users. The motherboards, CPUs, and RAM comprise the bulk of the cost. With all 6 nodes idling, LittleFe draws about 80 Watts of power (about the same as a light bulb). When running a CPU-intensive molecular dynamics simulation on every node, LittleFe draws about 88 Watts of power.

Current and Future Projects Moving forward our design goals are cheaper, smaller, more powerful, and easier to assemble and maintain. Contra Costa College and Kean University both have student-faculty teams working on next generation LittleFe prototypes based on the Efika, a small form-factor PowerPC, system-on-a-chip (SOC) system. The Efika boards offer low cost, low power consumption, and small size. We're also exploring partnerships with organizations interested in manufacturing LittleFe units.

More Information `http://LittleFe.net`

Software

Overview For the operating platform, Acme uses the Bootable Cluster CD (BCCD), a project of Paul Grays group at the University of Northern Iowa. The BCCD comes in two flavors, a live CD version and a liberated version. The live CD makes it possible to turn any x86-based lab into a teaching cluster in a matter of minutes. The liberated version, used by Acme, is permanently installed on LittleFes disk drive.

The BCCD contains hundreds of software applications to support parallel and distributed computing and computational science education, this greatly reduces the friction associated with setting-up an environment for HPC-based education.

Built-in Tools A small sample of the HPC development and clustering tools included in the BCCD include:

- gcc, g77, and development tools, editors, profiling libraries and debugging utilities
- Cluster Command and Control tools
- MPICH, LAM-MPI and PVM
- The X Window System
- OpenMosix with openmosixview and userland openMosix tools
- Torque and Maui scheduler support
- octave, gnuplot, Mozillas Firefox and about 1,400 userland utilities
- Network configuration and debugging utilities
- Ganglia and other monitoring packages

Built-in Science Science software and curriculum modules available via the BCCDs list-packages tool include:

- GROMACS molecular dynamics primarily for biological molecules
- mpiBLAST parallel implementation of the basic local alignment search tool
- MrBayes - Phylogenetic reconstruction using Bayesian analysis
- FFTW - Fastest Fourier Transformation in the West

Current and Future Projects Groups at the University of Northern Iowa and Earlham College are working on curriculum module and next generation BCCD development. Grid interoperability, both science software and curriculum modules, through the Open Science Grid and the TeraGrid are close on the horizon for Acme.

More Information `http://bccd.cs.uni.edu`

Science Education

Overview The Computational Science Education Reference Desk (CSERD) is a collection of reviewed curriculum modules designed to support high performance computing and computational science education. CSERD is a Pathways project of the National Science Digital Library (NSDL) and supported by the National Science Foundation (NSF).

Goals and Motivation Our goal is science education for 9-16 students and faculty from primarily undergraduate institutions. This proves a difficult task when the majority of educators and scientists do not have a common definition of computational science education, nor is there an agreement on how, when, or where it should be offered. Many educators and scientists think of education and science as two totally disparate worlds.

Educational needs are very similar to science needs: The software in the materials must work, the science must be correct, and the results must be accurate, useful, and applicable. CSERD gives teachers and scientists alike access to resources that meet these three guidelines by reviewing them in the following context:

- Verification Is the model logically correct and does it follow from the physical and mathematical laws used?
- Validation Did the model correctly predict the modeled phenomena?
- Accreditation Does the model or simulation fulfill its educational purpose?

What's Available Resources in CSERD range over multiple scientific disciplines and cover materials for primary, secondary, and post-secondary education. For HPC education, we have tutorials and example code available to illustrate different levels of parallelism: embarrassingly parallel parameter searches, domain decomposition programs that require only nearest neighbor communication, and tightly coupled N-body examples. Each of these codes shows a parallel programming example in the context of solving a science problem.

Current and Future Projects Student/faculty teams at Kean University, the University of Northern Iowa, and Earlham College are working on identifying, developing, and testing curriculum modules for inclusion into CSERD in a variety of natural science disciplines. Through the SC Education Program and the Humanities Research Institute of the University of California we are beginning to explore how computational techniques might be applied to teaching and research in a wider range of disciplines. The SC Education Program is also sponsoring our work with the Navajo Technical College at Crownpoint Reservation. Working with colleagues at NTC we are looking for ways to include computational methods into a variety of classes across the science curriculum.

More Information `http://cserd.nsdl.org`