

Innovative Technology for Computer Professionals

Computer

NOVEMBER 2011

<http://www.computer.org>

CODESIGN FOR SYSTEMS AND APPLICATIONS

- HUMAN EAR RECOGNITION, P. 79
- CROWDSOURCING MAPS, P. 90
- AUTOMATED PERSONAL ASSISTANTS, P. 112





chist-era call 2011

Announcement of the 2nd International Call for Research Projects within the ERA-Net CHIST-ERA

Deadline for proposal Submission: 17th of January 2012 (17:00 GMT)

CHIST-ERA Stands for *European Coordinated Research on Long-term Challenges in Information and Communication Sciences and Technologies ERA-Net*

The CHIST-ERA initiative is looking for highly innovative and multidisciplinary collaborative projects in ICST. Especially CHIST-ERA is open to new ideas and original solutions, involving interdisciplinary skills. In addition, the transformative research done in CHIST-ERA shall explore new topics with potential for significant scientific and technical impacts. The projects must involve at least three partners from three different and participating countries. In the Call 2011, two new and hot topics are addressed:



From Data to New Knowledge

The challenge is to produce new computational concepts, models, tools and methodologies to automatically and reliably extract new knowledge from large amounts of heterogeneous, unstructured data. Typical data include multilingual and multimedia data such as found on the web (text, speech, image, video, ...) and data generated by human organisations in the course of scientific, industrial or service activities (medical data, 3D object representations, advanced manufacturing data, ...). Though much activities are already going on in the field, existing systems are far from offering a highly reliable extraction of knowledge from any type of data, and basic research is still needed to explore new concepts and models for challenging tasks such as machine reading, processing of noisy data, multiscale model handling...

Green ICT, towards Zero Power ICT

The decrease of energy consumption for computation or communication is an important challenge for future. In addition this renders possible the design of autonomous systems scavenging their own energy from their environment. A broad range of solutions are currently envisaged from component to system of systems levels, where research is needed in various areas. Project proposals should address the issue of energy consumption in computation, information, sensing or communication systems from a global system perspective. Highly innovative approaches are expected at any of the system layers, from the nano-scale level to the architectural, software or protocol layers...

Call Information: www.chistera.eu

Disclaimer:

the information in this announcement is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user therefore uses the information at its sole risk and liability.



Computer

Innovative Technology for Computer Professionals

Editor in Chief

Ron Vetter
University of North Carolina
Wilmington
vetterr@uncw.edu

Associate Editor in Chief

Sumi Helal
University of Florida
helal@cise.ufl.edu

Associate Editor in Chief, Research Features

Kathleen Swigger
University of North Texas
kathy@cs.unt.edu

Associate Editor in Chief, Special Issues

Bill N. Schilit
Google
schilit@computer.org

Computing Practices

Rohit Kapur
Synopsis
rohit.kapur@synopsys.com

Perspectives

Bob Colwell
bob.colwell@comcast.net

Web/Multimedia Editor

Charles R. Severance
csev@umich.edu

2011 IEEE Computer Society President

Sorel Reisman
s.reisman@computer.org

Area Editors

Computer Architectures

Tom Conte
Georgia Tech
Steven K. Reinhardt
AMD

Distributed Systems

Jean Bacon
University of Cambridge

Graphics and Multimedia

Oliver Bimber
Johannes Kepler University Linz

High-Performance Computing

Vladimir Getov
University of Westminster

Information and Data Management

Naren Ramakrishnan
Virginia Tech

Multimedia

Savitha Srinivasan
IBM Almaden Research Center

Networking

Ahmed Helmy
University of Florida

Security and Privacy

Rolf Oppliger
eSECURITY Technologies

Software

Robert B. France
Colorado State University

Web Engineering

Simon Shim
San Jose State University

Column Editors

Discovery Analytics

Naren Ramakrishnan

Virginia Tech

Education

Ann E.K. Sobel

Miami University

Entertainment Computing

Kelvin Sung

University of Washington, Bothell

Green IT

Kirk W. Cameron

Virginia Tech

Identity Sciences

Karl Ricanek

University of North Carolina, Wilmington

In Development

Chris Huntley

Fairfield University

Industry Perspective

Sumi Helal

University of Florida

Invisible Computing

Albrecht Schmidt

University of Stuttgart

The Known World

David A. Grier

George Washington University

The Profession

Neville Holmes

University of Tasmania

Security

Jeffrey M. Voas

NIST

Social Computing

John Riedl

University of Minnesota

Software Technologies

Hal Berghel

Lero—the Irish Software Engineering Research Centre

Advisory Panel

Carl K. Chang

Editor in Chief Emeritus

Iowa State University

Hal Berghel

University of Nevada, Las Vegas

Doris L. Carver

Louisiana State University

Ralph Cavin

Semiconductor Research Corp.

Rick Mathieu

James Madison University

Naren Ramakrishnan

Virginia Tech

Theresa-Marie Rhyne

Consultant

Alf Weaver

University of Virginia

Publications Board

David A. Grier (chair),
Alain April, **David Bader**,
Angela R. Burgess, **Jim Cortada**, **Hakan Erdogmus**,
Frank E. Ferrante, **Jean-Luc Gaudiot**, **Paolo Montuschi**,
Dorée Duncan Seligmann,
Linda I. Shafer, **Steve Tanimoto**,
George Thiruvathukal

Magazine Operations Committee

Dorée Duncan Seligmann (chair), **Erik R. Altman**,
Isabel Beichl, **Krishnendu Chakrabarty**, **Nigel Davies**,
Simon Liu, **Dejan Milojević**,
Michael Rabinovich, **Forrest Shull**, **John R. Smith**, **Gabriel Taubin**, **Ron Vetter**, **John Viega**,
Fei-Yue Wang, **Jeffrey R. Yost**

Editorial Staff

Judith Prow
Managing Editor
jprow@computer.org
Chris Nelson
Senior Editor

Contributing Editors

Camber Agrelius
Lee Garber
Bob Ward

Design and Production

Larry Bauer
Design
Olga D'Astoli
Cover Design
Kate Wojogbe
Jennie Zhu

Administrative Staff

Products and Services Director
Evan Butterfield
Senior Manager, Editorial Services
Lars Jentsch

Manager, Editorial Services
Jennifer Stout
Senior Business Development Manager
Sandy Brown
Senior Advertising Coordinator
Marian Anderson

Circulation: *Computer* (ISSN 0018-9162) is published monthly by the IEEE Computer Society. IEEE Headquarters, Three Park Avenue, 17th Floor, New York, NY 10016-5997; IEEE Computer Society Publications Office, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720-1314; voice +1 714 821 8380; fax +1 714 821 4010; IEEE Computer Society Headquarters, 2001 L Street NW, Suite 700, Washington, DC 20036. IEEE Computer Society membership includes \$19 for a subscription to *Computer* magazine. Nonmember subscription rate available upon request. Single-copy prices: members \$20; nonmembers \$99.

Postmaster: Send undelivered copies and address changes to *Computer*, IEEE Membership Processing Dept., 445 Hoes Lane, Piscataway, NJ 08855. Periodicals Postage Paid at New York, New York, and at additional mailing offices. Canadian GST #125634188. Canada Post Corporation (Canadian distribution) publications mail agreement number 40013885. Return undeliverable Canadian addresses to PO Box 122, Niagara Falls, ON L2E 6S8 Canada. Printed in USA.

Editorial: Unless otherwise stated, bylined articles, as well as product and service descriptions, reflect the author's or firm's opinion. Inclusion in *Computer* does not necessarily constitute endorsement by the IEEE or the Computer Society. All submissions are subject to editing for style, clarity, and space.

Innovative Technology for Computer Professionals

Computer

www.computer.org/computer

CONTENTS

ABOUT THIS ISSUE

The essence of the codesign challenge for exascale systems is to use the key design criteria of embedded systems—cost and power consumption—while creating systems that are useful and effective over the broad range of applications needed to advance science. The cover features in this special issue have been selected to cover a cross-section of the codesign space and the relevant concerns and challenges.

COVER FEATURES

19 GUEST EDITORS' INTRODUCTION

19 Codesign for Systems and Applications: Charting the Path to Exascale Computing

Vladimir Getov, Adolfo Hoisie, and Harvey J. Wasserman

The clock speed benefits of Moore's law have ended, and researchers must codesign future exascale HPC systems and applications concurrently in an integrated manner to achieve higher performance under stringent power and reliability constraints.

22 Rethinking Hardware-Software Codesign for Exascale Systems

John Shalf, Dan Quinlan, and Curtis Janssen

The US Department of Energy's exascale computing initiative has identified hardware-software codesign as a central strategy in achieving more agile hardware development. Hardware simulation and code analysis tools that facilitate deeper collaboration between hardware architects and application teams will be an essential component of the codesign process.

31 Codesign for InfiniBand Clusters



Sayantana Sur, Sreeram Potluri, Krishna Kandalla, Hari Subramoni, Dhabaleswar K. Panda, and Karen Tomko

Codesigning applications and communication libraries to leverage underlying network features is imperative for achieving optimal performance on modern computing clusters.

37 Codesign Challenges for Exascale Systems: Performance, Power, and Reliability

Darren J. Kerbyson, Abhinav Vishnu, Kevin J. Barker, and Adolfo Hoisie

The complexity of large-scale parallel systems necessitates the simultaneous optimization of multiple hardware and software components to meet performance, efficiency, and fault-tolerance goals. A codesign methodology using modeling can benefit systems on the path to exascale computing.

COMPUTING PRACTICES

44 The iPlant Collaborative: Cyberinfrastructure to Feed the World

Dan Stanzione

As plant biology becomes a data-driven science, new computing technologies are needed to address many formidable challenges. The iPlant Collaborative provides cyberinfrastructure for researchers and developers to collaborate in creating better tools, workflows, algorithms, and ontologies.

PERSPECTIVES

53 Defending against Buffer-Overflow Vulnerabilities

Bindu Madhavi Padmanabhuni and Hee Beng Kuan Tan

A survey of techniques ranging from static analysis to hardware modification describes how various defensive approaches protect against buffer overflow, a vulnerability that represents a severe security threat.

RESEARCH FEATURE

61 Algorithmic Trading

Giuseppe Nuti, Mahnoosh Mirghaemi, Philip Treleven, and Chaiyakovorn Yingsaeree

Traders increasingly use automated systems for one or more stages of the trading process, yet the secrecy and complexity of the algorithms prompt providing an overview of how these systems work.

For more information on computing topics, visit the Computer Society Digital Library at www.computer.org/csdl.

IEEE Computer Society: <http://computer.org>
 Computer: <http://computer.org/computer>
computer@computer.org
 IEEE Computer Society Publications Office: +1 714 821 8380

Flagship Publication of the IEEE Computer Society

November 2011, Volume 44, Number 11

6 The Known World

 The Honest Give-and-Take
David Alan Grier

9 32 & 16 Years Ago

Computer, November 1979 and 1995
Neville Holmes

NEWS

11 Technology News

Turning on the Lights for Wireless
 Communications
Lee Garber

15 News Briefs

Lee Garber

MEMBERSHIP NEWS

70 IEEE Computer Society Connection

73 Call and Calendar

COLUMNS

75 Green IT

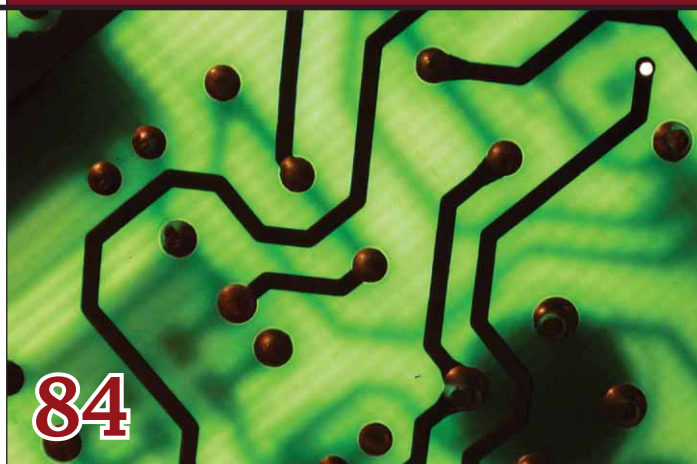
End-to-End Energy Management
**Yung-Hsiang Lu, Qinru Qiu, Ali R. Butt,
 and Kirk W. Cameron**

79 Identity Sciences

Human Ear Recognition
Arun Ross and Ayman Abaza



See www.computer.org/computer-multimedia for multimedia content related to the features in this issue.



83 Industry Perspective

Opportunities in the Mobile Search Market
**José Luis Gómez-Barroso, Claudio Feijóo,
 and Ramón Compañó**

87 Hard Issues

If Anything in This Life Is Certain,
 It's That You Can Kill Any ISA
Shubu Mukherjee

90 Social Computing

Crowdsourcing Maps
Mikhil Masli

112 The Profession

Automated Personal Assistants
Kai A. Olsen and Alessio Malizia

DEPARTMENTS

4 Elsewhere in the CS

78 Computer Society Information

94 Career Opportunities



Reuse Rights and Reprint Permissions: Educational or personal use of this material is permitted without fee, provided such use: 1) is not made for profit; 2) includes this notice and a full citation to the original work on the first page of the copy; and 3) does not imply IEEE endorsement of any third-party products or services. Authors and their companies are permitted to post the accepted version of their IEEE-copyrighted material on their own Web servers without permission, provided that the IEEE copyright notice and a full citation to the original work appear on the first screen of the posted copy. An accepted manuscript is a version which has been revised by the author to incorporate review suggestions, but not the published version with copyediting, proofreading and formatting added by IEEE. For more information, please go to: http://www.ieee.org/publications_standards/publications/rights/paperversionpolicy.html.

Permission to reprint/republish this material for commercial, advertising, or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to the IEEE Intellectual Property Rights Office, 445 Hoes Lane, Piscataway, NJ 08854-4141 or pubs-permissions@ieee.org. Copyright © 2011 IEEE. All rights reserved.

Abstracting and Library Use: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

IEEE prohibits discrimination, harassment, and bullying. For more information, visit www.ieee.org/web/aboutus/whatis/policies/p9-26.html.

ELSEWHERE IN THE CS

Computer Highlights Society Magazines

The IEEE Computer Society offers a lineup of 12 peer-reviewed technical magazines that cover cutting-edge topics in computing including scientific applications, design and test, security, Internet computing, machine intelligence, digital graphics, and computer history. Select articles from recent issues of Computer Society magazines are highlighted below.

Software

Lateness is the most common form of software project failure. Its causes can seem complex when viewed from ground level, but are surprisingly simple with a slightly more distanced perspective. In “All Late Projects Are the Same” in *Software*’s November/December issue, Tom DeMarco says that what’s really wrong with software folks is that they are continually beating themselves up for something that’s somebody else’s fault. DeMarco asserts that the louder the complaints about project lateness, the more likely it is that the project set out to deliver marginal value and was therefore kicked off under the false premise that it could be completed on the cheap.

Intelligent Systems

To help people live better in today’s digitally explosive environment, the authors of “Cyber-Individual Meets Brain Informatics” in the September/October issue of *IS* envision a Cyber-Individual (Cyber-I) that is the counterpart of a real individual in the physical world. Brain informatics, an emerging interdisciplinary field that systematically studies the human information processing mechanism, provides the principles of individual modeling that guide Cyber-I’s core design and intelligence upgrade. The Cyber-I is intended to create a powerful demand for brain informatics research on individual information-processing differences and provide a testbed for evaluating future results obtained from that research.

Computer Graphics

In “Digital-Content Authoring” in *CG&A*’s November/December special issue, guest editors Takeo Igarashi of the University of Tokyo and Radomir Mech of Adobe introduce recent advances in digital-content-creation techniques, ranging from 3D modeling to behavior authoring and image editing. Although the articles in this special issue address diverse problems, they provide a good overview of techniques common to authoring problems in general.

Computing

Given its leading role in high-performance computing for modeling and simulation, the US Department of Energy has a tremendous need for data-intensive science. The datasets it generates significantly outstrip current analysis capabilities. More comprehensive analysis would help scientists discover and identify unanticipated phenomena and expose shortcomings in current simulation methodologies and software. Also, real-time data analysis would enable intelligent design and refinement of experimental processes. “Data-Intensive Science in the US DoE: Case Studies and Future Challenges” in the November/December issue of *CiSE* locates the challenges and commonalities among three case studies and illuminates, in detail, the technical challenges involved in realizing data-intensive science.

SECURITY & PRIVACY

S&P’s September/October special issue on cyberwarfare addresses the use of cyberattacks as an instrument of warfare. The four papers selected for the issue, introduced by guest editors Thomas A. Berson of Anagram Laboratories and Dorothy E. Denning of the US Naval Postgraduate School, address topics relating to the use of cybermilitias in cyberwarfare, policy and legal issues concerning state use of cybercapabilities, military principles for conducting cyberwarfare, and

strategic deterrence of cyberattacks against national infrastructure.



Much research into pervasive computing has been devoted to systems comprising a small number of devices that interact with a single individual or a small group of users. Body-worn sensors and sensors embedded in the user's environment are used to infer user location, activity, and information about the user's immediate surroundings, shaping the concept of context awareness. In the October-December issue of *PvC*, guest editors Paul Lukowicz, Tanzeem Choudhury, and Hans Gellersen assert that as technology becomes truly pervasive, we must proceed from the "single user, single system" perspective to large-scale heterogeneous systems that involve many devices and many individuals collaborating across different spatial and temporal scales.



In *IC*'s September/October issue, guest editor Craig W. Thompson of the University of Arkansas introduces nine articles on virtual world architectures—seven in the current issue and two to appear in future issues. The articles explore the current limitations of virtual worlds, deconstruct their architectures, and consider how the architectures might evolve to extend the technology's applications. Topics include the integration of 3D virtual world viewers with Web browsers as well as the federation and extension of virtual worlds, their accuracy, and relevant standards.



Although general-purpose CPUs have traditionally been the dominant player in both mainstream and high-performance computing systems, recent years have seen a major shift toward GPUs. Originally developed primarily for graphics and video applications, researchers are increasingly harnessing GPUs by using programming languages such as CUDA and OpenCL to provide large performance speedups for general-purpose applications. Guest editor David Brooks of Harvard University introduces *Micro*'s September/October special issue on GPUs versus CPUs with a discussion of advances and challenges in the field of hybrid CPU/GPU computing.



The July-September issue of *MultiMedia* features an article titled "Mobile Visual Search: Architectures, Technologies, and the Emerging MPEG Standard" in which

authors from Stanford University, Nokia Research Center, and Qualcomm review current mobile-search architectures and key component technologies. They also review MPEG activity to define a working draft for a new visual-search standard in the mobile context by February 2012.



Guest editors for the September/October special issue of *IT Pro* present six articles on the future of Web applications. According to their introduction, "In the coming years, billions of devices will be connected to the Internet, and they'll access and share information through the Web." Three articles focus on Web app development, one on digital diaries as an application scenario, one on the cloud computing ecosystem, and the sixth on the shrinking boundary between mobile and Web applications.



Multimedia Web Extras

Computer's November multimedia-enhanced digital edition includes the content listed below. You can also find these items online at www.computer.org/computer-multimedia.

Codesign for InfiniBand Clusters

Sreeram Potluri, one of the authors of "Codesign for InfiniBand Clusters," shares a deeper look into how the InfiniBand cluster works.

Extreme-Scale Computing

In an interview conducted during a recent meeting of the IEEE Computer Society's Magazine Operations Committee, Loyola University's George K. Thiruvathukal describes exascale, or extreme-scale, computing and some of its design challenges.

The History of Supercomputing Research

Larry Smarr, founding director of the National Center for Supercomputing Applications, provides a brief background into the history of supercomputing research.

The Known World (Behind the Scenes)

David Alan Grier gives a behind-the-scenes look at how he composes his popular "The Known World" column for *Computer*.

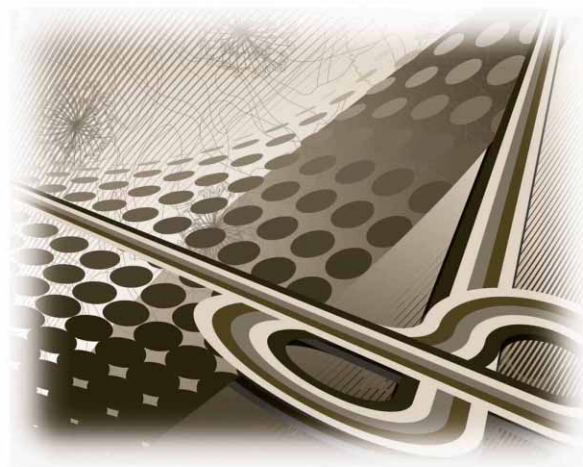
Tips for Patent Applicants

A podcast episode based on an article published in *Computer*'s September 2011 issue reviews basic points to keep in mind once you've successfully filed for a patent.

THE KNOWN WORLD

The Honest Give-and-Take

David Alan Grier
George Washington University



Information technology has not only expanded the scale and scope of global markets, it has also provided the means for probing the meaning of every give-and-take transaction.

The e-mail quickly became a constant nag, a dark angel visitant to remind me of my wrongs. Each time, it asked if I would evaluate my recent automobile purchase at Mr. Tony Pro's Auto Mall. Each time, I replied with a flick of the delete key. I was in no mood to give Mr. Pro any information that might strengthen his hold on me.

Under the best of circumstances, the negotiations over a new car are fraught with inequalities and offer all the advantages to the seller. The sticker price is a fiction. The invoice is no closer to the truth, as a regional office pays incentives on each car sold to better control of flow of products from factory to market. Every bit of information about automotive preferences or driving habits gives a chit of power to the dealer.

Often, the customer has no greater power than bluff and bluster, both of which were factors in my negotiation. However, while most customers buy a car only once every few years—or in my case, 14 years—the seller has the dominant position, having ample experience in offering a ready defense that sports a strong handshake, a firm smile, and a protest that he's doing all he can for you.

So when the honest give-and-take was done, and the new car was parked in front of my house, I was in no mood to give Mr. Pro any more information than was necessary. Although he could protest with all his might that he was merely trying to serve me better, I nevertheless harbored the fear that he was trying to gain the upper hand in a future market transaction.

THE VALIDITY AND THE VALUE

Information technology has not only expanded the scale and scope of global markets, it has also provided the means for probing the meaning of every give-and-take, for determining if the bargain was valid and if the result was valuable.

Although they're often treated as if they were the same thing, validity and value are very different concepts. Valid activities are those that are done well, that follow basic principles, that demonstrate the truth of the underlying concepts. Valuable activities are those that are important: they expand the field, find application to other disciplines, or generate a flow of money from satisfied customers.

As engineers, we tend to use experts to determine validity and markets to assign value. Experts review new ideas, test their underlying assumptions, and analyze how they were produced. If the results of this effort meet their standards, the experts declare the idea valid. Things can change. The field can evolve. However, as far as the experts can determine, the new idea is a valid addition to the body of knowledge.

