Spring 2004

csit

Integrating Advanced Computing with Science, Engineering and Liberal Arts

School of Computational Science & Information Technology

From Greek Frogs To Modern Computer Tools

INSIDE THIS ISSUE

A recent article in Science suggests that the number of whales before hunting started was much larger than anyone had suspected. These controversial results originated from calculations on genetic variation among today's whales. Peter Beerli at CSIT developed the method and the computer tool used in the study.

It all started with substantially smaller animals, though. When Peter Beerli, now an FSU professor, was a PhD student in Switzerland he used water frogs (*Rana esculenta* group) as models in his evolutionary studies. The Greek archipelago provided him with the ideal material. The frogs on the different islands descended from a common ancestor, but had been isolated from their relatives on other islands for about five million years, separated by the saltwater barrier. And once they were cut off, the frogs on each island began to vary from their neighbors because of the genetic mutation "clock" that always ticks at a slow, but predictable, rate.

Peter Beerli

Habitat of the water frog species Rana cretensis *of Crete, Greece*

Most mutations are minor and harmless, and might not even create any obvious external difference in the organism. They do, however, drift into the common gene pool of a population where they provide scientists with an important historic record. With the right computational methods, it is possible to peer into the past by analyzing the genetic variations of today.

At the time Peter Beerli began his studies, there were simple techniques available to compare the genetic variation between any two, isolated populations. But when he also wanted to take into account the migration rates between different mainland populations, he got stuck. The pairwise approach, that he was able to use for the isolated populations, was not good

continued on page 2

School of
Computational Science
& Information Technology

It is delightful to introduce our CSIT newsletter and describe what you will be reading in these pages as it grows and develops with us.

The CSIT mission has several facets—supporting the FSU scientists who use the supercomputers, catalyzing interdisciplinary research that uses computations, and performing innovative research on computational methods and approaches. Our newsletter—and associated features on our web site—will help bring them all to light.

We'll help keep you aware of the wide variety of computational research going on at FSU, inside and outside of CSIT. Computational work is an increasingly important part of many areas of scholarship and each edition of our newsletter will help publicize some of the important and interesting work being done with the FSU supercomputers and other computing resources.

And of course the newsletter will keep you informed of how CSIT itself is growing and changing. We have a nucleus of strong scholars who work at the interface of real world problems and the computations needed to solve them. We are planning a more formal educational program that goes beyond teaching individual courses, one that we hope will prepare a new generation of interdisciplinary computational scientists. And our newsletter will help keep you abreast of that progress.

Write us and tell us what you'd like to read and learn.

Joe Travis
Director, CSIT

enough when frogs from different populations were able to mingle. Peter Beerli had to develop new techniques and incorporate them in a computer program. His program, named MIGRATE, now allows scientists to estimate migration rates, as well as present and historic population sizes, using genetic data.

In the journal Science of July 2003, two other scientists described how they used MIGRATE to calculate historic population sizes of North Atlantic whales. Analyzing the genetic variation of today's whales, together with additional information on average generation time and breeding age, they got surprising results. Populations of humpback and fin whales before whaling were about ten times greater than previously thought. If the results are correct the current populations are a long way from recovery and conservation strategies need to be much more extensive.

That debate will go on for a length of time. Meanwhile, Peter Beerli continues to refine MIGRATE. Originally programmed for desktop computers, where an individual run of a big dataset would take weeks, the program has now been rewritten (parallelized) to run many times faster on the CSIT supercomputers and on computer clusters. Faster calculations make it possible to use much larger datasets than before, thus getting more reliable results. Biologists across the globe now use MIGRATE to find clues on evolution and ecology hidden in the DNA of birds, insects, fish and humans.

beerli@csit.fsu.edu

Peter Beerli

A water frog from Crete (Rana cretensis)

Grid Computing Is Hot

The term Grid Computing designates the effort to effectively use globally distributed computers on single large tasks. Existing computational infrastructure represents a tremendous potential of cheap computing power, connected via the Internet. Many people have been challenged by the possibility of using this capacity, and Grid Computing has become a major research topic in Computer Science. Thus it is good news that FSU, through CSIT, is developing its own Grid thanks to new external funding.