In assigning value, markets look beyond validity. An idea can be clever, be completely valid, and represent a substantial intellectual achievement, but still have no value whatsoever. Value need not be measured in terms of the profit a concept can produce in the public market or the numbers of workers its production can employ. An idea can be valuable if it provides tools that expand the field, has applications to other technical subjects, or merely simplifies the task of creating other new ideas.

Yet, markets are influenced by human factors that can temporarily mask value. One party can completely mislead the other into assigning an inflated value to a product. Those who work in the marketplace, such as those who offer automobiles for

sale, need to protect their decisions by controlling the flow of information with a paired set of marketing tools: the suggestion system and the recommendation system.

SOLICITING SUGGESTIONS

Suggestion systems search for valuable ideas from customers or employees. Recommendation systems present ideas with the intent that customers or employees will find value in them. Both have long histories that could probably be traced, with little difficulty, to transactions on the Silk Road. However, both have been the object of careful analytic study and engineering practice. They show how we systematically try to extract or create ideas.

Suggestion systems are processes that solicit feedback from customers or employees. They're characterized by the familiar, although commonly mocked, suggestion box, usually a container with a slot in its top that accepts slips of paper offering new ideas. With the rise of the modern factory, suggestion systems acquired the trappings of an engineering discipline. They had a theory of operations, a set of best practices, and a professional society, the National Association of Suggestion Systems, which was located in downtown Chicago.

By the middle of the 20th century, the directors of NASS felt that their technology had become crucial to modern industrial management. "Having once and for all demonstrated its undeniable worth," they explained, "the Suggestion System is here to stay, to flourish and thrive and add its full, fair quota to the steady onward march of American Progress through the years."

The standard model for suggestion systems was fairly straightforward. A suggestion box collected ideas from customers or workers. These ideas were sorted by a suggestion clerk, who took them to the appropriate managers for validation. These man-

agers determined if the suggestions were valid and estimated the amount of money that might be earned or saved from each. From this review, the validated suggestions moved to a senior manager or operations committee, who determined the value of each suggestion by looking at the potential income or savings in the context of the entire company. They selected the most valuable suggestions for implementation and rewarded the individuals who had put their ideas in the box.

Despite its confidence in suggestion systems, NASS was forced to admit that suggestion systems often failed to produce any value for

Both suggestion systems and recommendation systems have been the object of careful analytic study and engineering practice.

the company. These systems were often expensive to operate and often met resistance from managers and engineers.

World War II produced the one brief period when suggestion systems were effective. It was no ordinary time. Many managers were open to the suggestions of outsiders as they were in new leadership roles and had no stake in existing production or operational systems. All felt the urgency of the war. Many, if not most, had relatives or neighbors in combat. By the middle of the conflict, all war production plants were required "to provide machinery whereby each man may submit ideas and suggestions for doing the job better." This requirement came with the promise that these systems would "tap a vast new reservoir of ideas, welding our productive genius into a united effort for victory."

Yet, suggestion systems died a quick public death after the war.

NASS declared that "nearly 95 percent of all attempts to operate suggestion systems were unsuccessful." Although the organization continued operations for a decade, it then disappeared.

Suggestion systems soon became objects of ridicule by the public. Cartoons showed suggestion boxes sitting over trash cans, shredders, and toilets. The Computer Society's suggestion box currently sits locked on a lunchroom shelf. No one knows what might be in the box or where the key might be.

SUGGESTIONS TO RECOMMENDATIONS

The Internet has considerably simplified the process of soliciting and organizing suggestions. Questionnaires can be sent by mail and collated into a central database. Ideas can be sent to managers for required validations. The news of a successful idea can be spread to entire workforces or customer populations. Modern technology has turned the suggestion system into the recommendation system—the software that gathers data on customers and identifies products that each might find valuable.

Recommendation systems, sometimes known as collaborative filtering systems, have proven more valuable than their progenitors. Most large Internet retailers have embraced such systems. It's unusual to purchase items without being told what others have bought, what offerings go well with our selections, or what new product might be of interest to us. One major retailer, Netflix, even sponsored a contest to develop a new recommendation algorithm and, in 2009, rewarded an approach that it identified as a substantial improvement over the prior state of the art.

Recommendation systems have succeeded where suggestions systems did not for a pair of fundamental reasons. First, they address the problems of mass consumption,

THE KNOWN WORLD

which are often simpler to grasp and easier to manipulate than the problems of mass production or mass distribution, which were the common focus of suggestion systems. Second, they exploit one of the fundamental strengths of information processing: the ability to record what we actually do, rather than the ideas we propose.

According to a recent survey of the technology, most recommendation systems utilize four types of data. Only the first, rating data, comes from the sort of questionnaires the auto dealership asked me to complete. The remainder is data that describes us—that is, demographic data—or data that describes our actions in the marketplace, such as behavior or transaction data. It's one thing to provide misleading responses on a questionnaire, as it only means that we're lying to others. It's quite something else to engage in misleading market behavior, as such actions mean that we're lying to ourselves.

ACTIONS RATHER THAN WORDS

For the questions from Mr. Tony Pro's business, I provided the most useless of lies. I scrolled through his form and checked the neutral rating of 4 for each query. However, he ultimately turned the tables on me by

capturing a little bit of my identity. After completing the questionnaire, his system began to bombard me with notices to have my new car checked by his skilled mechanics. The notices ranged from enticing to threatening, from promising new features for the car to threatening to void the warranty should I drive far more than the engineering specifications recommended.

One day, I received a notice that I would get a free tune-up if I took the car to the shop in the middle of the week. It was a little hook that deftly hid the great truth that there's nothing offered for free in this world that doesn't require payment at some future date. After delivering my car to the dealership, I found a chair in the corner of the showroom, where I began to work on some research on my laptop.

After 20 or 30 minutes, an employee asked me if I would like to use the wireless connection. Grateful for the opportunity, I accepted his offer and then asked him about his job at the dealership. I indicated that I had worked for the auto industry during my college years and was interested in how the work might have changed.

"I'm the data mechanic," my new friend replied.

"Data mechanic?" I asked. "Does that mean you manage the computers and networks in the office?"

"Yes," he responded, "and also the diagnostic and data collection systems in the shop."

A moment passed before I realized what he had told me. My car was sitting in a mechanic's stall, freely volunteering my habits as an owner and driver. It couldn't quite reveal every detail about every trip, but it could download information about the number of miles I had driven, the range of my travels, and the way I operated the car. Perhaps I hit the brakes too hard. Maybe I accelerated too fast from a stopped position. Most likely I dawdled on the freeway, forcing others into a line behind me. All of this information, every little bit, was being relayed to the employees of Mr. Tony Pro.

After this first visit, the e-mails from Mr. Tony Pro changed their tone. Where they once suggested that I might be driving the car too much between visits, they now touted the value of maintaining low-mileage vehicles. He had clearly learned something from my car. He knew that I lived in the city and rarely drove during the week. The information that I had so carefully avoided providing to him in a questionnaire or during an exchange with his employees had been captured from my car's memory and now found a home in Mr. Pro's records. He gained one more chit that day in the economic give-and-take. **■**

*David Alan Grier, an associate professor of international science and technology policy at George Washington University, is the author of the upcoming book, *The Company We Keep*. Contact him at grier@gwu.edu.*

cn Selected CS articles and columns are available for free at <http://ComputingNow.computer.org>.

IEEE Intelligent Systems

THE #1 ARTIFICIAL
INTELLIGENCE
MAGAZINE!

IEEE *Intelligent Systems* delivers the latest peer-reviewed research on all aspects of artificial intelligence, focusing on practical, fielded applications. Contributors include leading experts in

- Intelligent Agents • The Semantic Web
- Natural Language Processing
- Robotics • Machine Learning

Visit us on the Web at
www.computer.org/Intelligent

32 & 16 YEARS AGO

NOVEMBER 1979

HEALTHCARE (p. 4) “While computer technology offers a unique resource for containing costs and improving the quality and availability of care, many government officials, health research planners, concerned citizens, and even computer professionals claim that medical computing has increased costs while offering insufficient proof of improved health care. They are skeptical of the positive contributions information technology can make to cost containment and cite poor performances by computer designers over the last decade, who overestimated the power of the technology and underestimated the complexity of health care systems.”

AMBULATORY PATIENTS (p. 9) “For the past decade the Laboratory of Computer Science of the Massachusetts General Hospital has been involved in the development and implementation of a computer-based medical information system, COSTAR (for computer-stored ambulatory record). Designed to perform the data management functions needed by a group practice in the care of ambulatory patients, COSTAR supplements or replaces the traditional paper-based patient medical record with an integrated information system. COSTAR data management meets the medical, administrative and financial needs of ambulatory group practice, and can be adapted to meet the requirements of other group practices. COSTAR systems are now operational or being installed in over a dozen practices in the United States or abroad.”

HOSPITAL MANAGEMENT (p. 28) “MATRIX ... is an on-line data base management system designed to allow hospitals to develop customized information systems for areas such as the ordering of medications and tests and the updating of patient records. The objectives of MATRIX ... are to improve the quality and lower the cost of patient care. It is designed to provide sufficient flexibility to support applications unique to health care institutions and to evolve with changes in medicine and technology.”

CLINICAL DATA BANKS (p. 34) “... Current data storage and retrieval technology should permit the development of large data banks representing collected clinical experience. These will aid the physician by linking specific treatments to health and medical processes and to patient outcome. ARAMIS, the American Rheumatism Association Medical Information System, is a national data bank for rheumatic diseases, a prototype for such systems. Its software (the Time-Oriented Databank System, or TOD) also supports national data banks in stroke and coma as well as smaller data bases in other specialties.”

ONCOLOGY MANAGEMENT (p. 42) “The Johns Hopkins Oncology Center is one of 21 Comprehensive Cancer Centers

established as part of the National Cancer Plan initiated by the Cancer Act of 1971. The center has major programs in laboratory and clinical research, education, and collaborative activities with community physicians. ... At any one time, there are 1500 patients being treated under one or more of several hundred formally established treatment plans called protocols. This article describes a clinical information system which assists in the management and care of these patients.”



COMPUTING IN CHINA (p. 60) “A question frequently asked is ‘How long will it take the Chinese to catch up with the industrialized nations?’ In view of China’s late arrival in the industrial age, the 10-year educational and managerial gap created by the Cultural Revolution, and the economic handicaps bedeviling the country, I believe the gap between China and the industrialized West could take as long as 30 years to close. China will make tremendous strides, but there is simply so much catching up to do that a significant closing of the gap will not come quickly. Thirty years, however, should give China time to become a formidable rival to the other industrialized nations of the world.”

TELECONFERENCING (p. 62) “... The combination of a CAD system data file and graphic terminal obviously offers yet another means for interpersonal communication. By using an appropriate command language, interactive graphics teleconferencing by computer is possible. When combined with a dial-up voice connection, an audio/graphics teleconferencing arrangement can offer participants both the ability to view the highest quality of drawings and the opportunity to discuss, edit, or exchange them.”

RELIABILITY TOOLS (p. 77) “The gap between DP hardware and software is taking on the dimensions of a canyon as hardware becomes cheaper, faster, and more reliable, and software becomes more expensive, cumbersome, and error-ridden. One of the reasons for the difference lies in the retention, in software design and development, of concepts and methodologies rooted in a dead past. So-called reliability tools are among the worst offenders in this regard.”

Editor: Neville Holmes; neville.holmes@utas.edu.au

32 & 16 YEARS AGO

NOVEMBER 1995

DEPENDABILITY (p. 5) “The implications for the future directions of software are interesting. There will be increased emphasis on fault-tolerant and highly available systems. Diagnosis and delivery of software fixes on line (from your car’s electronic computers to upgrades for your ISDN phone service and modem) will come into vogue. ... New products consisting of ‘middleware’ between the operating system and application software will be created to ensure that applications and data don’t crash or fail in a way that compromises use. Parallel databases, fault tolerant servers and networks, and self-healing computers will become the platforms necessary to deliver software.”

PERFORMANCE TOOLS (p. 21) “The primary motivation for using parallel computer systems is their high performance potential, but that potential is notoriously difficult to realize, and users often must analyze and tune parallel program performance. Parallel systems can be instrumented to provide ample feedback on program behavior, but because of the volume and complexity of the resulting performance data, interpreting these systems can be extremely difficult. Hence, performance tools are needed to help bridge the gap between raw performance data and significant performance improvements.”

PERFORMANCE PREDICTION (p. 47) “Other problems, such as perturbation of the program’s behavior and generation of vast seas of (mostly useless) data that require a performance expert to interpret, make performance analysis a tedious, error-prone, and time-consuming task. Performance prediction tools can significantly expedite this task by providing fast and accurate information to guide the programmer toward efficient data distribution strategies and/or profitable program transformations that will increase performance.”

EVENT TRACING (p. 57) “Just as a logic analyzer lets a hardware designer study signal transitions, software event tracing provides the raw performance data needed to understand all possible spatial and temporal interactions of parallel tasks. However, on parallel systems with hundreds of processors, application instrumentation of procedure calls, message passing, and input/output can quickly generate a large amount of performance data. ...”

WWW SERVERS (p. 68) “To support continued growth, WWW servers must manage a multigigabyte (in some instances a multiterabyte) database of multimedia information while concurrently serving multiple request streams. This places demands on the servers’ underlying

operating systems and file systems that lie far outside today’s normal operating regime. Simply put, WWW servers must become more adaptive and intelligent. The first step on this path is understanding extant access patterns and responses. ...”

NETWORK AS COMPUTER (p. 81) “Having usurped much of the mini and mainframe domain, the PC now faces a serious challenge to its dominance on the desktop computing scene. Larry Ellison, CEO of Oracle Corp. , has referred to the PC as ‘a ridiculous device’ and is not alone in viewing the Internet as a rising star that will push the PC out of the spotlight and into a supporting role. Sun Microsystems, which has long maintained that ‘the network is the computer,’ has thrown down the gauntlet by developing a product that can only accelerate this trend: a network programming language called Java.”

BACKFIRING (p. 87) “The availability of empirical data from projects that use both function-point and lines-of-code metrics has led to a useful technique called ‘backfiring.’ Backfiring is the direct mathematical conversion of LOC data into equivalent function-point data. Because the backfiring equations are bidirectional, they also provide a powerful way of sizing, or predicting, source-code volume for any known programming language or combination of languages.”

A NEW STANDARD (p. 89) “... While software has been established as an integral part of scientific and business disciplines, environments for developing and managing software have proliferated without a common, uniform framework for the software life cycle. This standard provides such a framework, so that software practitioners can ‘speak the same language’ when they create and manage software. Practitioners can use the framework to acquire, supply, develop, operate, and maintain software.”

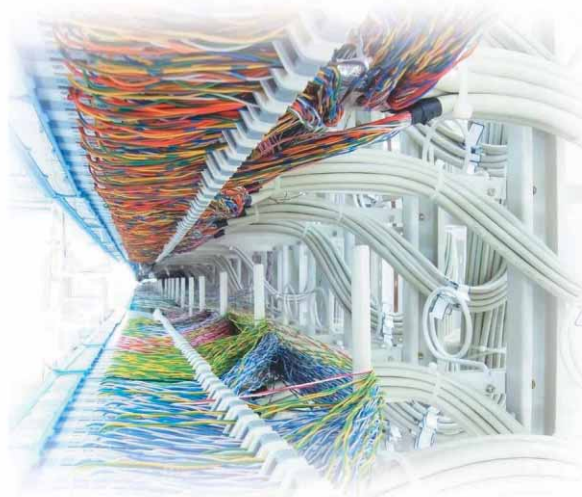
MUSIC (p. 91) “Since 1992, the IEEE Computer Society has supported the establishment of a Technical Committee on Computer-Generated Music. This vast interdisciplinary area of computer science and electrical engineering stretches from artistic music composed or played with computers to audio signal processing. CGM offers new possibilities for research and practice, and as the IEEE CS has argued, this signifies CGM’s greater importance in science and technology—not to mention music itself.”

PDFs of the articles and departments from Computer’s November 1979 and 1995 issues are available through the IEEE Computer Society’s website: www.computer.org/computer.

TECHNOLOGY NEWS

Turning on the Lights for Wireless Communications

Lee Garber



In the ongoing search for better ways to transmit data, researchers have turned to a wireless approach that has been discussed for years: visible light communications.

Researchers in industry and academia are always looking for newer, faster, and more energy-efficient approaches to communications. In their search, they have turned not to a new technology but to one more than 130 years old: visible light communications.

VLC uses modulated light signals rather than radio frequency (RF), microwave, or other types of signals to transmit data.

As the “VLC Projects and Products” sidebar discusses, not only are academic and corporate scientists developing VLC approaches, but several companies have already released products based on the technology. Vendors that have products or are in the process of developing them include Casio, Eurescom, France Telecom, NEC, Orange, Panasonic, Samsung, Sharp, Siemens AG, Telefonica, and Toshiba.

Schools with VLC research programs include the Sapienza University of Rome; Dortmund University of Technology; Ilmenau University of Technology; University of Athens; University of California,

Berkeley; and University of Oxford.

The technology could be utilized for purposes such as Internet access and various types of networking, as well as point-to-point, point-to-multipoint, and multipoint-to-point communications, explained University of Edinburgh professor Harald Haas.

Proponents say VLC offers several important advantages over current wireless systems including higher data rates, better security, more spectrum availability, and greater energy efficiency. They say the technology would be particularly useful in settings such as aviation, green computing, and healthcare facilities. The “Illuminating the Road Ahead” sidebar covers VLC implementation in more depth.

However, VLC also has weaknesses and faces obstacles to commercial success. Unlike Wi-Fi and numerous other wireless technologies, it is not yet proven in large-scale usage. In addition, VLC access points will have to be made affordable.

COMMUNICATING VIA LIGHT

Alexander Graham Bell, the telephone’s inventor, sent the first VLC

transmission on 3 June 1880 via another one of his inventions: the Photophone. In 1931, Bell Telephone Laboratories engineer—and subsequent vice president—Sergius P. Grace proposed using light for secure wireless communications.

However, it was many years before work on VLC began in earnest.

Driving forces

There are several reasons for the recent VLC push. For example, users are always looking for better performance, and VLC promises high data rates. Also, the technology works with LED lights, which are becoming increasingly popular.

Demand for wireless communications is continuing to increase, but the spectrum available for traditional, radio-based mobile approaches is shrinking rapidly.

Using radio communications equipment can be hazardous in various types of locations. For example, it can cause sparks, which makes it dangerous for use on oil platforms. And it can interfere with other radio equipment, which makes it inappropriate for use on aircraft.

TECHNOLOGY NEWS

VLC PROJECTS AND PRODUCTS

There are numerous visible light communications research projects, and VLC products are even beginning to appear, signaling the technology's growing importance.

D-Light. University of Edinburgh professor Harald Haas has been working on VLC since 2004. Haas and his colleagues call their approach both D-Light (data light) and Li-Fi (light fidelity)

Haas expects his technology will be useful for purposes such as Internet access, vehicle-to-vehicle communications, and machine-to-machine communications. He and his team plan to establish a spin-off company—VLC Ltd.—based on their research, which was funded by Scottish Enterprise, a government agency that works to stimulate economic growth in Scotland.

OMEGA. The European Union-funded OMEGA project—an interdisciplinary effort with participants from industry and academia that ran from 2008 to March 2011—researched user-friendly, high-bandwidth, home-area networks. The project explored options such as optical, power-line, infrared, and VLC networking.

OMEGA's VLC demonstration used 16 high-power, ceiling-mounted LED lamps. The researchers aggregated the output of four video players into an Ethernet stream, which they then modulated onto electrical current sent to the lamps. The lights then continuously broadcast four high-definition videos at 100 Mbits per second.

US National Science Foundation. The NSF operates the Smart Lighting Engineering Research Center in partnership with several US universities.

As part of this program, Boston University professor Thomas Little noted that he is working on a project that uses VLC to provide network connectivity and control among distributed light sensors. Little is also researching data-delivery approaches that use both VLC and radio communications.

RONJA. The Reasonable Optical Near Joint Access Free Space Optics device, which Czech Republic-based Twibright Labs developed, uses red light to transmit data up to 1.4 kilometers. It also uses infrared light, which has a wavelength longer than that of visible light, for transmissions up to 0.78 km.

Siemens. The German company is primarily developing high-speed VLC links based on commercial LEDs, focusing on modulation techniques, as well as LED-driver circuitry and analog receivers. Siemens has been working with one of the brightest LEDs available commercially, made by its Osram subsidiary.

Klipsch. This US manufacturer of sound equipment has introduced speakers that can receive music via data transmitted from LED lightbulbs.

The LightSpeaker system combines 10-watt LED lighting and a 20-watt wireless sound speaker into a unit that installs like a standard light bulb. A music source is plugged into a centrally located transmitter, which then sends different music streams to as many as four pairs of speakers.

LVX System. The company has developed an LED lighting fixture that offers an access point to whichever network an individual or organization is using.

The system is installed as part of the 2 x 2 foot lighting panels used in many commercial buildings, although LVX hopes to offer the system in other fixture models soon, noted CEO John Pederson. Customers install a standard networking cable that plugs into a board on the back of the fixture.

The system's current data rate is 0.5 to 3 Mbps at distances up to 10 feet. Pederson said this is adequate in office and airport settings, the company's target markets. A second-generation system, available in a year, will offer higher speeds and programmability, he noted.

LVX System has installed its first few light fixtures in six municipal buildings in St. Cloud, Minnesota.

adapted standard sources such as indoor and outdoor lighting, as well as illuminated signs and displays, street lamps, and vehicle headlights.

System elements. A VLC system's key elements include a visible light source such as an LED lamp, which acts as the communications channel and transmitter; driving circuitry to control the light source; a modulator to get data onto the light stream; and a photodetector to receive the incoming stream and demodulate the information into electrical signals for processing, noted Siemens spokesperson Sebastian Webel.

According to the University of Edinburgh's Haas, the VLC system converts the electrical data signal it receives into a rapidly varying stream of photons. To modulate the signal, the driving circuitry varies the LED's light intensity at multiple levels.

Haas said his team uses "subcarrier index-modulated orthogonal frequency-division multiplexing (SIM-OFDM) to vary the light intensity in a very subtle and distinct fashion." This enables high data rates by transmitting at least 10 bits for each transmission step, as opposed to the one bit that simple on-off switching allows.

The signal is sent to the photodetector, where the light excites electrons and induces an electrical current. The data must undergo an electrical-to-optical-to-electrical conversion because there currently is "no practical means of high-speed, all-optical data processing," Webel said.

Finally, Haas added, an algorithm at the receiving end processes the data stream and converts it into the ones and zeros of digital data.

Light sources. VLC systems can work with either a single light source of a specific color or by combining red, green, and blue light sources to produce the desired color.

Different LED colors have different efficiencies, interference levels, and manufacturing costs, noted Boston University professor Thomas Little.

Doctors are increasingly using wireless technology to transmit data from medical devices to PCs or laptops for analysis. However, there are powerful magnetic fields around devices such as magnetic resonance imaging scanners, so physicians can't use radio-based wireless tech-

nology to transmit MRI data they collect.

Under the hood

VLC uses high-frequency pulsed light instead of radio waves or other types of signals. Many of the implementations work with LEDs in

Most lighting-quality LEDs use blue LEDs with phosphor to create a white color, a relatively inexpensive approach that offers only one transmission channel per data stream. Mixed-color LEDs are more complex but offer multiple transmission channels in a single stream.

Using a bright LED reduces the number of lights necessary to provide the same signal strength, said Webel. However, he added, the challenge is that bright LEDs require high driving currents, which means researchers must develop new driver-circuit architectures.

Increasing throughput. VLC systems can increase throughput by creating multiple communication channels within a single light stream via the use of different colors of light or different modulation frequencies, Webel noted.

For example, SIM-OFDM can convert a serial data stream into many—perhaps 1,000 or more—parallel streams based on frequency.

Using multiple LEDs and photodetectors as transmitters and receivers, respectively, can further increase throughput, he added. This gives users the option of simultaneously sending separate data streams over each LED. In some cases, multiple LEDs are required just to provide the necessary lighting intensity.

Data rates. VLC throughput is increasing. A near-term goal of researchers is to achieve data rates of 100 Mbits per second.

The fastest Wi-Fi version, IEEE 802.11n, offers theoretical-maximum data rates of about 150 Mbps, although using OFDM along with multiple transmitters and receivers could boost the speed to 300 Mbps. Long-Term Evolution (LTE), which some vendors market as fourth-generation cellular technology, can theoretically transmit up to about 300 Mbps.

A VLC system that Germany's Siemens and Heinrich Hertz Institute developed has transmitted data at

ILLUMINATING THE ROAD AHEAD

If and when visible light communications is implemented more widely, it could be used in many settings and for many purposes:

Healthcare. Because, unlike radio communications, VLC doesn't experience problems with magnetic interference, the technology could enable doctors to wirelessly transmit data from magnetic-based medical devices, such as MRIs, to PCs or laptops for analysis.

Hazardous settings. VLC could enable wireless data communications in oil fields and mines, near gas pipelines, and in other environments where using RF equipment—which can create sparks—could be dangerous.

Commercial aviation. Because, unlike radio communications, VLC doesn't interfere with flight-related radio signals, it could enable wireless data communications by passengers on aircraft. Airlines generally don't let passengers use RF-based equipment on planes when in flight. VLC could also let airlines wirelessly offer entertainment and other content to passengers.

Green computing. VLC offers more energy efficiency than radio communications.

Military applications. VLC could enable fast, secure transmissions within vehicles and aircraft.

Underwater communications. RF doesn't work optimally underwater, but VLC functions well in such settings over short distances.

Automobiles. LED stoplights or railroad signals could transmit information to cars or trains. Cars could use LED lights to help occupants communicate with other vehicles, noted Siemens spokesperson Sebastian Webel.

Smart lighting. This approach, designed to create intelligent lighting systems that can be operated in an energy-efficient way, could use VLC as the infrastructure for illumination, control, and communications. VLC would require less wiring and energy than typical smart-lighting systems.

Sensors. VLC could be useful for communications in various types of sensor systems, noted Boston University professor Thomas Little.

Museums. VLC systems could illuminate an object in an exhibit and at the same time wirelessly provide information about it, noted University of Edinburgh professor Harald Haas.

about 500 Mbps over a range of a few dozen centimeters. Earlier experiments demonstrated data rates of 200 Mbps over a distance of 5 meters.

The researchers used a bright LED that they could modulate at a frequency at least twice as high as other LEDs that they tested. Webel said the higher frequency enabled greater data rates, although researchers are unsure why this is the case.

Haas said his research team hopes to achieve data rates up to 1 Gbps in nonlaboratory settings by late 2012.

Advantages

VLC offers several advantages over traditional, radio-based wireless systems.

Security. VLC is generally more secure than traditional RF technologies, according to Haas. First, only receivers within the visible cone of

transmitted light can receive data, making transmissions difficult to intercept.

Also, Haas said, walls and other obstacles can block the high-frequency VLC transmissions, meaning data is unlikely to leak out of an office or home.

Energy efficiency. VLC systems use less power in transmitting data than radio-based communications equipment. VLC doesn't use many of the types of energy-consuming communications equipment that radio-based systems employ, such as antennas and radio circuitry, Haas noted. In fact, he said, VLC's primary power-consuming element consists of the energy-efficient LED lights that are often already in operation.

VLC will also increase the demand for LED lighting, Haas added, which is more efficient than traditional fluorescent or incandescent sources.

TECHNOLOGY NEWS

Wider spectrum, more capacity.

There is much more visible light spectrum available than radio spectrum. The visible light spectrum extends from 400 to 790 THz, while the RF spectrum runs only from 3 KHz to 300 GHz.

Also, visible light doesn't have the shortage of available spectrum that radio-based wireless technologies are experiencing because of the growing use of mobile technologies, Boston University's Little said.

Although not all the visible light spectrum is optimal for communications, it still provides more opportunities than the RF spectrum for sending data without interfering with other communications.

POSSIBLE SHADOWS

Radio technology already works well and meets most wireless users'

needs, Siemens' Webel said. Thus, he explained, VLC proponents will need to either offer better services than RF or find novel applications.

VLC has not been widely used for nearly as long as Wi-Fi and other wireless technologies. This lack of maturity could hurt VLC adoption, at least for a few years.

Another challenge will be developing VLC technologies that don't add much cost to lighting systems, Little said. For potential vendors, the University of Edinburgh's Haas noted, finding investors and pilot-program customers to get the technology moving commercially could be a chore.

A key issue for widespread VLC implementation is making the access points inexpensive enough to compete with other wireless technologies' base stations.

VLC works best when the light intensity is high, which means long-distance transmissions or those that could be at least partially blocked by obstacles or weather conditions could be problematic. This could make VLC less useful for longer-range outdoor applications, or applications that must work between buildings or between multiple offices or rooms within a building.

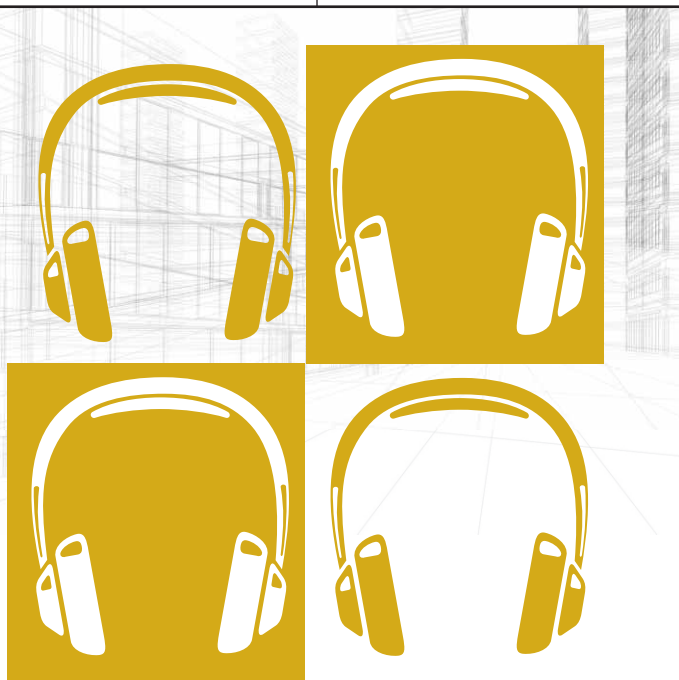
VLC could be subject to regulation as both a communications and a noncommunications technology, involving issues such as eye safety and illumination in traffic signals or car lighting. Proponents thus face the challenge of coordinating implementation across multiple types of standards and regulatory bodies.

Vendors could develop hybrid systems, using VLC when line-of-sight, visibility, and other conditions are favorable, and using RF technology at other times.

Haas predicted VLC will become popular because LED-lighting usage will increase. He noted that numerous countries have already banned incandescent light bulbs because of their energy inefficiency.

One of VLC's first commercial uses might be in applications requiring high-speed data transfer over short distances, such as file sharing, according to Little.

The technology will be popular where security is important and radio communications either are prohibited or face spectrum crowding, he added. **■**



LISTEN TO GRADY BOOCH "On Architecture"

podcast available at

cn <http://computingnow.computer.org>

Editor: Lee Garber, *Computer*;
l.garber@computer.org

cn Selected CS articles and columns
are available for free at [http://
ComputingNow.computer.org](http://ComputingNow.computer.org).

NEWS BRIEFS

System Identifies User Location without GPS or Wi-Fi

A Switzerland-based semiconductor vendor has developed a system for mobile devices that determines a user's location when neither GPS nor Wi-Fi is available. This could be valuable for users inside buildings in crowded urban downtown areas or in mountainous or heavily forested regions.

STMicroelectronics' system starts by taking input from three small, high-performance microelectromechanical systems, each on its own chip: a gyroscope, a compass-like sensor that measures the Earth's magnetic fields, and a sensor that estimates the user's altitude based on air-pressure readings.

Using multiple sensors provides enhanced motion- and location-based capabilities, said Benedetto Vigna, group vice president and general manager of STMicroelectronics' MEMS, Sensors, and High-Performance Analog Division.

The product employs a dedicated processor, along with STMicroelectronics' iNEMO filtering and predictive software engine, to integrate the different types of location-based information. It then utilizes dead reckoning to calculate users' positions in three dimensions (including altitude).

The technology yields information on users' linear acceleration, angular velocity, heading, and altitude. This enables them to identify the direction in which they're heading and their location.

According to STMicroelectronics, the system was designed to be energy-efficient, an important factor for battery-powered smartphones and other mobile devices.

The geomagnetic module—which measures $3 \times 5 \times 1$ mm—offers high-resolution, three-axis sensing of linear and magnetic motion.

The 3-axis digital gyroscope measures $4 \times 4 \times 1$ mm and doesn't

require continuous communication between the sensor and the host processor, which reduces power consumption.

The air-pressure sensor measures $3 \times 3 \times 1$ mm, operates at between 700 meters below and 10,000 m above sea level, and can recognize altitude changes as small as 0.3 m.

Malware Infects US Military Drone System

A virus has infected the system behind American Predator and Reaper military drones.

Initially, US Air Force officials expressed fear the virus could log the keystrokes of the aircraft's remote pilots, creating the possibility that hackers could obtain and sell or otherwise distribute classified information.

However, military cybersecurity specialists now explain, the virus was a common type of malware that steals online gaming logins and passwords. Thus, they say, the malware probably wasn't part of an attack targeting the drones. Instead, it may have worked its way from other systems onto the US Department of Defense (DoD) networks.



Air Force officials noted that the malware infected ground systems that are separate from the drones' flight controls and did not affect their operations. Thus, even after the virus was found, pilots continued to fly the drones, which the US military has used frequently in Afghanistan and Iraq.

The DoD's Host-Based Security System—a COTS-based application designed to monitor, detect, and counter known cyberthreats—found the malware recently.

Officials said that despite numerous efforts, they couldn't remove it from the computers at Creech Air Force Base in Nevada that control the aircraft. Instead, they had had to erase and reformat the drives that the drones' control systems use, a time-consuming process.

The malware incident isn't the drones' first identified security



STMicroelectronics has developed a product that uses three chip-based systems—a gyroscope, a compass-like sensor that measures the Earth's magnetic fields, and a sensor that estimates the user's altitude based on air-pressure readings—to determine a mobile-device user's location.

NEWS BRIEFS

NEW TRANSISTOR COULD LET DEVICES INTERACT DIRECTLY WITH LIVING THINGS

University of Washington scientists have built a transistor that uses protons to send information, potentially allowing the creation of devices that could communicate with living things. Such devices could monitor biological processes and eventually transmit signals that could control various functions.

Researchers are interested in using devices that can work directly with, for example, the human body, to help enable biological sensing or more effective prosthetics.

Electronic devices transmit information using electrons. Living things, on the other hand, use ions, which are positively or negatively charged atoms. Protons are positively charged hydrogen ions.

The challenge is finding a way to translate electronic signals into ionic and protonic ones and vice versa, noted assistant professor Marco Rolandi, the project's lead researcher.

"We found a biomaterial that is very good at conducting protons and allows the potential for interfacing with living systems," Rolandi said.

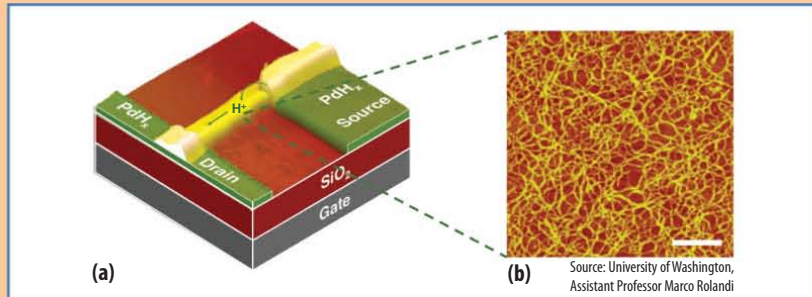
The University of Washington researchers developed a 5-micron-wide field-effect transistor that sends pulses of proton current.

"In our device, large bioinspired molecules can move protons, and a proton current can be switched on and off in a way that's completely analogous to an electronic current in any other field-effect transistor," Rolandi explained.

The device is made with maleic-chitosan, a substance typically obtained commercially from chitin, the structural element in crustaceans' external skeletons. According to the researchers, the material is easily obtained, simple to work with, and compatible with living tissue.

They also note that chitosan absorbs water and forms multiple hydrogen bonds within a transistor over which protons can easily hop.

The current prototype has a silicon base and thus couldn't be placed into a human body. However, the use of a biocompatible base could enable such implantation in the distant future.



University of Washington researchers have developed a protonic field-effect transistor (a) that uses protons to send information, which potentially allows the creation of devices that could communicate with living things. In the transistor, a voltage applied between the proton-transparent palladium hydride (PdHx) source and drain initiates a protonic-current flow along the maleic-chitosan channel, shown in yellow. When hydrated, maleic-chitosan nanofibers (b) form an extended hydrogen bond network along which protons hop. An electrostatic potential applied to the gate electrode turns the protonic current on or off.

problem. For example, many of the aircraft don't encrypt the video they transmit to US forces. In 2009, soldiers found hours of footage shot by drones on the laptops of captured Iraqi insurgents.

Supposedly, the drones' cockpits are not connected to the Internet, which should make them unable to transmit captured keystrokes to a

hacker and leave them immune from transmitted malware. In the past, though, the use of external storage drives has introduced problems to military networks.

Several years ago, experts say a worm infected Predator and Reaper drones via the removable hard drives that load map updates and transport-mission videos from one computer

to another. The DoD has ordered all drone units to stop using the drives.

New Technique Doubles Mobile-Network Throughput

Rice University researchers say they've developed a full-duplex wireless technology that could double network throughput inexpensively without requiring new hardware for devices or networks and without causing service interruptions.

Currently, mobile networks require devices to use different frequencies to send and receive data. Full-duplex technology lets mobile devices send and receive data on the same frequency, effectively doubling a network's capacity.

The Rice scientists—led by Ashutosh Sabharwal, associate professor of electrical and computer engineering—demonstrated that device makers could reliably add full-duplex to existing smartphones and still maintain signal quality.

The researchers added full-duplex as an additional mode to the existing hardware, meaning that device makers wouldn't be required to add new hardware.

Said Sabharwal, "Device makers love this because real estate inside mobile devices is at a premium."

In the past, the concern with using the same frequency to send and receive data was that the dual sets of transmissions would interfere with accurate reception of incoming signals.

The Rice researchers overcame this by repurposing an existing antenna that devices currently utilize for multiple-input, multiple-output technology. MIMO uses several transmitters and receivers, rather than just one of each, to increase wireless throughput. With MIMO, multiple signals that the Rice system transmits cancel each other out, enabling the recipient to accurately receive what's being sent.

Sabharwal said his team will add the full-duplex technology into

its wireless open access research platform, which is available to other scientists.

Although full-duplex wireless technology doesn't necessitate new cell towers, it would require new industry standards.

Thus, the researchers say, it probably won't appear for several years, when carriers begin using fifth-generation cellular technology. Major carriers are just beginning to roll out 4G networks.

The researchers say their work already has attracted the attention of wireless providers worldwide.

The Rice scientists also showed that full-duplex systems could operate in asynchronous mode, meaning that a node could begin to receive one signal while still transmitting another, further increasing throughput.

Scientists Unveil Haptic Pedestrian Navigation System

Japanese researchers have developed a pedestrian navigation system that uses haptics so that users can watch where they're going and not have to look at maps or a navigational device. They say their Hapmap system could be particularly useful for the visually impaired.

Keio University and University of Tokyo scientists developed Hapmap, which provides subtle, complex cues that accurately let users follow a winding path's curves without having to watch the small, battery-powered device.

Typically, pedestrian navigation systems are limited to simple cues such as "walk straight ahead," even if a pathway has many curves.

Hapmap's haptic output component, operated by a servo motor, resembles a small seesaw, which pushes into a user's hand. When the display tilts right or left, it tells the pedestrian to walk in the indicated direction. When the display doesn't tilt, the user walks straight ahead. The researchers say this gives users the

sensation of holding onto a railing that is guiding them along a path.

Hapmap includes a user-tracking system and motion-capture cameras to identify where the pedestrian is and which way a path is turning. This enables the system to automatically control the haptic feedback in real time and offer accurate, detailed navigational information.

In the future, the researchers hope to enable Hapmap use in conjunction with GPS and other navigation systems.

Securing Implanted Medical Devices from Hacking

Academic researchers have developed a system designed to prevent hackers from attacking implantable electronic medical devices such as heart pacemakers.


MIT and University of Massachusetts Amherst scientists say their system would keep hackers from being able to affect an implantable device's operations or steal patient information.

They note that implanted devices such as pacemakers, defibrillators, and insulin pumps increasingly

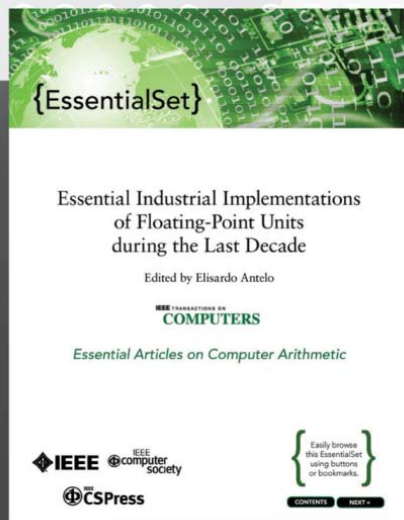
include wireless communication capabilities, used for purposes such as remote monitoring and diagnosis.

Research has shown that it might be possible to exploit the wireless capabilities to send commands to a device or intercept data that it transmits, although no such incidents have been reported.

The scientists developed a transmitter they call a *shield*, which patients could wear around their neck or wrist. The shield relays messages between an implanted device and authorized endpoints. It uses techniques such as signal jamming and encrypted channels to secure the communications and thereby block the interception of messages and the issuance of commands.

The researchers noted that the eventual commercial success of their technology would depend in part on how serious patients consider the threat of attack against their implanted devices. 

Editor: Lee Garber, *Computer*;
l.garber@computer.org



IEEE
CIPress

ESSENTIAL INDUSTRIAL IMPLEMENTATIONS OF FLOATING-POINT UNITS DURING THE LAST DECADE: VOLUMES 1 & 2

Edited by Elisardo Antelo

Surveys the industrial design of floating-point units during the last decade. This EssentialSet is broken into two volumes, sold separately.

PDF edition • \$15 each (\$9 members) • 103 & 79 pp.

Order Online:
computer.org/store



Computer DIGITAL

The IEEE Computer Society's
next-generation flagship publication

More value, more content, more resources

For computing professionals, keeping abreast of the industry's most exciting developments is a continuous process. Beginning in January, *Computer Digital* will offer you even more tools to accomplish that at no risk.

Computer Digital will deliver the same great peer-reviewed articles and columns as the print version of *Computer*. PLUS, it will be:

Mobile—Read the email-delivered issue anytime, anywhere, at your convenience—on your laptop, iPad, or other mobile device.

Searchable—Quickly find the latest information in your fields of interest. Access the digital archives, and save what's most relevant to you.

Linked—Click on table of contents links and instantly go to the articles you want to read first. Article links go

to additional references to deepen your new discoveries.

★ **Engaging**—Absorb the latest high-tech concepts as they come to life through related audio, video, animation, and more. Email authors directly. Even apply for jobs through convenient ad links.

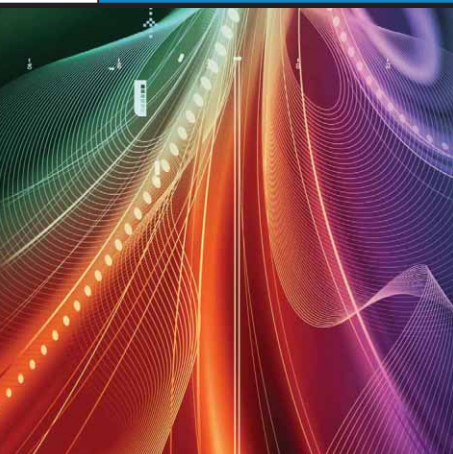
View Demo Now
computer.org/computer-demo

Current digital subscribers—Look for the *Computer Digital* January issue link coming your way in December.

Current print subscribers—Switch from print to digital by 8 December 2011 to receive the January issue. See link below.

Make the switch at computer.org/computerdigital

GUEST EDITORS' INTRODUCTION



Codesign for Systems and Applications: Charting the Path to Exascale Computing

Vladimir Getov, *University of Westminster*

Adolfy Hoisie, *Pacific Northwest National Laboratory*

Harvey J. Wasserman, *Lawrence Berkeley National Laboratory*

The clock speed benefits of Moore's law have ended, and researchers must codesign future exascale HPC systems and applications concurrently in an integrated manner to achieve higher performance under stringent power and reliability constraints.

Computational science has become a vital tool in the 21st century, central to progress at the frontiers of nearly every scientific and engineering discipline, including many areas with significant societal impact. A persistent need for more computing power has provided an impetus for the high-performance computing (HPC) community to embark upon the path to exascale computing.

The challenges associated with achieving efficient, highly effective exascale computing are extraordinary. Past growth in HPC has been driven by performance and has relied on a combination of faster clock speeds and increasingly larger systems. Achieving exascale performance under reliability and power constraints and in the presence of levels of parallelism increased by orders of magnitude will change the path of system and application development.

A recent DARPA study showed that even if it were technically feasible, exascale systems built following the current trajectory would require an energy budget in the hundreds-of-megawatts-per-hour range and reliability estimates that would render them impractical.¹ Thus, the clock speed benefits of Moore's law have ended, and the emphasis must now unavoidably yield to the goal of achieving performance under stringent power and reliability constraints.

EXASCALE COMPUTING CHALLENGES

The issues researchers will encounter on the path to exascale HPC are equally critical for all large-scale computing architectures and facilities, not just the largest ones or only those related to scientific computing. Workloads may differ, but energy challenges are common. Because power is the overriding hardware concern, energy efficiency will be essential across all computing scales. Furthermore, energy issues will affect all levels of the computing system, including processors, interconnects, algorithms, software, and programming models.