The Grid gets its name from the electrical power grid, which provides power to the user with minimum hassle. The user does not know whether the electricity comes from a huge power plant or from a local windmill. Ideally, the computer Grid should work the same way. However, accessing these widely distributed computers, represents a huge challenge. Not only is it a difficult task to write software that can run on desktop computers as well as supercomputers and on different networks, but, you have to deal with scheduling issues and security problems to make the Grid work.

One may think of Grid Computing as an elaborate extension of research in parallel and distributed computing, both of which represent methods of spreading parts of a large mathematical problem among different computers or processors. Software named Globus, developed at Argonne National Labs, has become the *de facto* standard as a platform for new Grid software development through the Globus Toolkit. Those interested in learning more about Globus and the Grid can log on at www.globus.org.

Michael Mascagni (CSIT/Computer Science) and Ashok Srinivasan (Computer Science) recently were awarded a $200,000 equipment grant from the Defense University Research Instrumentation Program (DURIP), and Matt Clausen (CSIT) was given a $80,000 equipment grant from Sun Microsystems, Inc. for creating the physical grid infrastructure at CSIT for FSU. In addition to this equipment funding, FSU has also committed some of its own resources to the Grid to enhance the usability at a distance of the neutron experimental facilities at Oak Ridge National Laboratory. This is in cooperation with NSF's recently funded Southeastern TeraGrid Extension for Neutron Science (SETENS) award. The FSU portion of this activity is being run by Dennis Duke (CSIT/Physics).

mascagni@fsu.edu

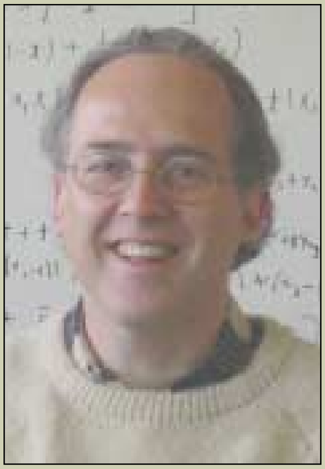
Macintosh Joins The Big Boys

In the ongoing race to create faster and more powerful computers, a supercomputer at Virginia Tech achieved a ranking of third in the world in November 2003. The computer is a cluster, i.e. a combination of many smaller computers, connected to work as one. Each of the 1100 components is a slightly modified Macintosh Power Mac G5 computer, the world's fastest personal computer, according to Apple. The idea of Virginia Tech was to create a supercomputer based on standard, off-the-shelf components.

The Top500 list comprises the 500 fastest computer systems being used today. The first list was published in 1993 and it has been updated twice a year since. Ranking is based on the ability to solve a specific system of linear equations. Though a large number of the supercomputers on the list are clusters, this is the only one made up of Macintosh computers.

Source: computing.vt.edu/research_computing/terascale/ and www.top500.org

Making Pictures Of The Invisible



While many of us are amazed at the development of computer graphics, few have an idea of the technology making it possible. How do the zeros and ones in the computer memory produce a colorful representation of a climate simulation, or 3-D images from inside a brain?

Professor Gordon Erlebacher at CSIT knows. For two decades, he and several FSU colleagues have been working on how to display complex objects and movements on the computer screen. Visualization is the tongue-twisting term for it, and FSU has a long tradition of research in this area. Visualization requires lots of computing power, available at FSU since the mid-1980's, when the university acquired its first supercomputers. Today, CSIT runs the Visualization Laboratory, which is a facility that serves a dual purpose: research into the latest visualization techniques, and software and hardware to support the needs of FSU scientists who desire to better understand their data through the visual senses. The unit is directed by professor Erlebacher and his colleague, computer science professor David Banks.

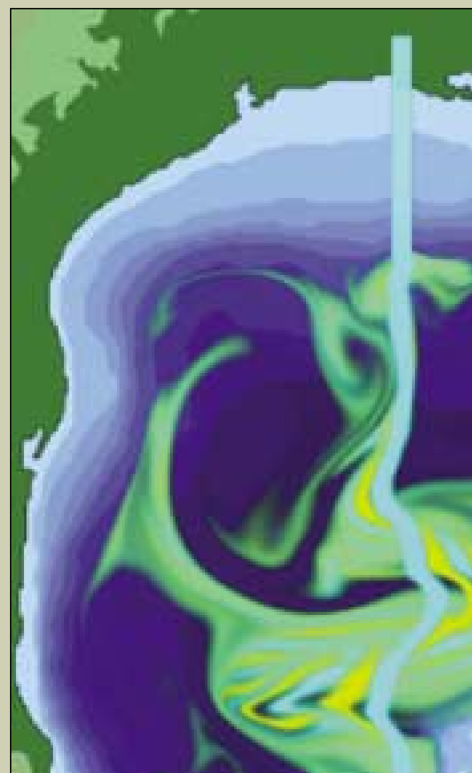
THE LOOK OF A STORM

In several fields where scientists need to analyze and understand huge amounts of information, visualization is an important tool. For many years, Gordon Erlebacher has cooperated with scientists working with the flow of fluids, like water in the oceans or air in the atmosphere. For example, to understand and predict the movement of a storm, a meteorologist would measure (or simulate) wind speed and direction at millions of points in space at regular intervals. At each of these points, the wind velocity can be displayed as an arrow with a certain direction and length, also known as a vector. A collection of vectors drawn at each point of a smooth surface (or volume) is called a vector field, and a proper choice of which vector to represent will produce useful information to the meteorologist. Animations of these vector fields, which often change as a function of time, can present researchers with valuable clues regarding the dynamics of the atmosphere. For example, in addition to tracking storm trajectories, and storm intensities, vector fields make possible the tracking of dispersants, such as pollutants. Similar versatility can be found in many fields of science, for example oceanography, (see figure).

The visualization itself is not just a service to the applied sciences, but also a field of intense research and development. An ever so sophisticated picture on a computer screen is made up of points, lines, and polygons, and animation is created by rotating, scaling and otherwise transforming these elements. Each step in the process is governed by an algorithm—a set of equations—developed by mathematicians like Gordon Erlebacher, or by computer scientists. Even the fastest computer would not be able to accomplish much without good algorithms.

GAMES PAVE THE WAY

The graphics card (also called graphics board or video card) is a key component in computer visualization. It



The CSIT group of Gordon Erlebacher includes Yunsong Wang, who graduated in fall with a Masters thesis in computer science, but still works on remote visualization. So does PhD student Zhenyu Lu, computer science, who is also taking visualization courses given by Dr. Banks. Another PhD student, Brian Bouta, is working on the use of geometric algebra to understand and visualize 3D time-dependent vector fields. Patrick Crawley, an undergraduate student, works with Dr Erlebacher on hardware-based graphics algorithms. The work on WEB-IS and middleware is done in collaboration with David Yuen (U. Minnesota) and Geoffrey Fox (Indiana U.).