Given the complexity of the increasingly daunting constraint space under consideration, successful optimization requires a new tack, a new approach, and a new set of design methodologies. For example, given the overwhelming performance and energy cost of data movement,² efficiency requires minimization of data movement—a task for all layers of the stack, from the hardware to the application software.

Similarly, optimization of the performance/power/reliability triad mandates rethinking of algorithms, programming models, and hardware in concert and requires an unprecedented level of collaboration and cooperation in hardware, system architecture, system software, and application codesign. This requires a completely new approach based on concurrent development and engineering in an integrated manner to a set of consistent overall design metrics, employing accurate, quantitative design methodologies.

THE CODESIGN APPROACH—BACKGROUND

For embedded systems,³ codesign traditionally has meant partitioning concepts in the design process to produce systems meeting stringent performance, verifi-

GUEST EDITORS' INTRODUCTION

cation, and other specifications within a shorter design cycle. The goal and methodology for doing this, as well as the benefits of this approach, have been well-established for many years. The key concept is meeting system-level objectives by exploiting tradeoffs between hardware and software through an integrated concurrent design process. An additional benefit accrues from automation or semiautomation of this concurrent design process, but the crucial part of the definition is concurrency: developing hardware and software at the same time on parallel paths.

What was perhaps left undefined was the precise nature of the interaction between hardware and software. This interaction evolved over the years with increasing use of improved design automation tools, faster application-specific integrated-circuit development tools that allow quick and inexpensive implementation of complex algorithms in silicon, and the use of reduced-instruction-set computing technology that allows the implementation of traditional hardware functionality in software.

The key concept is meeting system-level objectives by exploiting tradeoffs between hardware and software through an integrated concurrent design process.

Codesign in embedded systems came about in large part because a variety of factors led to the use of software in systems that had previously been entirely hardware-based. This increased the complexity of that software in microcontrollers, digital signal processors, and even general-purpose processors. Other factors included the decreasing cost of microcontrollers, rapidly increasing numbers of available transistors, the availability of advanced emulation technology, and the improved efficiency of higher-level language compilers for use in embedded systems. A key motivation was the need to support the growing complexity of embedded systems, which has an obvious parallel in exascale computing.

Embedded systems are characterized by running only a few applications that are completely known at design time, not being programmable by end users, and having fixed runtime requirements—meaning that additional dynamic computing power is not useful.⁴ Codesign considerations for such systems include cost, power consumption, predictability, and meeting time bounds.

In contrast, general-purpose computing systems are characterized by running a broad class of applications, being programmable by end users, and having the characteristic that faster is always better, which requires including cost and peak speed in their design criteria.⁴

The essence of the codesign challenge for HPC and exascale systems is to use the key design criteria of embedded systems—cost and power consumption—while creating systems that are useful and effective over the broad range of applications needed to advance science. “One-off” exascale systems will suggest failure.

CODESIGN FOR HPC SYSTEMS

In the HPC arena, codesign has also been used recently,⁵ and therefore it is not entirely new to exascale computing. Both the IBM BlueGene/L supercomputer (*IBM J. Research and Development*, vol. 49, no. 2/3, 2005) and IBM's PERCS project for DARPA's High-Productivity Computing Systems (HPCS) program⁶ have adopted the codesign approach. Two additional excellent examples of codesigned special-purpose supercomputers for molecular dynamics applications are the MDGrape system⁷ and the Anton supercomputer⁸ built by D.E. Shaw Research.

The IBM RoadRunner⁹ pointed the way toward the use of hybrid architectures by its inclusion of coprocessing elements along with general-purpose processors to accelerate a specific workload. The heterogeneity of the resultant architecture, which required a mixture of several programming models, posed significant challenges to ensure the utility of the coprocessor approach for designing HPC systems that are to be truly effective for a wide range of scientific applications. Still, the metric for success in these codesign examples was performance, without regard to power and reliability.

The HPC community currently finds itself needing to apply codesign methods on the path to exascale systems and applications.¹⁰ Therefore, key concepts that apply include

- employing a high level of abstraction to describe the system;
- using models to allow analysis and exploration of the system architecture, validate assumptions regarding the architecture, explore the design implementation performance parameters, and verify that tradeoffs made using high-level system models were worthwhile; and
- creating codesign methodologies and tools that designers can use to “tinker” with the platform, adding, subtracting, or changing parameters to determine the effect on the architecture and system performance.

A novel exascale concept is related to the necessity of rethinking the application software itself, including optimizing the algorithms and the codes for minimizing data movement for energy efficiency or for implementing resilience mechanisms. Hence, these virtual testbeds need to support initial optimization by both system and appli-

cation designers, before expensive and time-consuming actual implementations are necessary.

IN THIS ISSUE

The articles in this special issue have been selected to cover a cross-section of the codesign space and of the relevant concerns and challenges.

Architectural simulation


“Rethinking Hardware-Software Codesign for Exascale Systems” by John Shalf, Dan Quinlan, and Curtis Janssen describes a set of high-accuracy simulation tools that researchers can employ for low-level hardware and architecture codesign for a simplified application workload.

High-speed interconnects

In “Codesign for InfiniBand Clusters,” Sreeram Potluri and coauthors discuss a codesign approach that takes advanced features from the commodity InfiniBand network, incorporates the design into a state-of-the-art message-passing interface communication library, and then modifies applications to leverage these new features.

High-end systems

“Codesign Challenges for Exascale Systems: Performance, Power, and Reliability” by Darren J. Kerbyson and colleagues describes a comprehensive codesign methodology that uses analytical modeling to achieve maximum performance, power, and reliability for full systems and applications.

In its simplest definition, codesign is about anticipating and changing the future.¹¹ Early intervention in hardware designs, optimizing what is important, influencing the design, redesigning algorithms and system software, devising languages and programming models that reflect abstract machine models, writing code generators and auto-tuners, and modeling all of the above are the essence of the craft. Successfully meeting these challenges is essential for continued progress in computing performance. 

References

1. P.M. Kogge, ed., “ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems,” tech. report TR-2008-13, CSE Dept., Univ. of Notre Dame, 2008; www.cse.nd.edu/Reports/2008/TR-2008-13.pdf.
2. W. Dally, “Power, Programmability, and Granularity: The Challenges of ExaScale Computing,” *Proc. 25th IEEE Int’l Parallel & Distributed Processing Symp.* (IPDPS 11), IEEE, 2011, p. 878.
3. T.A. Henzinger and J. Sifakis, “The Discipline of Embedded Systems Design,” *Computer*, Oct. 2007, pp. 32-40.
4. L. Thiele, “Hardware/Software Codesign,” AS 2011 course materials; www.tik.ee.ethz.ch/tik/education/lectures/hswcd.
5. M. Wolfe, “Compilers and More: Hardware/Software Codesign,” *HPCWire*, 2 Nov. 2010; www.hpcwire.com/hpcwire/2010-11-02/compilers_and_more_hardware_software_codesign.html?layout=print.
6. R. Rajamony, L.B. Arimilli, and K. Gildea, “PERCS: The IBM POWER7-IH High-Performance Computing System,” *IBM J. Research and Development*, vol. 55, no. 3, 2011, pp. 3:1-3:12.
7. T. Narumi et al., “A 55 TFlops Simulation of Amyloid-Forming Peptides from Yeast Prion Sup35 with the Special-Purpose Computer System MDGrape-3,” *Proc. Conf. High-Performance Computing, Networking, Storage and Analysis* (SC 06), IEEE, 2006; doi:10.1145/1188455.1188506.
8. D.E. Shaw et al., “Millisecond-Scale Molecular Dynamics Simulations on Anton,” *Proc. Conf. High-Performance Computing, Networking, Storage and Analysis* (SC 09), ACM, 2009; pp. 39:1-39:11.
9. K.J. Barker et al., “Entering the Petaflop Era: The Architecture and Performance of Roadrunner,” *Proc. Conf. High-Performance Computing, Networking, Storage and Analysis* (SC 08), IEEE, 2008, pp. 1:1-1:11.
10. D. Maliniak, “Hardware/Software Co-Design Comes of Age,” *Electronic Design*, 10 July 2008; <http://electronicdesign.com/print/embedded/hardware-software-codesign-comes-of-age19301.aspx>.
11. K. Yelick, “Software and Algorithms for Exascale: Ten Ways to Waste an Exascale Computer,” invited talk, Oil and Gas High Performance Computing Workshop, Rice University, Houston, Texas, Mar. 2011; www.og-hpc.org/Rice2011/Workshop-Presentations/OG-HPC%20PDF-4-WEB/Yelick%20Exascale-SW-10ways-OilGas.pdf.

Vladimir Getov is a professor of distributed and high-performance computing at the University of Westminster, London. His research interests include parallel architectures and performance, autonomous distributed computing, and high-performance programming environments. He is a member of IEEE and ACM, a Fellow of the BCS, and Computer’s area editor for high-performance computing. Contact him at v.s.getov@westminster.ac.uk.

Adolfy Hoisie is a Laboratory Fellow and director of the Center for Advanced Architectures at the Pacific Northwest National Laboratory. His research focuses on performance analysis and modeling of systems and applications, areas in which he has published extensively. Hoisie is a member of IEEE. Contact him at adolfy.hoisie@pnnl.gov.

Harvey J. Wasserman is a member of the User Services Group at the National Energy Research Scientific Computing Center, the primary computing center for the US Department of Energy’s Office of Science, located at Lawrence Berkeley National Laboratory. Wasserman’s research focuses on workload characterization, benchmarking, and system evaluation. Contact him at hjwasserman@lbl.gov.

For videos related to this topic, see the following:

- “ASCR Discovery—Codesigning for Exascale,” www.youtube.com/watch?v=gXvh7WezxgE
- “Reconfigurable Exascale Computing,” www.youtube.com/watch?v=zTr4bepr5Xc

COVER FEATURE



Rethinking Hardware-Software Codesign for Exascale Systems

John Shalf, *Lawrence Berkeley National Laboratory*
Dan Quinlan, *Lawrence Livermore National Laboratory*
Curtis Janssen, *Sandia National Laboratories*

The US Department of Energy's exascale computing initiative has identified hardware-software codesign as a central strategy in achieving more agile hardware development. Hardware simulation and code analysis tools that facilitate deeper collaboration between hardware architects and application teams will be an essential component of the codesign process.

Over the years, users of high-performance computing (HPC) systems have seen dramatic increases in peak performance claims but with no commensurate improvement in application performance. Moreover, as these systems approach the exaflops scale, their designers are confronting enormous electrical power requirements. The cost of power is expected to exceed the procurement costs of HPC systems, and will thus ultimately limit their future practicality.

Unless system designers work aggressively with vendors and the scientific community to develop more energy-efficient solutions, these trends will lead to an HPC industry crisis.^{1,2} Traditional HPC system design methodologies have not had to account for power constraints or parallelism on the level designers must contemplate for exascale systems. Furthermore, these tectonic shifts in computer architecture are anticipated to radically change the programming model and software environment at all scales in future computing systems. The designers of HPC hardware and software components have an urgent need for a sys-

tematic methodology that reflects future design concerns and constraints.

The US Department of Energy's (DoE's) exascale computing initiative has identified hardware-software codesign as a central strategy to meet this need. The idea is to have a novel development partnership in which application scientists participate in a highly collaborative and iterative design process well before the vendor brings the system to market. Accelerating this cycle could improve energy efficiency and usable performance by orders of magnitude.

Through a joint project involving our three laboratories, we are assembling Codesign for Exascale (CoDEX), a comprehensive hardware-software codesign environment that will enable an unprecedented opportunity for application and algorithm developers to influence the direction of future architectures so that they meet DoE mission needs (<https://sites.google.com/a/lbl.gov/codex>). CoDEX combines three elements:

- highly configurable, cycle-accurate simulation of node architectures from Lawrence Berkeley National Laboratory's Green Flash project (www.lbl.gov/cs/html/greenflash.html);
- novel automatic extraction and exascale extrapolation of memory and interconnect traces using the Lawrence Livermore National Laboratory's ROSE compiler framework (www.rosecompiler.org); and
- scalable simulation of massive interconnection networks using Sandia Laboratories' Structural Simulation Toolkit (SST)/macro, a coarse-grained simulator (<http://sst.sandia.gov>).

These tools will enable a tightly coupled software-hardware codesign process that is highly suitable for complex HPC designs.

WHY CODESIGN AND SIMULATION TOOLS ARE KEY

HPC system development requires a more tightly integrated approach to designing future computing systems to meet the extreme power and performance challenges for the next decade. Codesign is an iterative process, in which the goal is to optimize both applications (algorithms and algorithmic implementations) and hardware for some combination of performance, power, and cost. An essential first step is to identify the application targets, which makes it possible to use target applications' performance on the target machine as a key success criterion—one in which hardware architects and applications teams are equally invested.


An effective codesign process also requires accurately predicting the performance, cost, and power consequence of any design tradeoff in algorithms or hardware configuration. The architectural simulation of computing systems plays an essential role in providing this predictive capability because it enables the quantitative exploration of HPC system design tradeoffs. HPC system development requires a more tightly integrated approach to designing future computing systems to meet the extreme power and performance challenges for the next decade.

The embedded processor market has refined codesign processes over the past 20 years to meet the demanding cost and power efficiency requirements of battery-powered consumer electronics (smartphones, MP3 players, and so on) as well as of power-sensitive high-performance embedded applications like avionics systems. Codesign's success is due largely to a continued focus on developing tools that make the methodology productive, cost-effective, and beneficial, such as automated processor synthesis tools, cycle-accurate simulators, and the automated generation of software tools (compilers and debuggers) from hardware specifications. These tools have broken through the slow pace of system design that held back the embedded industry for many years and that continues to impede critical advances in HPC design and implementation.

The partnership in DoE's vision furthers codesign practice by anticipating application needs. The multidisciplinary codesign team begins the process by identifying leading-edge, high-impact scientific applications and then using these as optimization targets instead of the traditional speed and feed targets (flops and bandwidth) and percentage of peak flops. The design process is application driven. Thus, rather than looking at what scientific applications can run on an exascale system after it ar-

rives, the team seeks to uncover what a system needs to meet the requirements of critical science problems. There was no thunderclap or equivalent to a sonic boom when we crossed the threshold to petaflops, nor will there be for the first system to exceed the "exaflops barrier." The success of the DoE exascale program hinges on the success of the science applications. As such, the process builds on DoE's understanding of application requirements and expands its already broad-based computational science portfolio.

In contrast to a traditional pipelined process, which optimizes isolated team-specific metrics, the codesign process requires collaboration and compromise from all teams to achieve a common optimization target. Application experts work in concert with computer architects and algorithm designers—all using sophisticated tools to accelerate the design loop to meet the common goal of maximizing scientific application performance



Codesign is an iterative process, in which the goal is to optimize both applications and hardware for some combination of performance, power, and cost.

within the constraints of power and cost. The methodology depends on a bidirectional optimization of design parameters, in which software requirements drive hardware design decisions, and hardware design constraints motivate changes in the software design to better fit within those constraints.

Thus, a deep analysis of application requirements drives key design decisions for the overall system architecture. It provides quantitative measures of application requirements and relates them to architectural parameters such as on-chip memory, memory bandwidth demands, and interconnection requirements.

Unfortunately, many hardware design choices are too costly to implement or do not improve energy efficiency enough to justify the cost—a condition that motivates software implementation changes. Architectural simulation provides an avenue to experiment with hardware configurations, programming models, and algorithms on exascale class machines before making the full system implementation available. Designers can integrate models of the hardware, runtime system, and application into the simulation to evaluate the tradeoffs among the design choices imposed on each aspect of the integrated system. Cycle-accurate simulation tools let designers quantify the performance impact on applications in light of specific hardware constraints.

COVER FEATURE

Table 1. Application surrogates and their definitions.

Surrogate	Description
Compact application	Small app with fewer features and simplified boundary conditions relative to a full app
Mini-application	Small, self-contained program that embodies essential performance characteristics of key apps
Skeleton application	Program that captures an app's control flow and communication pattern; can be run only in a simulator
Proxy application	General term for all other surrogates
Mini-driver	Small programs that act as drivers of performance-impacting library packages
Kernel	Program that captures an algorithm's node-level aspects

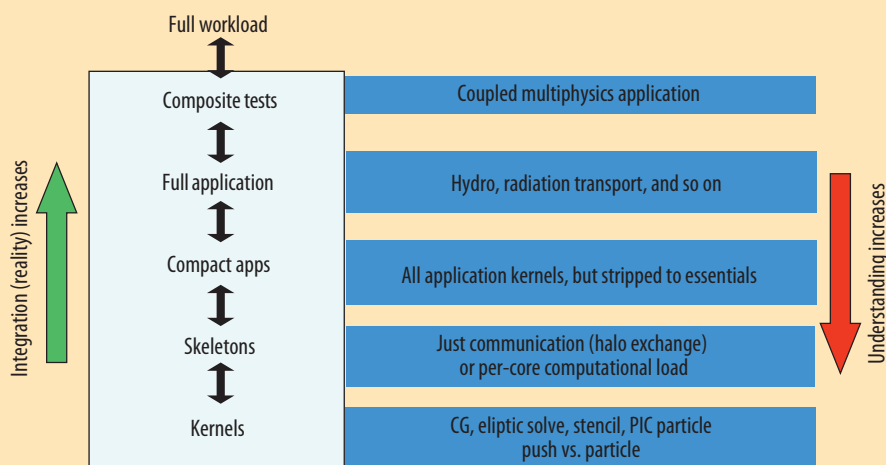


Figure 1. Interplay among the layers in the application hierarchy. In the CoDEx codesign process, kernels enable the rapid exploration of new languages and algorithms because of ease of rewriting, whereas full applications are harder to rewrite but ensure adherence to original application requirements.

Analysis of the hardware constraints in terms of cost, area, and power consumption serves as feedback to motivate changes in the application and algorithm design to better fit within hardware constraints.

CODESIGN TOOLS

The codesign process requires both hardware and software models. The hardware models capture the structure and function of the exascale system's hardware components. The software models consist of a variety of surrogates for the application software. Different surrogate types allow designers to implement different aspects of the codesign process. Table 1 lists several surrogates and their definitions. Figure 1 shows how the application reduction hierarchy relates to code analysis—as applications reduce to the kernel surrogate, understanding increases, and as the design moves up the hierarchy, its gets closer to reality, or the full application.

Figure 2 shows the codesign workflow using these different application representations. Compact mini-applications and mini-drivers and kernels run on actual hardware; skeleton applications run only in a simulated hardware environment. Tools that allow automated conversion among these surrogates are a key feature of the CoDEx environment.

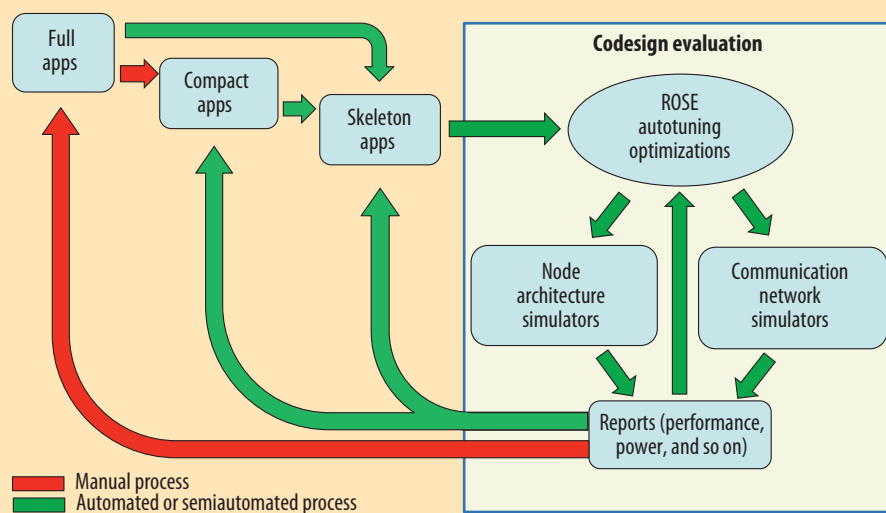


Figure 2. Hierarchy of code representations. The CoDEx codesign process uses a hierarchy of simplified surrogate code representations to provide hardware designers with actionable detailed information while still ensuring that the context for any insight remains faithful to the full application's requirements.

Deep code analysis

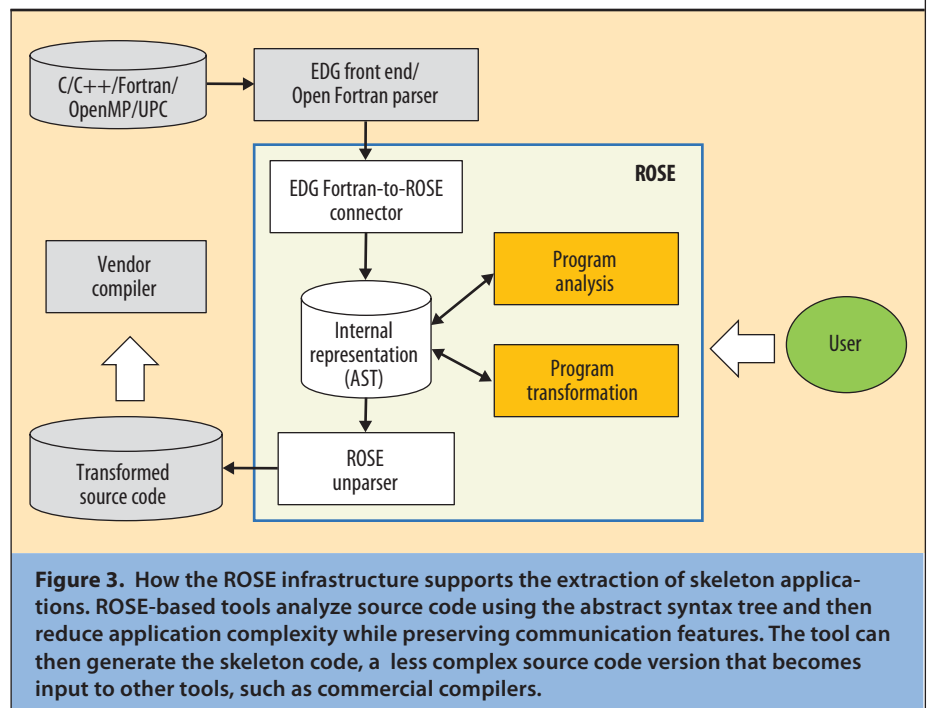
In our codesign process, the ROSE infrastructure forms a central part of the tool chain for automating the extraction of skeleton applications to support input to the simulators. As Figure 3 shows, the chain starts with the ROSE-based tool reading source code into an abstract syntax tree using an intermediate representation. The tool then analyzes the source code indirectly using the AST and subsequently transforms the AST to cut away parts of the application's representation to reduce complexity but preserve specific features, such as message-passing interface (MPI) communication. From the modified AST, the tool can generate a less complex source code version, which becomes the skeleton for input into hardware simulators.