Yunsong Wang

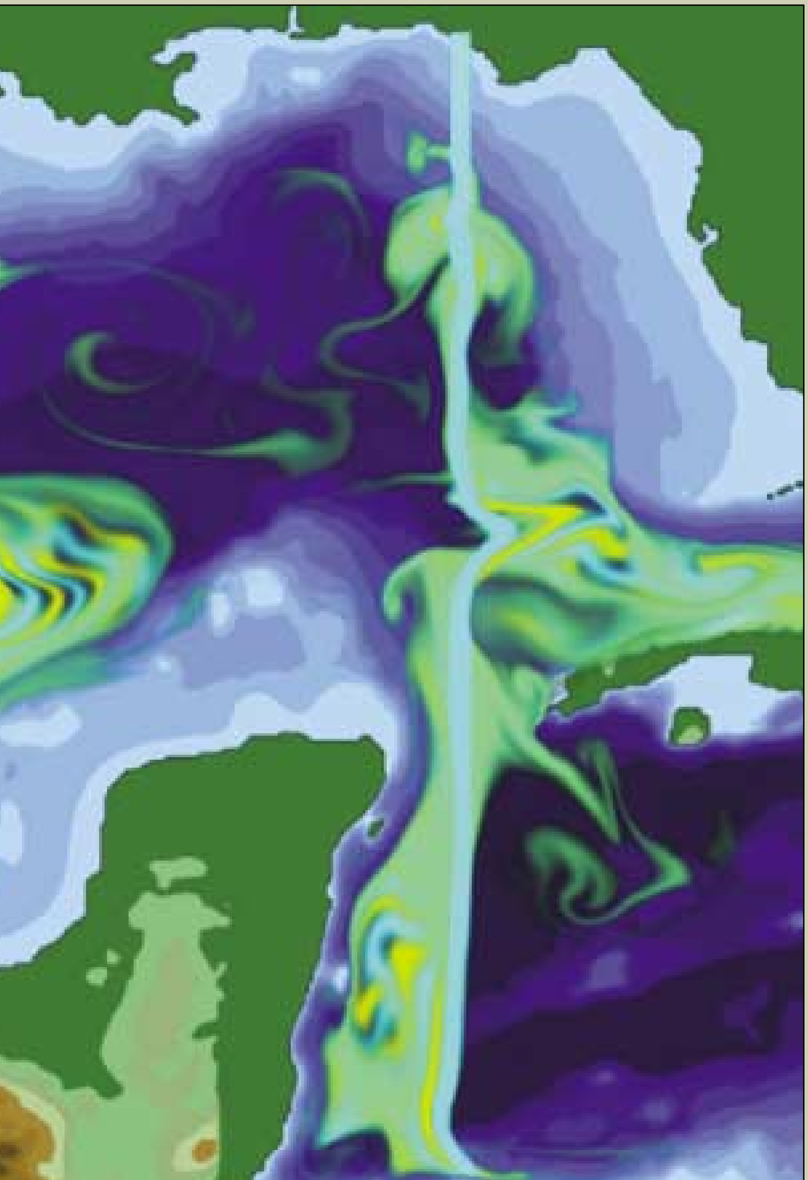
Zhenyu Lu

Brian Bouta

Patrick Crawley

takes the digital information—the zeros and ones—from the computer and transforms it into something that we can see on the computer screen. The popularity of computer games has resulted in very cheap and powerful graphics cards, which are essentially self-contained powerful supercomputers in their own right. In other words, modern graphics cards are not only responsible for acting as a conduit between computer memory and the faster graphics board memory. They are also fully programmable, allowing the programmer to bypass the CPU if they so desire. By recasting standards algorithms into a form where identical operations are executed for each pixel of an image, speed increases of one to two orders of magnitude over conventional CPU-based algorithms are common.

For the last several years, Professor Erlebacher has been developing some of the first algorithms dedicated to the display of time-dependent vector fields, based on dense representations, that run completely on the graphics board. On a $300 Radeon 9700, vector fields at resolutions of one megapixel can be displayed at over 40 frames/second. The challenge is to develop these algorithms to run on the most common and cheap consumer graphics cards (nVidia and ATI), a task that has proven difficult because the field is evolving at such a rapid rate. Gordon Erlebacher was recently awarded a PEG (Program Enhancement Grant) to keep developing these techniques and ensuring that they find their way into applications in Meteorology, Oceanography, and Geophysics.



THE INTERNET DREAM

Visualization across the Internet is another challenge of Gordon Erlebacher's. His graduate student Zhenyu Lu just started his graduate studies to continue the software work started by Yunsong Wang in his Master's thesis. The software is a client that is a Java applet called WEB-IS, developed together with students from University of Minnesota. WEB-IS currently allows a scientist to order data from a remote computer and display it on his or her own computer, the client.

The vision is however much more far-reaching. Ongoing research seeks to connect WEB-IS to a more general middleware infrastructure to isolate clients from services, which execute tasks on behalf of the clients. In the words of Gordon Erlebacher, "middleware is a term used for software that is like plumbing: it provides services that nobody wishes to know about, connecting inputs to outputs".

The dream is that, through properly designed middleware, a scientist should be able to access information, such as a particular earthquake dataset, from one or several locations around the world without any a priori knowledge of where the data is located. Thus, the data could be duplicated at multiple sites, yet accessed in a transparent manner by the user. The possibilities for global, scientific cooperation would increase tremendously. This is indeed a great dream and an inspiring mission for young programmers.

erlebach@csit.fsu.edu

The figure clarifies the dynamics of ocean currents in the Gulf of Mexico through the periodic release of blue and yellow dye along two vertical strips. The distance between strips of alternating color is a measure of the magnitude of the ocean current. Data courtesy J. O'Brien (COAPS).

New Employees And Visitors

The Computational Evolutionary Biology (CEB) group grew substantially during 2003. Three professors were hired: **Peter Beerli**, **Fredrik Ronquist** and **Gavin Naylor**. **Mark Holder** started as a post doc to work with David Swofford, and will be at CSIT for five more years due to ITR funding (page 7). Two lab technicians were also hired, **Tina Kesler** and **Julie Ryburn**. **Thomas Buckley** is visiting the group for about 4 months. He is a post doc working at Landcare Research, New Zealand's foremost environmental research organization.

CSIT has been reinforced by several post docs during the last year. Four of them joined Professor Hussaini's group in 2003, namely **Shannon Grady**, **Mohamed Jardak**, **Svetlana Poroseva**, and **Ali Uzun**. In January, **Marwan Al-Haik** joined the group. **Pierre-Antoine Absil** is working with Kyle Gallivan and Anuj Srivastava (Statistics) on large-scale computational problems on manifolds with applications to eigenvalue, model reduction, and statistical inference problems. Dr. Gallivan also hired **Younes Chahlaoui** as a post doc. He arrived in January to work on large-scale model reduction problems. **Marcus Garvie**, **Hugh MacMillan**, and **Catalin Trenchea** started to work with Dr. Gunzburger.

Long-term visiting faculty include **Dr. V. Bhujanga Rao** from India, who is working with Dr. Hussaini and will stay till August 2004. He is working on problems of hydrodynamic flow noise related to underwater bodies.

The Technical Support Group added three new members over the last year, **Elaine Dennison**, **Danny Laughlin** and **Bob Smith**. All were added to improve end user support. Elaine and Danny's work draws upon their skills in UNIX/Linux system administration. Bob's primary duties include service process analysis and oversight of service response.

Eva Ronquist, administration group, was hired in October to help publicize CSIT and computational science at FSU.

Ming Cai, Associate Professor of Meteorology, joined CSIT during the fall of 2003. Before moving to FSU he worked at the University of Maryland for more than 10 years. Dr. Cai works with climate modeling and prediction, and global warming.

Dr. Angela Kunoth, Professor of Mathematics, is visiting Dr. Max Gunzburger's group during a sabbatical from University of Bonn, Germany. Dr. Kunoth's research interests center around issues in numerical analysis for partial differential equations. With Dr. Gunzburger she works on wavelet methods for the efficient numerical solution of nonlinear optimization problems constrained by partial differential equations. She arrived in September 2003 and will stay at CSIT through March 2004.

Bob Smith, Danny Laughlin, and Elaine Dennison (from left)

5 Questions To Dr. Michael Navon

In November 2003, Dr Michael Navon visited Beijing, China, as an invited guest of the Academia Sinica and the Institute of Atmospheric Physics.

Q What was the destination of your trip?

A I visited IAP, the Institute of Atmospheric Physics, in Beijing. IAP is the highest academic organization in the basic research of atmospheric sciences in China, and with elite staff in science and technology.

Q Why were you invited?

A I have been working with students from China for many years, both from IAP and from elsewhere. Many of my former students now hold research positions in North America. Among them is Dr X. Zou, professor of meteorology at FSU.

Q Which research areas did you present?

A I gave a lecture about adaptive observations. These are additional meteorological data, collected in regions where we can assess that they will improve quality of forecast of catastrophic events such as hurricanes. Another lecture topic was reduced order modeling, which allows us to run complex models with less computer time and memory.

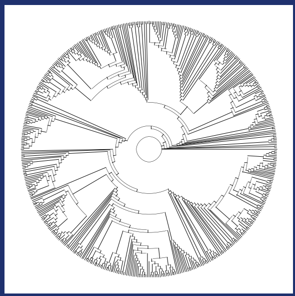
Q What is your impression of Beijing?