Research to provide a deep understanding of code performance depends on significant internal compiler analysis. Current work uses data dependence and control-flow information (def-use analysis) to support the generation of backward slices rooted in the application's MPI calls. Future work will expand on def-use analysis and use a more general dataflow framework in ROSE to support the generation of skeletons that capture a range of behaviors—not just MPI communication. Future work will also incorporate interprocedural analysis based on the system dependence graph to track the use of variables across function interfaces and provide more aggressive simplification levels for large-scale applications.

These techniques to automate generation of skeleton applications are central to how codesign can incorporate current scientific software into the process of evaluating modern numerical techniques for future exascale architectures. Research to support this work is an ongoing collaborative effort of the ROSE project and Galois Inc., with iterative results available on the ROSE project website (www.rosecompiler.org).

Coarse-grained interconnection simulation

Coarse-grained simulations let designers study large-scale systems in a way that captures the complex interactions among hardware components, including interactions that arise only at the largest scales. The SST/macro³ and Warwick Performance Prediction simulator⁴ are examples of simulators designed specifically for studying exascale class machines. Our codesign process uses



the SST/macro simulator. The simulator software is modular, permitting the study of various computational and communication models. SST/macro is distributed under an open source license and is downloadable from <http://sst.sandia.gov>.

Designers can use the simulator in two ways. One is to replay a trace of a previously run MPI application through the simulator, allowing its execution time to be estimated on new hardware. The application mimics a real application's control-flow and messaging pattern except with values from performance models instead of the computational and message-passing costs.

Alternatively, the simulator can use a skeleton application^{5,6} that domain experts provide or ROSE-based analysis tools generate. Traditionally, an application analyst must manually extract skeleton applications from full applications by removing all noncommunication code yet preserving all the code necessary to compute the communication topology and message sizes. This task is tedious and labor intensive. Because communication skeletons are central to modeling communication performance, the CoDex project team has developed technology that enables the ROSE compiler to automate a skeleton's extraction.

Both ways of using the simulator have unique benefits. The tracing approach provides an easy way to understand existing application performance on new architectures. The skeleton application approach allows studies at extreme scales and permits the evaluation of new programming model ideas and notional algorithms without a complete implementation.

COVER FEATURE

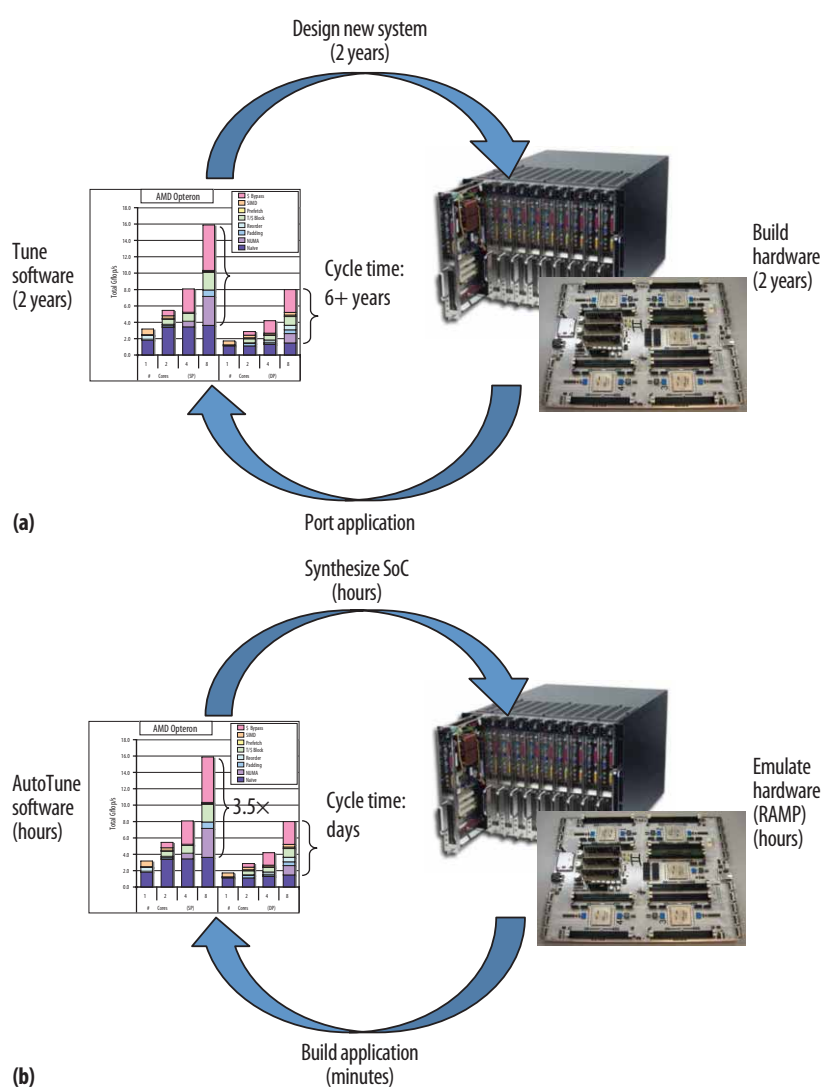


Figure 4. Comparison of the CoDEx codesign cycle and a conventional design cycle. (a) Conventional design cycles last four to six years. In contrast, (b) the CoDEx codesign process features rapid synthesis tools to generate prototype designs, FPGA-accelerated emulation, and software autotuning, which could reduce that time to one or two days.

Rapid design synthesis

The embedded processor market has refined codesign processes by developing tools that make hardware-software codesign productive, cost-effective, and beneficial. Our codesign approach extends many of the tools designed for the rapid synthesis of embedded designs so that they are suitable for exploring HPC design alternatives.

One of our key extensions is to the Tensilica Xtensa Processor Generator (XPG) tool chain to work as our rapid prototyping platform for the HPC simulation. XPG provides an end-to-end solution for quickly creating simple, semi-custom processors from optimized, power-efficient building blocks. The XPG's customizable instruction set,

communication interfaces, and memory hierarchy make it ideal for exploring novel chip multi-processor designs. Its ability to extend the instruction set to add application-specific functionality produces a streamlined processor with scratchpad memories, advanced communication features, and custom operational codes that facilitate advanced communication and synchronization. XPG's ability to automatically generate C/C++ compilers, debuggers, and functional models enables fast software porting and rapid testing with a new architecture. It also speeds language development by enabling researchers to target source-to-source translation technology (targeting the compilers that XPG automatically generates), which is substantially easier than crafting a new compiler from scratch.

Accelerated hardware emulation

Design synthesis tools such as XPG produce gate-level Register Transfer Language, a list of gates with the circuit connections between them. Designers can use gate-level RTL as a chip-design layout, which in turn can be mass-produced at a chip fabrication facility. An alternative is to load the gate-level RTL for a potential processor design onto field-programmable gate arrays (FPGAs) for full cycle-accurate emulation. Architectural emulation platforms based on FPGAs, such as Palladium,

are a common way to support the rapid prototyping and evaluation of application-specific IC designs. Once the design synthesis tool maps gateware components onto the FPGA's programmable logic blocks, the resulting system looks for all practical purposes like the actual hardware, except that it runs at a much lower clock frequency (but orders of magnitude faster than software simulation). The emulated system can boot conventional operating systems such as Linux and use the actual compilers targeted for the final development platform.

FPGA emulation is hundreds to thousands of times faster than typical software emulation environments, yet its dramatic speed increase does not sacrifice accuracy. On

the contrary, FPGA emulation is arguably more accurate than a software simulation environment, since it truly represents the hardware design. The direct mapping to FPGAs on the hardware emulation platform and copious performance data provide a fast, accurate performance emulation environment. In this environment, designers can benchmark real codes, thus ensuring that application developers are intimately involved in the codesign process.

Gate-level RTL design also provides accurate feedback on the design's power consumption because it reflects the actual gates that would produce the real chip design. In contrast, power estimation is much more difficult to verify in a software-only simulation environment.

Our codesign process adopts the Research Accelerator for Multiple Processors (RAMP) technology, the result of a cooperative effort among six universities to build a novel emulation system for parallel processors (<http://ramp.eecs.berkeley.edu>). RAMP emulation technology uses the Berkeley Emulation Engine 3 (BEE3) boards, which in turn use Xilinx FPGAs. Each board has inter-FPGA links on the same board and between boards. Consequently, designers can scale up to simulate a single many-core socket, or even larger-scale clusters. The Berkeley RAMP Blue project demonstrated emulation of more than 1,000 cores using a stack of 16 BEE2 boards.⁷

With FPGA-accelerated emulation capabilities, a designer can generate a new system design (tape out) every day; system build time is in minutes instead of years. Designers can also evaluate alternative design scenarios in minutes—which used to take months or even years with conventional development processes.

Overall, the performance and increased confidence in the timing model of FPGA-accelerated hardware emulation platforms enhance interaction at all levels of system development. As Figure 4 shows, these capabilities can facilitate and ultimately accelerate the iterative process of science-driven system design.

CODESIGN STRATEGY

Codesign depends on a strategy that involves integrated design teams as well as sophisticated tools. Application experts work in concert with computer architects and algorithm designers to accelerate the design loop. The optimization target for the process is the delivered application performance for the combined hardware and software environment, rather than tangentially related metrics such as peak flops or byte-to-flops ratios. Therefore, it is essential to clearly identify the application up front that will serve as the common measure of success for all aspects of the system design.

The iterative design methodology in Figure 4 depends on a bidirectional optimization of design parameters: deep analysis of application requirements drives key design decisions for the overall system architecture, while ar-

MULTIRESOLUTION MODELING

Modeling, simulation, and compiler analysis all play synergistic roles in the codesign process to cover a broad space of design parameters. As Figure A shows, a multiresolution approach using multiple modeling methodologies is essential to cover the extreme scale and provide enough fidelity to inspire confidence in the design choice.

Tools such as cycle-accurate hardware simulation offer extreme detail in their modeling capability, but limit system scale to node size or small clusters. Software simulators, such as SST/macro, can expand the system scale that a model can handle, but they must neglect some detail to achieve that scalability. Likewise, constitutive models and other empirical modeling methods can cover much larger systems, but by definition can model the effects only as parameters. Most modeling is through empirical models because they are faster to construct and evaluate, but software-based and cycle-accurate models are also necessary because they can verify that the simpler model has included all the important effects and has not neglected anything essential but unanticipated.

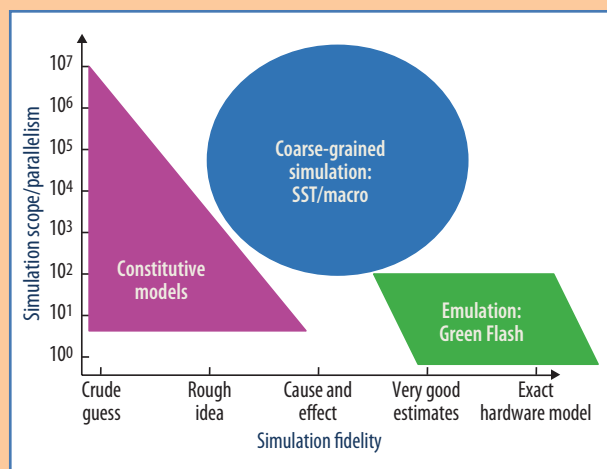


Figure A. Multiple modeling approaches cover both the scale and accuracy required to understand system design tradeoffs. The two key axes for simulation and modeling techniques are model fidelity (horizontal axis) and the scale of the system that simulation can handle.

chitectural simulation provides quantitative feedback on potential design tradeoffs for both hardware and software implementations. This tightly integrated design optimization cycle will enable orders-of-magnitude improvement in energy efficiency and usable performance.

RESULTS

We have applied our codesign process in several projects to develop climate modeling and seismic imaging applications, as well as to automatically cotune hardware and software in the same process.⁸

Our first project was Green Flash, the codesign of a machine for kilometer-scale climate modeling.⁹ The three central features of the Green Flash design process were fast

COVER FEATURE

Table 2. Comparison of energy efficiency in the Intel Nehalem, Nvidia Fermi, and Green Wave, a codesigned many-core processor implementation.

Architecture	Intel Nehalem	Nvidia Fermi	Green Wave
Total nodes	127,740	66,823	75,968
Megapoints per watt	4.27	6.28	32.63
Communication overhead (%)	9.50	43.00	16.00
Total megawatts	38.20	26.10	5.00

FPGA-based hardware emulation, rapid design synthesis tools, and software autotuning technology.

Green Flash benefited greatly from FPGA-based emulation to facilitate tightly coupled codesign processes. These processes in turn enabled the rapid development of an application-driven many-core chip design targeted at scientific applications. Finally, as the “Multiresolution Modeling” sidebar describes, the project demonstrated the effectiveness of using a multidisciplinary hardware-software codesign process that facilitates close interactions among application and computer scientists and hardware engineers in developing a system tailored to scientific computing requirements.

Another project, Green Wave, conducted in collaboration with the Fraunhofer Institute for Industrial and Applied Mathematics (www.itwm.fraunhofer.de), extended the codesign process to study reverse-time-migration problems in seismic imaging—among the most demanding commercial applications for large-scale HPC clusters.¹⁰ The team modeled the full node design, including the memory subsystem, as well as the interconnection design's performance. Green Wave demonstrated how a hardware-software codesign process can achieve an order-of-magnitude increase in energy efficiency over comparable GPU- and CPU-based systems.

Table 2 shows how the codesigned Green Wave implementation compares with the Intel Nehalem and Nvidia Fermi designs for the seismic imaging survey. To ensure a fair comparison, the Green Wave design used a 45-nm process technology, the same memory technology as the Nehalem, and limited chip area to 240 mm², about the same as the die size for the Nehalem and about half the die size for the Fermi.

The project team optimized Green Wave using a codesign process for a typical seismic imaging survey. The survey, which must complete in a week, requires approximately 12,000 separate data analysis runs performed on a 30 × 20 × 10 megapoint computational grid. The first row of Table 2 shows the system size required to meet this computational rate; the last row indicates how many megawatts a system of that size would consume. The metric of value for these simulations is the number of megapoints per second—the number of grid points that the HPC cluster can process per second for this seismic survey. The second row of the table translates this into an energy efficiency metric of megapoints per watt of power consumed. The third row indicates how much of the computational time is spent on inter-processor communication using the same InfiniBand interconnection technology. (We collected detailed communication timing information from an instrumented MPI implementation.)

As the table indicates, relative to a conventional HPC design, the codesigned system offers 5 to 7.6 times higher performance given similar constraints on chip area, lithographic scale, and memory and interconnection technology. We believe expanding the codesign process to these components could yield another order-of-magnitude improvement, since the Green Wave processor performance optimizations shift the primary sources of energy consumption to the memory and interconnection components.

Our experience using architectural simulators reflects six years of prototyping application-tailored chip designs for the Green Flash (kilometer scale climate modeling) and Green Wave (high-resolution seismic imaging) projects. We have generalized the tools and methodologies developed to accelerate the codesign process for those projects, and these generalizations provide many of the hardware simulation capabilities for the CoDEX project.

Both the Green Flash and Green Wave codesign projects relied on rapid synthesis to prototype new interprocessor communication and synchronization services to support advanced execution models. For example, we synthesized direct word-granularity interprocessor communication services to support fine-grained synchronization primitives

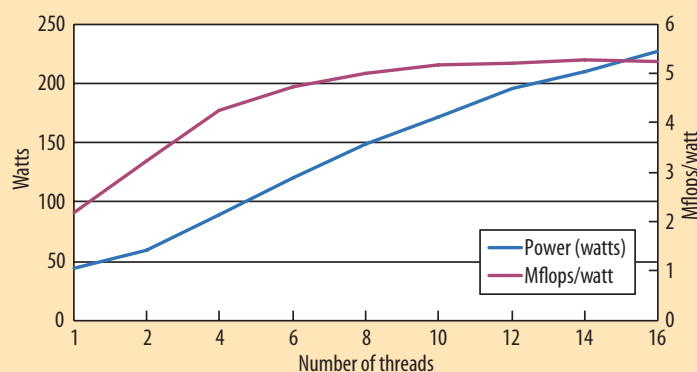


Figure 5. OpenMP transformations using ROSE. The evaluation used a simulator on a numerical kernel (Jacobi iteration), evaluating both power and Mflops per watt for various numbers of threads.

for noncache-coherent global-address-space languages. The more flexible and generalized CoDEX environment can automatically expose this new capability in the compiler to make it immediately accessible to our software stack developers for experimentation. These automated rapid prototyping tools will enable us to make substantial progress in developing novel hardware support for locality management, resilience, security, and interprocessor communication on a limited budget.

Advanced compiler environments also play a significant role in our architectural analysis. For example, we use ROSE to analyze sample applications and extract the MPI use as an input code to support the evaluation of message passing, which in turn supports the SST/macro simulator. This work supports the CoDEX project's charter to automate the application of codesign to support the evaluation of DoE applications.

The simulator represents a range of proposed future node architectures, with ROSE automating the complex code transformations needed to evaluate the potential of each design point. Figure 5 shows both power and Mflops per watt for different numbers of threads using a node simulator. Automating the exploration of both hardware and software design alternatives is essential in the goal to accelerate the iterative codesign cycle.


Codesign for exascale systems provides research challenges along all axes of the hardware, system software, and application design space. Our codesign process features hardware simulation and compiler-driven code analysis for model development, which allows designers to consider complete application codes rather than proxy kernels. Deep code analysis using compiler-assisted model development, interconnection simulation, and both software and FPGA-based hardware emulation systems with rapid design synthesis tools plays a central role. The cotuning environment enables the rapid search of many design alternatives to find the optimal hardware-software solution. Such an extensive search would not be practical or even possible with traditional decoupled design processes.

We are rapidly evolving the architectural simulation framework we have brought together in the CoDEX project to improve usability, increase integration with other simulators, model faults and power consumption, support alternative programming models, and quantify the uncertainty in our simulation results. These efforts will provide a platform for ROSE-based tools to analyze ever more complex application programs. Above all, we are placing a heavy emphasis on validation to ensure that our simulation results provide an accurate basis for exascale system design.

To simulate failure modes, we are expanding our ability to automatically inject faults into any system design com-

ponent. Designers can use this feature to harden software infrastructure in the face of transient errors or to test new models for fault detection, resilience, and recovery. Similarly, they can use SST/macro to insert simulated hardware failures into the interconnection model to test detection and correction methods for the interconnection protocol.

Our hardware simulation environment enables us to better understand how to build a programming environment together with the hardware that can exploit new features. For example, we are planning to study language and hardware design for invasive uncertainty quantification, in which probability density functions represent floating-point values. Another direction is to deepen the understanding of the programming involved in integrating NVRAM directly into the memory hierarchy—a likely trend for future exascale system designs. Hardware simulation enables us to explore the consequences and benefits of adding more intelligence to the memory chips to perform operations on data in situ within the memory subsystem.

In short, it is an exciting time to be involved in computer architecture, and the future of codesign for exascale computing is bright. We envision processes that unite multidisciplinary teams of application scientists, mathematicians, and computer architects in creating machines that can tackle problems critical to societies worldwide. 

Acknowledgments

The work described in this article was supported by the US Department of Energy's Office of Advanced Scientific Computing Research. Lawrence Berkeley National Laboratory is supported by the DoE Office of Advanced Scientific Computing Research under contract DE-AC02-05CH11231. Lawrence Livermore National Laboratory is supported by the DoE Office of Advanced Scientific Computing Research under contract DE-AC52-07NA27344. Sandia National Laboratories is a multiprogram laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the DoE's National Nuclear Security Administration under contract DE-AC04-94AL85000. We also acknowledge Chris Rowen of Tensilica and Martin Deneroff for insightful guidance and support.

References

1. P. Kogge et al., "Exascale Computing Study: Technology Challenges in Achieving Exascale Systems," IPTO tech. report TR-2008-13, 2008, DARPA; www.cse.nd.edu/Reports/2008/TR-2008-13.pdf.
2. K. Asanovic et al., "The Landscape of Parallel Computing Research: A View from Berkeley," tech. report UCB/EECS-2006-183, 2006, EECS Dept., UC Berkeley; www.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-183.pdf.
3. C.L. Janssen et al., "A Simulator for Large-Scale Parallel Architectures," *Int'l J. Parallel and Distributed Systems*, vol. 1, no. 2, 2010, pp. 57-73.
4. S.D. Hammond et al., "WARPP: A Toolkit for Simulating High-Performance Parallel Scientific Codes," *Proc. 2nd Int'l Conf. Simulation Tools and Techniques (Simutools 09)*, Inst.

COVER FEATURE

- for Computer Sciences, Social-Informatics, and Telecommunications Eng., 2009, pp. 1-19.
5. V.S. Adve et al., "Compiler-Optimized Simulation of Large-Scale Applications on High-Performance Architectures," *J. Parallel and Distributed Computing*, vol. 62, no. 3, 2002, pp. 393-426.
 6. R. Susukita et al., "Performance Prediction of Large-Scale Parallel Systems and Applications Using Macro-Level Simulation," *Proc. Int'l Conf. High-Performance Computing, Networking, Storage, and Analysis (SC 08)*, ACM, 2008, pp. 1-9.
 7. A. Krasnov et al., "RAMP BLUE: A Message-Passing Many-core System in FPGAs," *Proc. Int'l Conf. Field Programmable Logic and Applications (FPL 07)*, 2007; <http://ramp.eecs.berkeley.edu/Publications/RAMP%20Blue%20FPL%202007.pdf>.
 8. M. Mohiyuddin et al., "A Design Methodology for Domain-Optimized Power-Efficient Supercomputing," *Proc. Int'l Conf. High-Performance Computing, Networking, Storage, and Analysis (SC 09)*, ACM, 2009, pp. 1-12.
 9. D. Donofrio et al., "Energy-Efficient Computing for Extreme-Scale Science," *Computer*, Nov. 2009, pp. 62-71, 2009.
 10. J. Krueger et al., "Hardware-Software Codesign for Energy-Efficient Seismic Modeling," *Proc. Int'l Conf. High-Performance Computing, Networking, Storage, and Analysis (SC 11)*, ACM, to be published, 2011.

John Shalf is a staff computer scientist at Lawrence Berkeley National Laboratory and leads the Advanced Technology Group at the National Energy Research Supercomputing Center. His research interests include computer architecture, programming models, and frameworks for large-scale scientific application development. Shalf received an MS in electrical and computer engineering from Virginia Tech. He is a member of the IEEE Computer Society and the American Association for the Advancement of Science. Contact him at jshalf@lbl.gov.

Dan Quinlan is the leader of the ROSE project at Lawrence Livermore National Laboratory's Center for Advanced Scientific Computing. His research interests include program analysis, source-to-source compiler construction for general-purpose languages, programming models, parallel algorithms, and many-core and cache-based optimizations. Quinlan received a PhD in computational mathematics from the University of Colorado. Contact him at dquinlan@llnl.gov.