A The IAP is located in a region with very impressive research centers and high technology. The ancient city of Beijing itself is growing into a bustling new metropolis. It has a population of over 13 million, many new high-rise buildings, and the city is in the middle of the preparations for the 2008 Olympics.

Q In what ways will the cooperation continue?

A My host, Professor Jiang Zhu, assistant director of IAP, and I agreed upon new research collaboration concerning two grant applications. In addition, I agreed to provide guidance to a postdoctoral fellow who will be recruited to IAP.

navon@csit.fsu.edu



NSF Funding For "Tree of Life" Software Development

A team of CSIT researchers has been awarded $1.8 million to develop software for NSF's "Tree of life" project, which aims at exploring the evolutionary relationships among all species of living organisms. The FSU/CSIT effort, led by CSIT professor David L. Swofford, is part of a larger informatics project, awarded $11.6 million by NSF's Information Technology Research program. The informatics project as a whole is oriented towards the analysis of the tremendous amounts of data that "Tree of life" generates. The FSU award will be used for graduate student stipends, programmer salaries, and a 128-processor computational cluster. About half of the FSU award will be subcontracted to collaborators at four other institutions.

Reconstructing the "Tree of life" has long been a high priority for biologists, but doing so requires an extraordinary computational effort. The ITR-funded project will bring together researchers from computer science, biology, and mathematics to solve the difficult computational challenges and to deliver useful software to the research community. This intense collaboration would not occur during the course of normal professional pursuits, and CSIT's multidisciplinary focus makes it an ideal home for the project.

swofford@csit.fsu.edu

School of Computational Science & Information Technology

Middle School Students Make Planets Dance

Every Friday morning for the past semester, CSIT has been honored by the presence of a couple of bright students, Luka Sharron and Evan Fuller. The two sixth graders have been working with their mentors, Professors Dennis Duke and Jorge Piekarewicz, to simulate planetary movement on the computer screen in a project called "The Dance of the Planets".

Luka and Evan had no previous experience in programming, but they learned quickly under their mentors' guidance. They started by collecting data for each planet. Among other variables, they had to find out how long it takes for a planet to complete its orbit around the sun, and at what speed the planet moves in different parts of the orbit. They wrote an Excel program to process the data, and when they ran it the different planets started to move gracefully on the computer screen.

Luka Sharron and Evan Fuller (from left)

Together with seven other student groups, they presented their work to an audience of parents, and fellow students on January 13. Next semester the students will take on new projects at the FSU Magnet Lab.

Since 1991, the middle school mentorship program has been directed by Joan Crow, who is a teacher at the School of Arts and Sciences in Tallahassee, a public charter school. CSIT and its predecessor SCRI have been involved for several years.

Seminars on Materials Science

The year 2003 marked the launch of the "CSIT/ Martech Series", which fosters collaboration and understanding in the field of Computational Materials Science, an area where FSU has considerable expertise. Physicists, engineers, and applied mathematicians have participated, and attendance has been high.

The seminar series was initiated by Michael Mascagni and is being run in association with FSU, Martech. Seminars are monthly, and all are invited. For schedules of coming seminars, go to http://www.martech.fsu.edu/ Seminars/ or http://www.csit.fsu.edu/seminars/seminars.php.
mascagni@csit.fsu.edu

CSIT—School of Computational Science and Information Technology

The mission of CSIT is to be the focal point of computational science at the Florida State University. The school supports and develops a variety of high performance computing facilities, accessible to the university community. CSIT is designed to overlap with existing departments and schools to provide a venue for interaction among faculty and students across many disciplines.

Please visit our website at www.csit.fsu.edu!

CSIT
Dirac Science Library
Florida State University
Tallahassee, FL 32306-4120
Telephone: 850 644-1010
Fax: 850 644-1593

Director: Dr. Joseph Travis
850 644-7024

Editor: Eva Ronquist
850 644-0196
evaron@csit.fsu.edu

News from CSIT will be issued three times a year. Free subscriptions and single copies can be ordered from the editor.