Curtis Janssen is a Distinguished Member of the Technical Staff at Sandia National Laboratories. His research interests include technologies required for high-performance scientific computation: programming models, performance analysis, and machine architecture. Janssen received a PhD in theoretical chemistry from the University of California, Berkeley. Contact him at cljanss@sandia.gov.



Enroll now.

ON THIS BATTLEFIELD, EDUCATION IS YOUR BEST DEFENSE.

Cyber attacks are being waged all over the world, creating an unprecedented demand for trained professionals to protect our country's data assets and develop cybersecurity policies. Help meet the demand with a bachelor's or master's degree in cybersecurity. Whether you plan to work for Cyber Command taking down cyber terrorists or for private industry battling hackers, UMUC can help you make it possible.

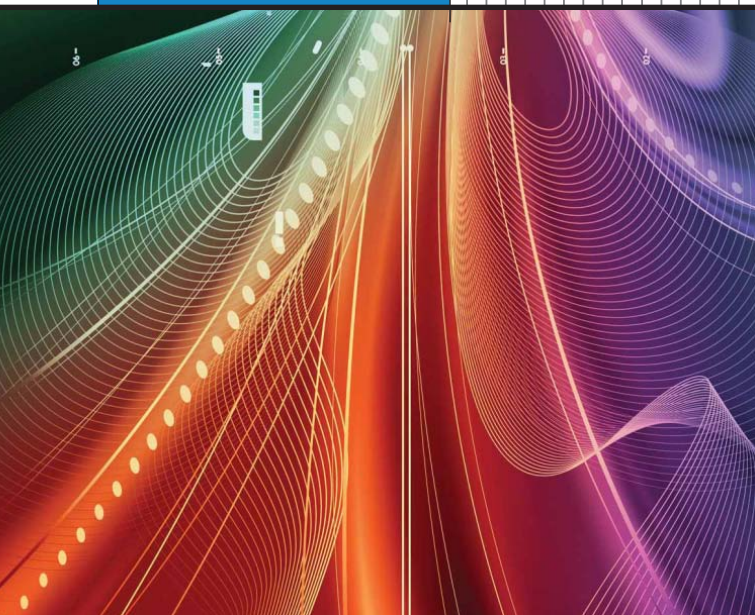
- Designated as a National Center of Academic Excellence in Information Assurance Education by the NSA and DHS
- BS and MS in cybersecurity and MS in cybersecurity policy available
- Programs offered entirely online
- Interest-free monthly payment plan available, plus financial aid for those who qualify

CYBERSECURITY

800-888-UMUC • umuc.edu/cyberwarrior

 UMUC
University of Maryland University College
Copyright © 2011 University of Maryland University College

COVER FEATURE



Codesign for InfiniBand Clusters

Sayantan Sur, Sreeram Potluri, Krishna Kandalla, Hari Subramoni, and Dhableswar K. Panda
Ohio State University

Karen Tomko, *Ohio Supercomputer Center*

Codesigning applications and communication libraries to leverage underlying network features is imperative for achieving optimal performance on modern computing clusters.

Scientific computing can take credit for many of the technological breakthroughs of our generation. It finds use in fields ranging from drug discovery and aerospace to weather prediction and seismic analysis. Scientific computation often deals with large amounts of data, and its algorithms must compute results from mathematical models. Due to their compute- and data-intensive nature, these applications are often parallel, that is, they perform calculations simultaneously on multiple computers.

The TOP500 list ranks supercomputing sites across the world. Recently, top systems have crossed the petaflops (10^{15} floating-point operations per second) barrier. Experts expect it to reach exaflops (10^{18}) levels by the turn of the decade. This fast growth in high-performance computing comes from the low cost of commodity components: general-purpose processors from Intel, AMD, and IBM; graphic processors from Nvidia and others; I/O buses such as PCI Express; and interconnection networks such as InfiniBand.

As machines grow ever larger and more powerful, communication and application stacks must also evolve. An underlying design principle in high-performance computing is to expose, not hide, system features that lead to better performance. However, as system complexity increases, communication library developers must expose them in a manner that does not overwhelm application developers with detail.


To keep scaling applications on increasingly more powerful systems, it is imperative to explore new architectures from the system point of view and new programming paradigms from the application point of view. Here, we explore a codesign approach that takes advanced features from a commodity network (InfiniBand), incorporates the design into a state-of-the-art message-passing interface (MPI) communication library, and then modifies applications to leverage these new features.

APPLICATION SCALING BOTTLENECKS

In practice, parallel applications experience the additional overheads of messaging. One cost is the unnecessary synchronization of processors—a sending process must ensure that the receiving process successfully receives the message. Adopting novel programming model concepts, such as one-sided communication, can solve this oversynchronization problem. In one-sided communication, the processes allocate space to store messages before the actual exchange takes place.

COVER FEATURE

Another factor is the overhead from processors communicating in groups (collective communication), which often involves large volumes of data and communications scheduled to optimize network usage. Although communication scheduling yields better network utilization, one delayed process can delay all the other processes in the collective operation. Additionally, performing the communication scheduling tasks—such as waiting for messages and forwarding them—leads to lost processor cycles. During these tasks, the main CPU cannot perform useful work, lowering overall efficiency. With the help of advanced intelligent networks, the CPU can offload the communication scheduling and progress tasks to the network adapter, freeing up the CPU for useful tasks.



If the architecture's major components, messaging libraries, and applications evolve separately, the result will be a loosely coupled system.

At the same time, if the architecture's major components, messaging libraries, and applications evolve separately, the result will be a loosely coupled system. The weakest link would then limit performance. Thus, it is important to codesign these components and extract the maximum performance from a system.

INFINIBAND NETWORK ARCHITECTURE

InfiniBand (www.infinibandta.org) is a switched interconnect standard used by more than 40 percent of the TOP500 supercomputing systems (www.top500.org). Current generation InfiniBand network cards and switches with quadruple data rate (QDR) speed can deliver 32-Gbps end-to-end bandwidth with about 1- to 1.5- μ s latency.

One of InfiniBand's major features is remote direct memory access. Using RDMA, one process can remotely read or write the memory contents of another process without any involvement of the remote processor. When this powerful feature is used intelligently in communication library design, it can reduce synchronization requirements and improve latency.

The ConnectX-2 network interface is the latest InfiniBand adapter from Mellanox. Along with all the standard InfiniBand features, it offers Core-Direct, a new network offloading feature. Using this feature, a process can create arbitrary lists of send, receive, and wait operations, then post them onto the network adapter's work request queue. The network interface then executes the tasks without involving the host processor. With such task lists, MPI library developers can design nonblocking collective operations.

MESSAGE-PASSING INTERFACE

MPI has been the dominant parallel programming model for the past couple of decades. It has been widely ported, and several open source implementations are available offering very good performance and scalability. As a result, all modern supercomputers support MPI. InfiniBand offers a low-level interface with several types of queue pairs, with varying levels of services. With this interface, developers of upper-level software, such as MPI implementations, can design flexible and high-performance connection management, buffer management, coalescing strategies, and so on.

MVAPICH2 (mvapich.cse.ohio-state.edu) is a high-performance open source implementation of the MPI 2 standard on InfiniBand, the Internet Wide Area RDMA Protocol (iWARP), and RDMA over converged Ethernet (RoCE). MVAPICH2 has a systematic internal design that achieves very good scalability by exploiting various InfiniBand features, such as unreliable datagrams, shared receive queues (SRQs), and extended reliable connections (XRCs). It also uses connection management strategies, such as on-demand connections, and buffering strategies for message coalescing to improve memory efficiency.

MVAPICH2 combines all these optimizations into one unified runtime that offers high performance. To the best of our knowledge, MVAPICH2 has the most scalable runtime on InfiniBand. Over the past 10 years, researchers have used MVAPICH2 as a state-of-the-art MPI for investigating communication runtimes. Developers have also used it as a production MPI library on several InfiniBand clusters worldwide.

PARALLEL SCIENTIFIC APPLICATIONS

Scientific applications employ a wide range of numerical techniques. In our work, we use two different application techniques: finite difference methods and Fourier transforms.

Anelastic wave propagation

Researchers at the Southern California Earthquake Center (SCEC) use the anelastic wave propagation code by Olsen, Dey, and Cui.¹ AWP-ODC is a community model that uses a staggered-grid finite-difference method to solve the 3D velocity-stress wave equation. This method decomposes the volume representing the model's ground area into 3D rectangular subgrids to parallelize the code. Each processor performs stress and velocity calculations for its portion of the grid, applying boundary conditions at the volume's external edges if the subgrid is on the boundary. Ghost cells, comprising a two-cell-thick padding layer, manage the most recently updated wave-field parameters exchanged from the edge of neighboring subgrids.

Researchers have used this code to carry out some of the most detailed simulations to date of earthquakes along the

San Andreas Fault, including the well-known TeraShake and SCEC ShakeOut simulations. This application was a finalist for the Gordon Bell Prize in 2010.²

Parallel Three-Dimensional Fast Fourier Transforms Library

The Parallel Three-Dimensional Fast Fourier Transforms library from the San Diego Supercomputer Center is a portable, high-performance, open source implementation based on the MPI programming model

(code.google.com/p/p3dffft). Researchers have used P3DFFT in direct numerical simulation of turbulence applications.³

P3DFFT leverages the fast serial FFT implementations of either IBM's Engineering and Scientific Subroutine Library (ESSL) or the Fastest Fourier Transform in the West (FFTW) library for efficient one-dimensional FFT calculation. The FFT computations require two costly MPI_Alltoall communication operations to perform matrix transpose operations. It is possible to restructure the P3DFFT library to leverage our proposed implementation of the MPI_lalltoall operation.⁴

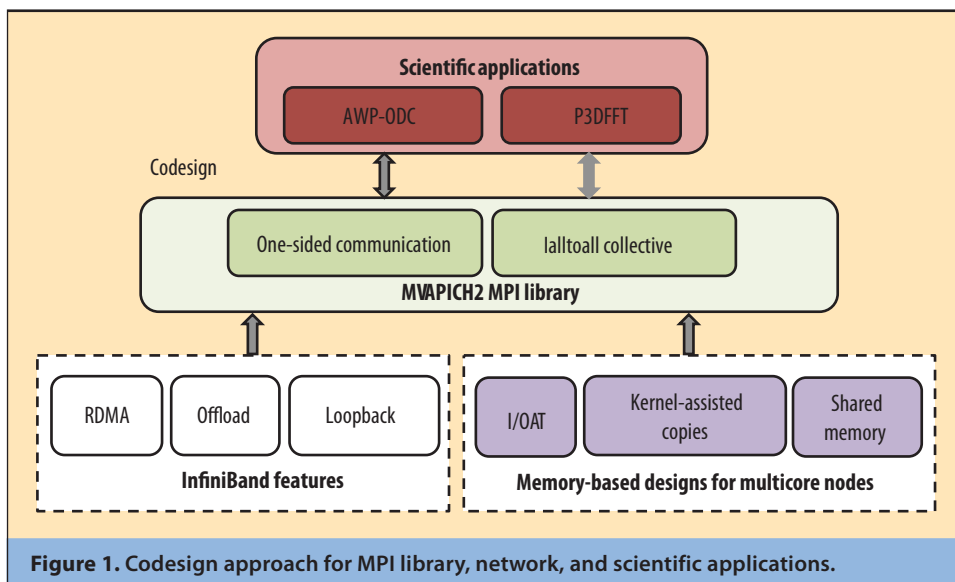
MPI LIBRARY DESIGN

To start codesigning an application with an MPI library, we first enhanced the library to incorporate novel network features. As Figure 1 shows, the InfiniBand features include RDMA, offload, and loopback that an application can leverage in the MPI layer. Modern multicore computing platforms also provide several additional features for optimal data transfer between the cores within a node. An application must also exploit these features for the best performance.

Improving one-sided communication with RDMA

MPI's one-sided model aims at reducing synchronization overheads inherent in communication using send or receive. Each process exposes a region of its memory (a window) to all the other processes in its communication group. Every process can then directly read from, write to, or update window memory at any other process.

The origin process provides the parameters required for communication without any intervention from the target. The communication operations are nonblocking, allowing data transfer to proceed asynchronously and leaving the processor free to do other useful work. The semantics



of one-sided communication form a perfect match with the InfiniBand network's RDMA operations. For example, the MPI library can directly map an MPI_Put (write to a window) onto an RDMA write call. The asynchronous nature of both RDMA operations and the MPI_Put allows for efficient computation and communication overlap.⁵

The application controls issue and completion of one-sided communication calls through synchronization operations. Although all communication operations require only the participation of the origin process, synchronization operations can either be active (requiring the participation of both origin and target) or passive (participation of origin only). MPI_Win_lock and MPI_Win_unlock calls provide passive synchronization, which is usable only point to point. That is, a process must individually lock and unlock a window at each process it wants to communicate with. This leads to inefficiencies in applications that have a fixed communication pattern involving multiple target processes.

MPI_Win_fence is an example of active synchronization with collective semantics that require the participation of all processes in the communicator. MPI provides a flexible mode of active synchronization with MPI_Win_post, MPI_Win_wait, MPI_Win_start, and MPI_Win_complete. Using these calls, MPI allows a subgroup of processes in the communicator to synchronize. Doing so leads to better performance, particularly in applications where a process communicates only with a small set of other processes: for example, its neighbors in the process grid.

Designing nonblocking alltoall exchange with collective offload

The Alltoall personalized exchange is the most communication-intensive collective operation in the MPI 2.2 standard. With N processes, the latency of a large-message Alltoall operation is proportional to N^2 , which significantly

COVER FEATURE

```

1D FFT in  $x$  for  $V_1$ 
transpose  $x$  and  $y$  of  $V_1$ 
1D FFT in  $y$  for  $V_1$ 
Initiate  $y$  and  $z$  transpose with MPI_Alltoall of  $V_1$ 
do  $V_j = V_2$  to  $V_n$ 
  1D FFT in  $x$  for  $V_j$ 
  transpose  $x$  and  $y$  of  $V_j$ 
  1D FFT in  $y$  for  $V_j$ 
  Initiate  $y$  and  $z$  transpose with MPI_Alltoall of  $V_j$ 
  Wait for transpose complete for  $V_{j-1}$ 
  1D FFT in  $z$  for  $V_{j-1}$ 
enddo
Wait for transpose complete for  $V_n$ 
1D FFT in  $z$  for  $V_n$ 

```

Figure 2. Algorithm for the forward transform in the redesigned multivariable, pipelined, overlapped FFT routines.

affects the performance and scalability of various scientific applications. A potential strategy to improve application performance would be to design a high-performance, nonblocking implementation for the Alltoall exchange, which applications can leverage to overlap the Alltoall communication with computation. It is possible to use host-based approaches to design nonblocking collectives (libNBC).⁶ However, such methods require the host processor to process the collective operation, which directly limits the overlap. It also is not very portable.

MPI libraries can leverage the network offload feature in the ConnectX-2 InfiniBand adapter to design nonblocking collectives. However, the current ConnectX-2 interface limits the size of the task lists that a process can post to the adapter, directly affecting the scalability of collective operations, such as MPI_Alltoall.

To overcome this limitation, we divide the entire Alltoall operation across multiple task lists and rely on a lightweight thread to post them. Since the adapter can execute task lists independently, the progress thread is active for a very short duration, minimizing its contention with the application thread. We create a separate queue pair (trigger_qp), a completion queue (trigger_cq), and a completion channel (trigger_comp_channel) to let a process communicate with itself through InfiniBand's blocking progression mode. At the end of a task list, a process enqueues a send task to itself on the trigger_qp. The offload-progress thread posts the task list, calls `ibv_get_cq_event` on the trigger_comp_channel, and schedules itself into a sleep state. The adapter executes the task list and finally executes the send on trigger_qp. This generates a network interrupt on trigger_comp_channel, signaling the progress thread to post the next task list (if any).

LEVERAGING THE IMPROVED MPI LIBRARY

Scientific applications that have nearest-neighbor or many-to-many communication patterns and computa-

tion that can overlap communication will benefit from the improved MPI library.

One-sided communication in AWP-ODC

AWP-ODC expends most of its execution time to compute and exchange two variables: velocity and stress. Both of these variables have multiple components, each of which corresponds to a data grid. During the exchange phase, each process sends its data grid boundaries to the neighbors in all directions and receives boundary data from them. The computation of each individual component within velocity and stress are independent of one another. However, there is a dependency across the velocity and stress components.⁷ This understanding of the data dependencies and our knowledge of RDMA-based one-sided designs in MVAPICH2 forms the basis for our codesigned version of AWP-ODC.

We use MPI 2 one-sided communication primitives to provide buffers into which neighboring processes can use the Put operation to directly place their data. This enables each process to complete data exchanges without synchronizing with its neighboring processes. The processes only need to synchronize before the point where they use the new data. MVAPICH2 maps the Put operation directly to RDMA. This helps overlap the transfer of one component with the computation of others.

Typical commodity compute nodes have 16 to 48 cores on each node. A significant portion of communication is intranode using shared memory. The shared-memory communication channel requires the CPU to make the data copies. This does not allow for overlap between the communication and any computation. Kernel-assisted schemes provide better copy performance but still suffer from the lack of overlap. Input/output acceleration technology provides overlap using direct memory access.⁸ InfiniBand provides a loopback communication model, which lets processes on the same node communicate through the network adapter. Our design uses this channel to enable overlap. Because the communication is completely hidden under useful computation, the increased latencies do not negatively impact application performance.

Collective offload and overlap in P3DFFT

The Cooley-Tukey algorithm for one-dimensional FFTs is computationally efficient, but this algorithm's butterfly pattern of memory accesses makes it a challenge to scale. P3DFFT first performs a one-dimensional FFT along the x dimension, followed by a transpose of the x and y dimensions. It then repeats the same pattern across the y and z dimensions, followed by a one-dimensional FFT along the z dimension. The original data array is typically distributed along the x dimension, with the y and z dimensions split among processors in rows and columns of a two-dimensional processor grid. There are two expensive

Alltoall operations to transpose data among row and column (col) communicators. Typically, the row communicator maps to cores within a node or adjacent nodes. The col communicator spans multiple nodes. In our work, we focus on replacing the MPI_Alltoall on the col communicator with MPI_lalltoall.

Figure 2 shows the FFT routines (forward and back) after we restructured them. Loop index j runs over the variables for transformation. During the j iteration, we overlap the FFT operations and the XY(row) transposition for j with the YZ(col) transposition of the $(j - 1)$ iteration, which relies on the MPI_lalltoall operation.

CODESIGN EXPERIMENTS

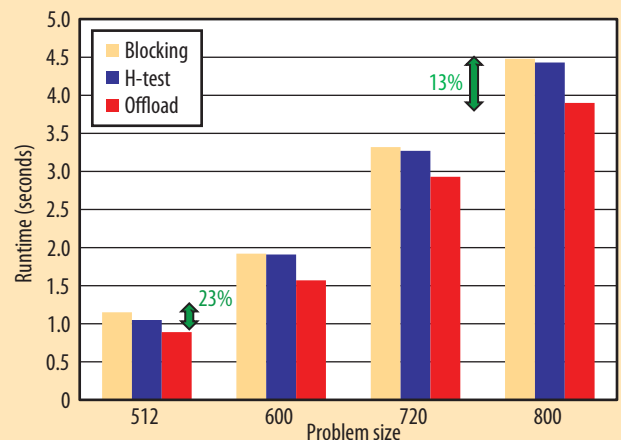
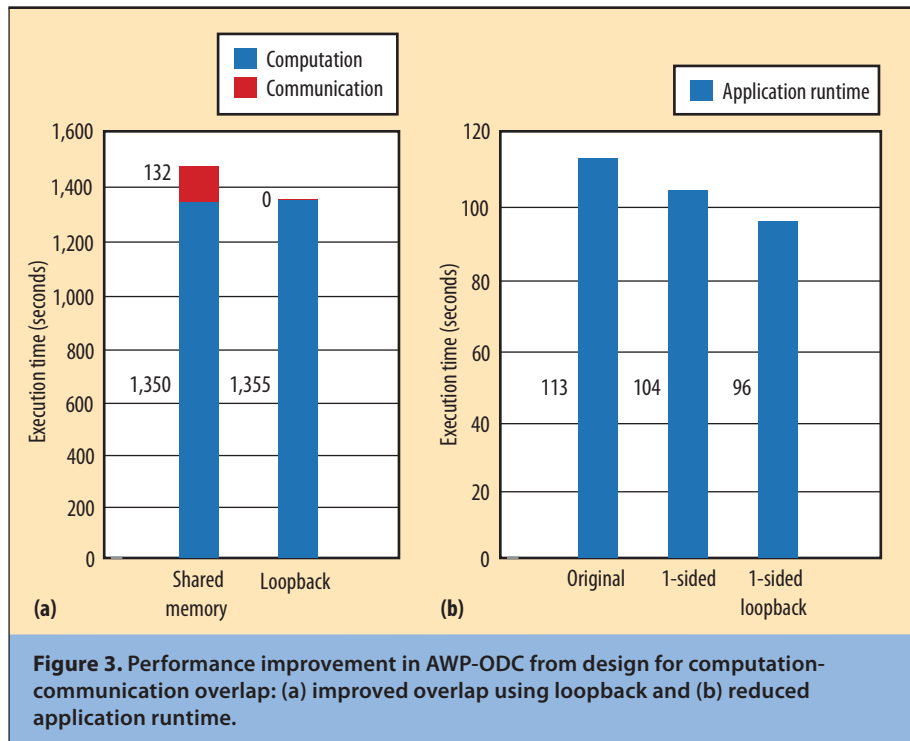
To assess the improved MPI library, we measured AWP-ODC and P3DFFT performance when using our restructured codes.

AWP-ODC improvements

We first evaluated the effectiveness of loopback transfers in overlapping communication and computation. We ran the AWP-ODC application with 48 processes on a pair of AMD Magny-Cours machines (24 processors per node) connected using quad data rate (32 Gbps) InfiniBand interconnects. Each process operates on data grids of $128 \times 128 \times 128$ elements.

As Figure 3a shows, there is a 132-second overhead when processes on a node communicate through shared memory. When these processes use the loopback channel, the network adapter handles data movement asynchronously, completely overlapping communication with computation. Therefore, the communication time drops to zero while the computation time remains the same.

The platform for our next experiment, the Ranger supercomputer at the Texas Advanced Computing Center, is one of the largest InfiniBand clusters available for open science research. Each Ranger node contains 16 processors and 32 Gbytes of memory; the nodes connect via a single data rate (8 Gbps) InfiniBand network. Because of Ranger's slower network, minimizing communication costs is important for application efficiency. Figure 3b compares the performance of the original and our enhanced versions of AWP-ODC on 8,192 processors. Using our overlap design and the loopback channel produces an improvement of 15 percent in the total application runtime.



P3DFFT library improvements

This experiment demonstrates the benefit of collective communication offload to a network adapter. We ran it on a 512-core cluster. Each node has eight Intel Xeon cores running at 2.53 GHz with 12 Mbytes of L3 cache and 12 Gbytes of memory. The nodes connect through a quad data rate (32 Gbps) InfiniBand network. We used the test_sine kernel to evaluate the benefits of our modified P3DFFT library. Figure 4 compares the application runtimes of the baseline blocking version with the library redesigned for overlapped collective communication, using host- and network-offload-based MPI_lalltoall implementations. We


COVER FEATURE

ran our test on 128 cores while varying problem size N between 512 and 800.

The kernel with our proposed MPI_Ialltoall consistently outperforms the one with blocking MPI_Alltoall by about 10 to 23 percent. It outperforms the kernel that uses the nonblocking MPI_Ialltoall operation using the host-based approach by about 10 to 17 percent.

The high-performance computing field is forging ahead with complex system architecture designs.

Predictions indicate that computing power will have surpassed the exaflops level by the turn of the decade. Providing balanced system performance will require cohesive design of the processor, memory hierarchy, network architecture, and topology. At the same time, developers must modify applications to fully leverage system features.

Our work is a step toward this goal. We will continue to enhance MVAPICH2 by efficiently exposing features of new system architectures through codesign. We will optimize MPI communication for modern system components such as accelerators so that applications can use them effectively. We also will advance the codesign principle by working with scientists and engineers to redesign their applications alongside MVAPICH2, taking complete advantage of the system capabilities. As a member of the MPI Forum, we will use our experiences from MVAPICH2 and application codesign in improving the MPI communication standard to address the community's changing needs. 

Acknowledgments

This research is supported in part by US Department of Energy grants DE-FC02-06ER25749 and DE-FC02-06ER25755; National Science Foundation grants CCF-0621484, CCF-0702675, CCF-0833169, CCF-0916302, CCF-0926691, and 0937842; grants from Intel, Mellanox, Cisco, QLogic, and Sun Microsystems; and equipment donations from Intel, Mellanox, AMD, Appro, Chelsio, Dell, Fujitsu, Fulcrum, Microway, Obsidian, QLogic, and Sun Microsystems.

References

1. K.B. Olsen, "Simulation of Three-Dimensional Wave Propagation in the Salt Lake Basin," doctoral dissertation, Univ. of Utah, Salt Lake City, 1994.
2. Y. Cui et al., "Scalable Earthquake Simulation on Petascale Supercomputers," *Proc. ACM/IEEE Int'l Conf. High Performance Computing, Networking, Storage and Analysis (SC 10)*, IEEE, 2010, pp. 1-20.
3. D.A. Donis, P.K. Yeung, and D. Pekurovsky, "Turbulence Simulations on O(104) Processors," *Proc. 3rd Ann. TeraGrid Conf. (TeraGrid 08)*, 2008; www.sdsc.edu/us/resources/p3dfft/docs/TG08_DNS.pdf.
4. K. Kandalla et al., "High-Performance and Scalable Non-Blocking All-to-All with Collective Offload on InfiniBand Clusters: A Study with Parallel 3D FFT," *Computer Science—Research and Development*, Springer, 2011, pp. 237-246.
5. W. Jiang et al., "High Performance MPI-2 One-Sided Communication over InfiniBand," *Proc. IEEE/ACM Int'l Symp. Cluster Computing and the Grid (CCGrid 04)*, IEEE, 2004, pp. 531-538.
6. T. Hoefler et al., *Non-Blocking Collective Operations for MPI-2*, tech. report, Open Systems Lab, Indiana Univ., 2006.
7. S. Potluri et al., "Quantifying Performance Benefits of Overlap Using MPI-2 in a Seismic Modeling Application," *Proc. Int'l Supercomputing Conf. (ICS 10)*, ACM, 2010, pp. 17-25.
8. P. Lai, S. Sur, and D.K. Panda, "Designing Truly One-Sided MPI-2 RMA Intra-Node Communication on Multi-Core Systems," *Computer Science—Research and Development*, Springer, 2010, pp. 3-14.

Sayantana Sur is a research scientist in the Department of Computer Science and Engineering at Ohio State University. His research interests include high-speed interconnection networks, high-performance computing, fault tolerance, and parallel computer architectures. Contact him at sayantana_sur@gmail.com.

Sreeram Potluri is a PhD student in the Network-Based Computing Laboratory at Ohio State University. His research interests include high-speed interconnects, accelerator technologies, parallel programming models, and high-end computing applications. Contact him at potluri@cse.ohio-state.edu.

Krishna Kandalla is a PhD student in the Department of Computer Science and Engineering at Ohio State University. His research interests include high-performance computing, high-speed interconnects, and MPI collective communication. Contact him at kandalla@cse.ohio-state.edu.

Hari Subramoni is a PhD student in the Department of Computer Science and Engineering at Ohio State University. His research interests include high-performance computing, high-speed interconnects, high-performance data transfers in InfiniBand WAN scenarios, and differentiated quality of service in high-performance computing. Contact him at subramon@cse.ohio-state.edu.

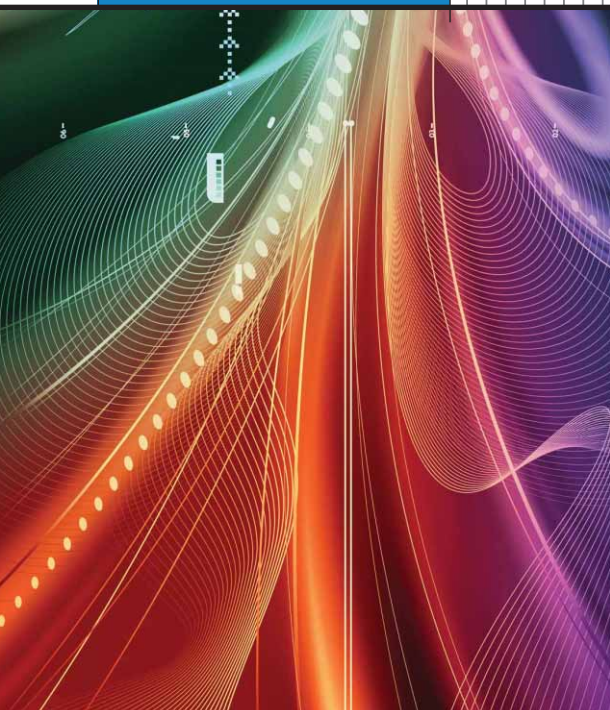
Dhableswar K. Panda is a professor of computer science and engineering at Ohio State University. His research interests include parallel computer architecture, high-performance networking, InfiniBand, exascale computing, virtualization, and cloud computing. Contact him at panda@cse.ohio-state.edu.

Karen Tomko is a senior research scientist at the Ohio Supercomputer Center. Her research interests include high-performance computing, application optimization, and accelerator technologies. Contact her at ktomko@osc.edu.



Selected CS articles and columns are available for free at <http://ComputingNow.computer.org>.

COVER FEATURE



Codesign Challenges for Exascale Systems: Performance, Power, and Reliability

Darren J. Kerbyson, Abhinav Vishnu, Kevin J. Barker, and Adolfo Hoisie

Pacific Northwest National Laboratory

The complexity of large-scale parallel systems necessitates the simultaneous optimization of multiple hardware and software components to meet performance, efficiency, and fault-tolerance goals. A codesign methodology using modeling can benefit systems on the path to exascale computing.

Given the complexity of current systems in terms of scale, memory hierarchy, and interconnection network topology, optimizing for performance alone is a monumental task. Systems with hundreds of thousands of processor cores are now commonplace, memory hierarchies routinely consist of three levels of cache, and interconnection networks with high-dimensional meshes, fat trees, and hierarchical fully connected topologies will dominate in the near term.

Optimization to an architecture is just one of many phases in an application's life span. However, this common process is considered to be "one way" in that it works for already designed and implemented architectures. Future exascale systems and applications will have additional performance, power, and resiliency requirements that represent a multidimensional optimization challenge. A codesign process can optimize two or more factors in

concert to achieve a better solution, ultimately leading to highly tuned exascale systems and workloads.

THE CODESIGN PROCESS

As Figure 1 shows, five factors in the codesign process contribute to the complexity of extreme-scale systems:

- Multiple *algorithms* used for a calculation can exhibit different computational characteristics. For example, using a uniform resolution of a data grid could lead to an implementation with memory and communication characteristics requiring computation exceeding that of a more complex adaptive mesh refinement implementation.
- The *application* represents the implementation of a particular method and comprises a component of the overall workload of interest. Using several applications in concert can explore multiple aspects of a physical system simultaneously, such as climate simulations that consider land, sea, and atmospheric components together.
- The *programming model* underlies the application and defines the way it expresses computation. The two commonly used approaches for expressing parallelism are *process-centric*, in which interprocess communication is expressed explicitly—for example, the message-passing interface (MPI)—or *data-centric*,

COVER FEATURE

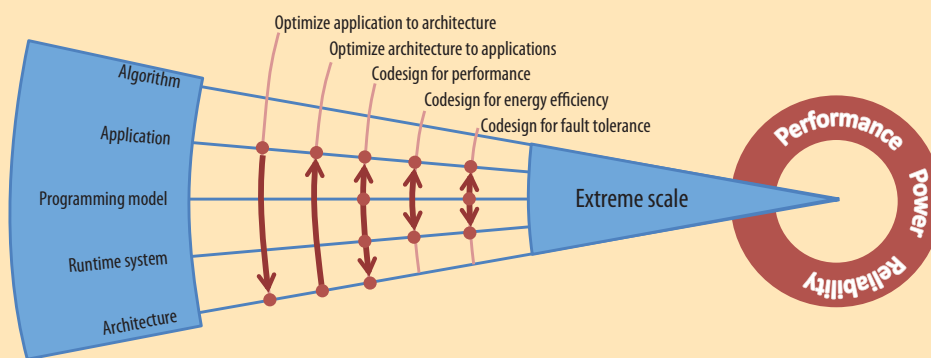


Figure 1. The five factors that contribute to complexity in extreme-scale systems: algorithms, applications, programming models, runtime systems, and architectures. The arrows show examples of the interplay of these factors in design and codesign at an extreme scale.

in which access to any data across the system can occur from any location—for example, Global Arrays, Unified Parallel C (UPC), and Co-Array Fortran (CAF).

- The *runtime system* is responsible for ensuring that application requirements are dynamically satisfied and mapped onto system resources. It includes process and data management and migration.
- The *architecture* includes the processor core's micro-architecture, the arrangement of cores within a chip, memory hierarchy, system interconnect, and storage subsystems.

No codesign process to date has covered all these factors comprehensively, but some notable cases have addressed a subset of them and their corresponding tradeoffs:

- *Optimization of an application to an architecture.* For an already implemented system architecture, this process requires mapping application workloads to architecture characteristics. It is commonplace in application development and software engineering but is not considered codesign.
- *Optimization of an architecture to an application.* Given an already implemented application, this process optimizes architecture design to achieve high performance. An example is the performance-guided design of large-scale IBM Power7 systems;¹ again, this is not considered codesign.
- *Codesign for performance.* Enabling application and architecture to best match each other unlocks the potential to achieve the highest performance in a new system. An example is the design of an application with the first petaflops system—the IBM Roadrunner.²
- *Codesign for energy efficiency.* Energy consumption by extreme-scale systems will increasingly become a design constraint and a notable cost factor. The largest systems today consume more than 10 megawatts,

incurring an operational cost of roughly US\$10 million per year. Our experiences in codesign for energy involved an application built to provide information about expected periods of idleness and a runtime aimed at lower power consumption.³

- *Codesign for fault tolerance.* A critical factor in extreme-scale system operation is fault tolerance. Traditional methods that use checkpoint-restart mechanisms, which store partial results to disk and can be pulled back if a fault occurs, might not scale well as system sizes increase. However, the use of selective methods, such as replicating just the critical data across system memory, can help reconstruct state from failed nodes and enable job execution to continue. We use codesign involving the application, programming model, and runtime system to improve resiliency.

Several of our own experiences with codesign methodologies have resulted in improved performance, power efficiency, and reliability.

PERFORMANCE

We draw on our past experience to design and deploy the IBM Roadrunner, the first petaflops system, which was also the first extreme-scale hybrid architecture design.² From the outset, we envisioned Roadrunner as an accelerated system using conventional compute nodes, each one hosting accelerators capable of high processing rates. Three years prior to Roadrunner's deployment, we explored three different classes of accelerators: Clearspeed multi-SIMD processors, GPUs, and the IBM Cell Broadband Engine.⁴ Although we ultimately used the IBM Cell, it was not a foregone conclusion.

The use of quantitative and accurate performance modeling of full applications facilitated Roadrunner's codesign.⁵ Each model encapsulated the first-order per-

formance characteristics of an application and was parameterized in both system and application factors. These factors form the variables that allow exploring the application and system design space simultaneously. Typical model parameters include the application's processing requirements (computation, memory, and communication) and hardware resource capabilities such as processing rate, data movement to memory, interconnection of accelerator to host, and interconnections between nodes.

Many model inputs use empirical data when system components are available for measurement or simulated results for codesign in future systems. Because they are analytically based, models can readily explore the design space, and they can be used for full applications with low overhead at runtime.

The key aspect in the Roadrunner design was to determine how useful, if at all, accelerators would be on the workload of interest; we did not want to consider just the peak speeds and feeds or simulate kernels that might not be correlated to actual application performance.

The key architectural characteristics include the number of cores available on each accelerator, the capabilities of each core, and the high communication costs incurred when data moves between host and accelerator memories. Figure 2 shows an overview of the architecture. The increased cost of communication becomes clearer, given the transfer path from one accelerator memory to another on a distance node that requires three steps: transfer from accelerator to host, host to host, and host to accelerator. Many current large-scale accelerated systems correspond to the architecture shown in Figure 2, but Roadrunner is the only system with two accelerators, each containing two processors in each compute node.

We used a particular application of interest—deterministic transport, which uses wavefront algorithms—in the codesign.⁶ In this application, the processing of a data block

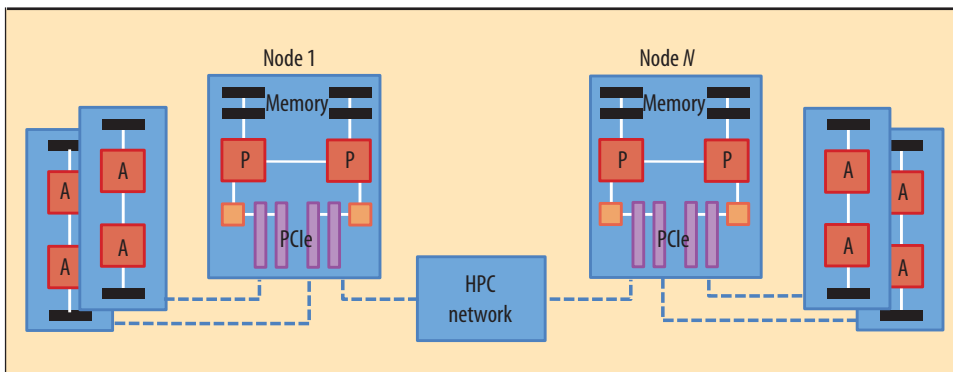


Figure 2. Example architecture of an accelerated system. Each compute node contains a conventional host node with two processors (P) and local memory, as well as one or more accelerators (two are illustrated here), each with one or more accelerator processors (A).

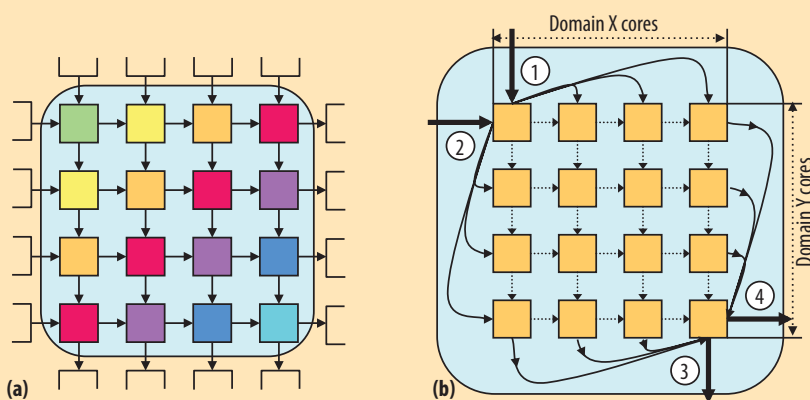


Figure 3. Processor interaction on a subset of a 4×4 logical processor-core arrangement: (a) default structure requiring four communications across a domain boundary in each direction in each step, and (b) communication structure designed for accelerated systems requiring one communication per several steps.

can only happen after a processor core receives data from upstream neighbors. Figure 3a shows an overview of this for a small, 16-processor-core example arranged in a logical 4×4 two-dimensional array. The colors denote that processors on the same diagonal are handling the same block of their local data, and arrows denote the dependencies (communications) between processors. Each step in the operation involves the processing of a data block and the communication of data to downstream neighbors. In Roadrunner, each accelerator core is considered to be a member of the logical two-dimensional processor array.

To design a wavefront application for an accelerated system, we had to significantly alter the communication structure so that it could overcome the high costs of data transfer between accelerators. We reduced the number of messages between accelerators by combining the required communications within a processor-core domain that occur once for several steps, as Figure 3b illustrates. This also resulted in extra computation steps, which dwarfed the communication savings at a small scale.⁷

COVER FEATURE

Table 1. Wavefront application performance after implementing both system and application.

Accelerator core count	Default time (sec)	Optimized time (sec)	Improvement (%)
8,192	1.84	1.96	-6.0
16,384	2.22	2.28	-2.6
32,768	2.81	2.56	9.6
65,536	3.66	3.18	15.2
97,920	4.15	3.27	27.0

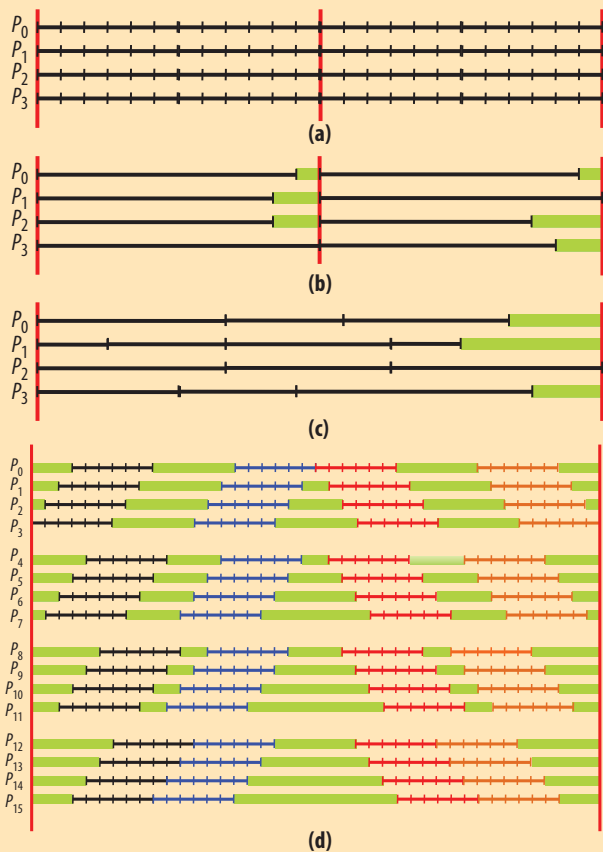


Figure 4. Examples of idle periods in several classes of applications: (a) load balanced over processors and iterations; (b) load imbalanced, varying over application iterations; (c) task-based processing with starvation at end of an iteration; and (d) algorithmic dependencies leading to idle periods on all processors at different times. A red barrier denotes application synchronization points, green regions denote idle times, and ticks along the timelines denote application steps or tasks.

As Table 1 shows, running the wavefront application on Roadrunner resulted in performance improvements at larger scales.

Although the codesign process determined both the best accelerator configuration and the best implementation of the application for Roadrunner, a further complementary activity happened in the development of the reverse

acceleration model,⁸ also guided by performance modeling. Using this programming model with a small runtime system, we programmed each accelerator core directly with a subset of the MPI. Thus, each accelerator core became an MPI rank, and activities that the accelerator could not handle, such as I/O, were offloaded to the host. The design and implementation of the reverse acceleration model was not part of the Roadrunner codesign process, but its capabilities were assumed to be available during the design exploration.

POWER

Current techniques to reduce energy consumption include reducing the power state and throttling down system components such as processor cores during idle periods in the processing flow. The power state is a function of the frequency at which the component operates and is often coupled with the supply voltage. Runtime systems can use dynamic voltage and frequency scaling (DVFS) to alter the power state dynamically but at a cost of thousands of processor-clock cycles.

A key issue is how and where to take advantage of these power-saving features. Researchers have proposed several methods, but their effectiveness depends on the parallel structure of applications—for example, application-transparent techniques for saving energy have successfully exploited the idle periods prior to global operations.⁹ Figures 4b and 4c show such idleness, which is caused by placing different computational requirements on individual processors.

In certain applications, idle periods are not associated with load imbalance—for example, the parallel activity of deterministic transport in Figure 4d has a defined pattern in which idleness occurs at different times and for different durations due to the wait for incoming data from other processors. This is not due to any inefficiency in the application but is algorithmic, caused by data dependencies that must be satisfied before computation can proceed. Runtime systems have particular difficulty in automatically identifying this activity as it is not associated with global synchronizations; this is especially true when the behavior is time varying. Application-specific information can identify a priori when processor cores will wait for incoming data and thus can be placed in a low-power state to save energy.

The recently proposed energy template approach³ is a codesign between the application that quantifies expected periods of idleness on a per-core basis and the runtime system that affects changes in the processor power state. Central to an energy template is an analytical model that guides the runtime system in making informed decisions about when to idle a processor core. The energy template principles are to

- represent an application-specific sequence of active and idle states for each processor core,
- contain the rules associated with the transition from one state to another using a model,
- use triggers that enable state transitions by transparently monitoring application activity, and
- enable the runtime system to make informed decisions about when to alter an individual processor core's power state.

NW-ICE, a power-instrumented testbed at the Pacific Northwest National Laboratory, illustrates the value of the energy template approach when using a wavefront application. As Figure 4d shows, idleness naturally occurs in this application because processor cores need to wait for data from upstream neighbors. Figure 5 shows the execution time, power consumption, and energy used in the wavefront application on a single rack of NW-ICE containing 28 nodes, each with two sockets of the quad-core Intel Harpertown processor. Only two power states for each core are available on NW-ICE, an idle and an active state, with a differential of 11 watts per core.

At most, we find a 4 percent difference in the time the wavefront application takes running with and without the energy template. This variation is small in performance analysis terms and occurs due to the extra overhead induced by the energy template and runtime. On the other hand, the magnitude of a power savings of 8 percent is significant and must be contrasted with the peak possible savings of 23 percent on the test system. The energy saved—the product of the time and the power—increases with processor count and should further increase in larger systems and with increased difference in power states.⁹

RELIABILITY

Fault tolerance is imperative to realizing the dream of sustainable computing on extreme-scale systems: failure rates increase with system size, yet the cost of fault recovery must be independent of scale. To address these challenges, we designed a fault-tolerance system that

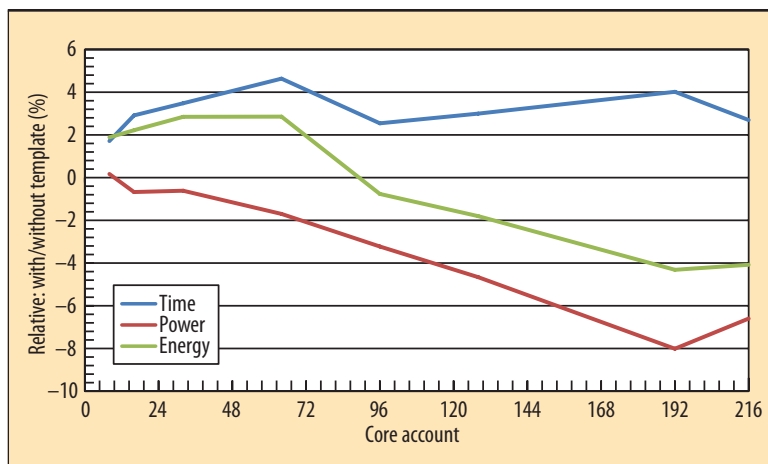


Figure 5. Comparison using an energy template on a wavefront application at multiple processor core counts on a single NW-ICE rack. Relative power use and energy consumption decrease with system size.

involves the application, programming model, and runtime—a clear codesign. Its primary focus is the replication of critical application data across the system to enable continued job execution in the presence of node failures. The codesign utilizes a task-based approach that applies the Global Arrays (GA) programming model,¹⁰ in which a task is designed as a unit of computation with input, output, and dependencies to other tasks. A task's self-contained properties make one-sided communication a perfect model for describing data dependencies, simplifying the implementation of many applications. GA itself was codesigned in conjunction with the NW-Chem computational chemistry application in the 1990s and is under active development at the Pacific Northwest National Laboratory.¹¹

The self-containment and data-centric nature of task-based execution models has important implications for fault tolerance. Our approach leverages these properties because it is scalable and the cost of recovery is proportional to the degree of failure. There are several requirements for achieving fault tolerance:

- Critical application data must be accessible even if the compute nodes that contain a portion of the global data become inaccessible. This is possible by using selective replication of critical data that is both read from and written to, and by using Reed-Solomon encoding of read-only data.
- A fault-tolerance management infrastructure must support continued execution.¹² FTMI includes highly reliable fault detection that leverages the features of modern high-performance interconnects, a fault-resilient process manager that enables continued application execution, fault-tolerant synchronization—a necessary and sufficient collective communication primitive for partitioned global address space (PGAS)

COVER FEATURE

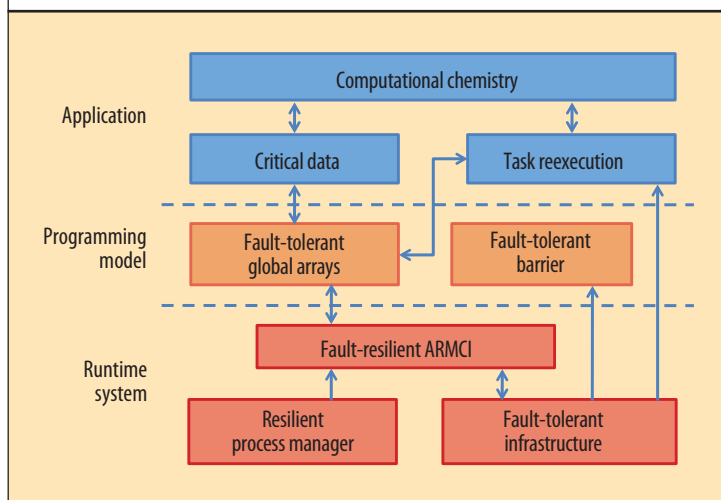


Figure 6. A fault tolerance codesign that includes interactions between the application, programming model, and runtime system.

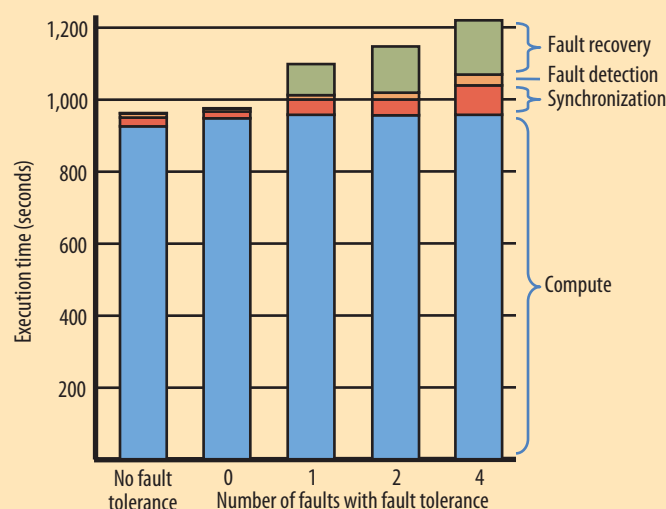


Figure 7. Performance evaluation of the fault-tolerant coupled-cluster (CC) and Global Arrays (GA) using 4,096 processors. The performance impact of zero to four faults during application execution is shown.

models, and fault information propagation that reduces detection overhead.

- To ensure correctness, data must be in a consistent state at the time of recovery. This is possible by adding metadata to each task that records state transitions and that the system distributes and replicates as critical data to ensure the task state can be recovered. Equally critical is the component that maintains consistency during data writes. FTMI handles this by synchronizing individual writes to remote locations in memory. Many components use this fault detection, including collective communication, the PGAS data store layer, and the application layer for state transition.

Figure 6 shows an overview of a fault-tolerant system we designed in conjunction with computational chemistry domain requirements; it is also applicable to many applications that use PGAS programming models. Many methods commonly used for computational chemistry trade off time and accuracy, including the coupled-cluster (CC) method.¹⁵ CC's noniterative nature makes it inherently difficult to save intermediate states and use traditional checkpoint-restart methods for fault tolerance. In a CC calculation, the total amount of critical data is proportional to N^4 —several orders of magnitude smaller than the complexity of the computation proportional to N^7 , where N is equal to the number of basis functions and typically ranges between 100 and 500. The computation within CC is task-based and utilizes GA. We codesigned FTMI with the fault-tolerant version of CC, providing necessary tools for fault detection, containment, and large-scale recovery.

FTMI is currently in use on large-scale Cray and InfiniBand systems. Figure 7 shows the execution time of fault-tolerant CC for the Uracil molecule, using 4,096 processes of an AMD/InfiniBand 2310 compute node system at the Pacific Northwest National Laboratory. In the absence of failures, overhead is negligible, making the fault-tolerant implementation highly effective compared with traditional checkpoint-restart methods. In the presence of one node failure, the overhead is 15 percent; the overall recovery cost in checkpoint-restart includes the time to restart all the processes proportional to the system's size.

A major challenge for high-performance computing as it marches to exascale levels is in providing practical and integrated approaches for codesign that consider performance, power, and fault tolerance in concert, as well as algorithms, applications, programming models, runtime systems, and hardware architecture. Although codesign methodologies represent a challenge for the architect, the exponentially increased degrees of freedom to consider the complexities, scale, and costs will warrant using these methodologies to achieve maximum system productivity, as our examples have demonstrated. Accurate predictive tools including analytical modeling have proven to be ideal vehicles to use in such codesign. **□**

Acknowledgments

This research is supported by the US Department of Energy's Office of Advanced Scientific Computing Research, grants #59493 and #59542. The Pacific Northwest National Laboratory is operated by Battelle for the US Department of Energy under contract DE-AC05-76RL01830.

References

1. K.J. Barker, A. Hoisie, and D.J. Kerbyson, "An Early Performance Evaluation of Power7-IH HPC Systems," *Proc. ACM/IEEE Conf. Supercomputing* (SC 11), IEEE CS, 2011, pp. 1-11.
2. K.J. Barker et al., "Entering the Petaflop Era: The Architecture and Performance of Roadrunner," *Proc. ACM/IEEE Conf. Supercomputing* (SC 08), IEEE CS, 2008, pp. 1-11.
3. D.J. Kerbyson, A. Vishnu, and K.J. Barker, "Energy Templates: Exploiting Application Information to Save Energy," *Proc. IEEE Int'l Conf. Cluster Computing* (Cluster 11), IEEE CS, 2011, pp. 1-9.
4. D.J. Kerbyson and A. Hoisie, "A Performance Analysis of Two-Level Heterogeneous Processing Systems on Wave-Front Algorithms," *Unique Chips and Systems*, E. John and J. Rubio, eds., CRC Press, 2007, pp. 259-279.
5. K.J. Barker et al., "Using Performance Modeling to Design Large-Scale Systems," *Computer*, Nov. 2009, pp. 42-49.
6. K.R. Koch, R.S. Baker, and R.E. Alcouffe, "Solution of the First-Order Form of the 3D Discrete Ordinates Equation on a Massively Parallel Processor," *Trans. Am. Nuclear Soc.*, vol. 65, 1992, pp. 198-199.
7. D.J. Kerbyson, M. Lang, and S. Pakin, "Adapting Wave-Front Algorithms to Efficiently Utilize Systems with Deep Communication Hierarchies," *Parallel Computing*, vol. 6, 2011, pp. 550-561.
8. S. Pakin, M. Lang, and D.J. Kerbyson, "The Reverse Acceleration Model for Programming Petascale Hybrid Systems," *IBM J. Research and Development*, vol. 53, no. 5, 2009, pp. 8:1-8:15.
9. B. Rountree et al., "Adagio: Making DVS Practical for Complex HPC Applications," *Proc. 23rd Int'l Conf. Supercomputing* (ICS 09), ACM, 2009, pp. 460-469.
10. J. Nieplocha, R.J. Harrison, and R.J. Littlefield, "Global Arrays: A Nonuniform Memory Access Programming Model for High-Performance Computers," *J. Supercomputing*, vol. 10, no. 2, 1996, pp. 169-189.
11. J. Nieplocha et al., "Advances, Applications and Performance of the Global Arrays Shared Memory Programming Toolkit," *Int'l J. High-Performance Computing and Applications*, vol. 20, no. 2, 2006, pp. 203-231.
12. A. Vishnu et al., "Fault-Tolerant Communication Runtime Support for Data Centric Programming Models," *Proc. Int'l Conf. High-Performance Computing* (HiPC 10), IEEE, 2010, pp. 1-9.
13. H.V. Dam, A. Vishnu, and W.D. Jong, "Designing a Scalable Fault Tolerance Model for Computational Chemistry: A Case Study with Coupled Cluster Perturbative Triples," *J. Chemical Theory and Computation*, vol. 7, no. 1, 2011, pp. 66-75.

Darren J. Kerbyson is a Laboratory Fellow at the Pacific Northwest National Laboratory. His research interests include the analysis and modeling of performance, power, and fault resiliency of future systems. Kerbyson received a PhD in computer science from the University of Warwick, UK. He is a member of IEEE. Contact him at darren.kerbyson@pnnl.gov.

Abhinav Vishnu is a member of the High-Performance Computing Group at the Pacific Northwest National Laboratory. His research interests include designing scalable, energy-efficient, fault-tolerant programming models and

communication runtime systems on high-speed interconnects. Vishnu received a PhD in computer science from the Ohio State University. He is a member of IEEE. Contact him at abhinav.vishnu@pnnl.gov.

Kevin J. Barker is a member of the High-Performance Computing Group at the Pacific Northwest National Laboratory. His research interests include developing performance modeling methodologies and tools for high-performance computing systems and workloads as well as understanding how current and future architectures impact performance. Barker received a PhD in computer science from the College of William and Mary. Contact him at kevin.barker@pnnl.gov.

Adolfy Hoisie is a Laboratory Fellow and director of the Center for Advanced Architectures at the Pacific Northwest National Laboratory. His research focuses on performance analysis and modeling of systems and applications, areas in which he has published extensively. He is a member of IEEE. Contact him at adolffy.hoisie@pnnl.gov.

cn Selected CS articles and columns are available for free at <http://ComputingNow.computer.org>.

Nokia, Inc.

has the following positions in
Sunnyvale, CA:

Principal Engineer, Browsing

Work with browser engine/UI on mobile & embedded platforms; code optimization for ARM processor; broad knowledge of Mozilla architecture; development of multi-process mobile browser & multi-process NPAPI plug-ins rendering implementation; & other duties/skills required. [Job ID: NOK-SVII-PEB]

Senior Software Engineer

Involves software development; exp. with Qt to involve development of C++/Qt applications with OOPS principal; exp. with C++ to involve object oriented design & development; extensible & adaptive framework development using C++, implementation knowledge of C++ design patterns; & other duties/skills required. [Job ID: NOK-SVII-SW]

Mail resume to:
Nokia Recruiter,
3575 Lone Star Cir, Ste 434,
Ft Worth, TX 76177
& note specific Job ID#.

COMPUTING PRACTICES

The iPlant Collaborative: Cyberinfrastructure to Feed the World

Dan Stanzione, *University of Texas at Austin*



As plant biology becomes a data-driven science, new computing technologies are needed to address many formidable challenges. The iPlant Collaborative provides cyberinfrastructure for researchers and developers to collaborate in creating better tools, workflows, algorithms, and ontologies.

The importance of computation for scientific discovery is well-established, and funding agencies and organizations have been investing in large-scale computation to support scientific research for decades. One of the most advanced such efforts is the Extreme Science and Engineering Discovery Environment (XSEDE), formerly the TeraGrid, a five-year, \$120 million project supported by the National Science Foundation (NSF) that provides thousands of researchers around the world with access to 16 supercomputers and advanced digital resources.

The iPlant Collaborative (www.iplantcollaborative.org) represents a new approach to large-scale investments in scientific computation in several ways. iPlant is among the first NSF-funded cyberinfrastructure (CI) projects¹ to tackle truly data-driven, rather than simulation-driven, science. In addition, rather than providing researchers with one large supercomputing system, database, or software tool, the project offers a comprehensive CI architecture for future plant science. Finally, rather than being tasked with tackling a specific scientific question, iPlant uses a unique synthetic approach to define a community-driven set of grand challenges to address through the CI. The iPlant project is relevant to computing professionals in multiple ways, as plant science presents some enormously challenging computational problems.

Given increasingly limited science funding and the rich history of computation-driven successes in so many other scientific fields, why make this computational investment in plant biology now? The answer is threefold.

First, as the sidebar “The Importance of Plants” describes in more detail, plants are critical to a sustainable future: supporting an estimated world population of 9.3 billion by 2050 will require doubling food production.² Second, plant biology is at a stage where computation will be key to delivering new scientific results. For example, current technology trends in DNA sequencing enable a single, small laboratory to produce one terabyte of sequence data every few days, and this rate is increasing at a pace that easily exceeds Moore’s law.³ Third, plant biology as practiced today is a data-driven endeavor⁴ and thus can be a model for the deployment of CI, which has long served simulation-driven sciences, for data-driven computational sciences.

COMPUTATIONAL CHALLENGES IN PLANT BIOLOGY

Much to the surprise of many in the field, plant biology has indeed become a computational science. However, it is radically different from the “traditional” computational sciences such as fluid dynamics, materials science, and astrophysics, in which the creation and execution of

algorithms representing mathematical models of the underlying physical systems embodies the computational aspect of the science. These simulations can require enormous scale, and in fact drive the development of supercomputers. The inability to deliver a sufficient amount of floating-point capability to simulation codes is the primary barrier to scientific progress.

In plant biology, and in the life sciences in general, the picture is somewhat different. Certain aspects of biology, like protein folding and molecular dynamics, require the same large-scale solutions of equations as other computational fields. In most situations, however, there are no known equations that can model biological systems.

The vast amount of information known about genomes is usually represented in text as genetic sequences. Genomes encode a host of features including genes, most of which encode proteins. Proteins, in turn, participate in metabolic pathways and regulatory networks, which eventually lead to the myriad diversity of form and behavior observed in plants and other organisms. While no set of differential equations predicts these associations, statistical methods offer a means to connect genomic, metabolomic, proteomic, and pathway information to organism characteristics, thereby revealing the organizing principles of living systems.

The barriers to constructing explanatory and predictive models in plant science fall roughly into three categories: data volume, computational complexity, and data integration.

Data volume

In the 1980s, a reasonable PhD project in plant science was to clone a gene, which on average represented 2,000 bases of DNA—the information-encoding strings of bases that make up an organism's genome. Thirty years later, graduate students sequence and assemble genomes comprising millions or billions of DNA bases. Instead of measuring levels of DNA-encoded messenger RNA for single genes using physical means called Northern blots (a few Kbytes of image data per experiment), they have moved to measuring the expression of all genes at once, first using microarray technology (5-10 Mbytes) and most recently DNA sequencing (25-100 Gbytes).

The capacities of DNA sequencing technology have grown superexponentially over the course of the past decade, leaping ahead of the ability of Moore's law to keep pace with the increased computing demand for all this data.

THE IMPORTANCE OF PLANTS

Plants sustain nearly all other forms of life on Earth. Photosynthesis—the hallmark energetic process of plants and blue-green algae—accounts for nearly 98 percent of the planet's atmospheric oxygen.

In addition to providing the basis for our entire food chain, plant products also account for most of our clothing and contribute a substantial amount of our construction material. A strong argument can be made that our mastery of breeding plants through agriculture was the basis for the rise of modern civilization from its hunter-gatherer origins. Advances in the technology through which humans cultivate plants has, particularly in the past century or two, dramatically reduced the fraction of the population involved in food production, freeing the rest of us to develop other things, like software.

Unfortunately, it is a bad time for progress to slow. As the population continues to increase, urbanization is reducing available cropland, and climate change is putting productive cropland at risk. In many regions of the world, supplies of fresh water for irrigation are becoming more limited.¹ And, perhaps most importantly, the protein-rich Western diet—which is sustained by cattle and poultry fed with plants—is spreading around the globe.

To get some idea of the magnitude of this issue, consider the world's three staple crops: rice, wheat, and corn. Together, these three grasses make up 70 percent of all agricultural production. In Europe, where the staple crop is wheat, about 1 pound of wheat must be produced per person per day. In Asia, about the same amount of rice is needed. In the US, however, about 3.5 pounds of corn are produced per person per day to sustain the protein-intensive diet.

The unintended consequences of the globalization of a protein-rich diet and other aspects of US lifestyle are putting tremendous stress on already-fragile food markets. As recently as 2008, spikes in the price of food staples, largely attributed to increased biofuel production, led to food riots in developing countries, and the rise in commodity prices in early 2011 suggests that additional food security unrest is imminent.² Overcoming these challenges will require vast improvements in plant production, just as current methods are showing diminished returns.

A partial answer lies in adopting computational approaches to advance a more evidence-driven model for decision making, both at the scientific and policy level. We know that plants can be more productive than they are now. There are plants that thrive with little water or in poor soil with little fertilizer. What are not yet well understood are the mechanisms to predictively engineer plants that will provide optimal yields in suboptimal conditions. Inexpensive and easily attainable gene sequencing, comprehensive metabolic and biochemical assays, and image-based phenotyping are starting to provide the data to help understand these phenomena, but it is a lot of data.

References

1. P. Rogers, "Facing the Freshwater Crisis," *Scientific Am.*, Aug. 2008, pp. 46-53.
2. J. Von Braun, "Rising Food Prices: What Should Be Done?," *EuroChoices*, Aug. 2008, pp. 30-35.

When a researcher collects a plant species in the Amazon River Basin, the half-terabyte of data representing its raw, unassembled genome sequence can be available for analysis the following week. It is not just in DNA sequencing, though, that data volumes have increased dramatically: comprehensive measurements of protein levels in a cell result in hundreds of Gbytes of mass spectrophotometric data per sample, and image-based phenotyping (trait measurement) analyses can generate millions of frames of multichannel photographic data or smaller numbers of extremely high-resolution photographs.

COMPUTING PRACTICES

iPLANT CYBERINFRASTRUCTURE ORIGINS

The iPlant CI design rests on principles expounded at the National Science Foundation-sponsored workshop “History and Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures” (<http://deepblue.lib.umich.edu/handle/2027.42/49353>). As defined by this workshop, “Cyberinfrastructure is the set of organizational practices, technical infrastructure and social norms that collectively provide for the smooth operation of research and education work at a distance. All three are objects of design and engineering; a cyberinfrastructure will fail if any one is ignored.” Thus, iPlant is not only a technical product, but also a virtual organization.

One of the project’s unusual aspects is that it was created not to address a particular issue but rather to address questions designated by the plant science community after the project began. The initial task was to define and prioritize the grand challenges. After a kick-off conference in April 2008, iPlant hosted a series of workshops throughout the year that resulted in the submission of a half-dozen formal proposals to its board of directors. In April 2009, the iPlant board announced that the project would proceed in developing CI around two grand challenges:

- constructing the phylogenetic tree for all green plant species (iPTOL, the iPlant Tree of Life), and
- clarifying the relationship between genotype information and expressed phenotypes in plants (iPG2P, iPlant Genotype to Phenotype).

iPlant assembled working groups consisting of project staff and members of the plant science community to define the CI required for each of these challenges. The CI’s goals were to leverage both the existing base of bioinformatics tools and the physical resources of existing cyberinfrastructures—primarily the TeraGrid (now XSEDE).

iPlant’s role was to create discovery environments, portals through which researchers could collaborate and make use of the large-scale data and resources available. The CI provides software and interface layers to unify these tools and offers transparent access to these resources via consistent interfaces, use of standards, and, critically, integration of disparate datasets and types. The CI created based on these requirements consists of a core architecture and extensions to provide the functionality described in the two grand challenges.

Computational complexity

The storage, transmission, and basic manipulation of large datasets is challenging enough, but running even simple algorithms on them can involve significant computational requirements. For example, successfully assembling a plant genome from billions of subsequences (fragments) generated during the DNA sequencing process often requires constructing and traversing a Tbyte-size de Bruijn graph, at hundreds of hours of CPU time per iterative attempt.⁵ Interpreting proteomic datasets requires applying sophisticated signal-processing and pattern-recognition algorithms to half-Tbyte-size spectrophotometric datasets. The large sizes of datasets in image-based analysis methods amplify the inherent difficulties in automated feature extraction. And while experiments that test interactions between genetic variants and quantitative traits such as plant height or grain yield start with modest datasets, due to the requirement to consider combinatorial factors, they end up with estimated compute requirements that would tax modern cluster systems.⁶

Data integration

A related issue complicating the computational space in plant biology is that bioinformaticians develop algorithms as needed using various software and hardware platforms, often with minimal standards guidance. These algorithms in turn import, export, and are configured using bespoke file and data formats as well as unique invocation semantics.

Researchers often store and analyze biological data close to the physical site where it was collected, leading to the proliferation of databases. There are more than 1,300 molecular biology databases that represent dozens if not hundreds of site-specific query and representation interfaces. The operational and data integration requirements for assembling composite datasets, chaining together collections of analytical steps, and, most importantly, stably reproducing these scientific workflows are often even more technically challenging than the specific analyses under consideration.

Attempts by the biology community to establish common standards abound in the form of ontologies, data standards, abstracted workflow managers, and so on, but the barrier to adoption is high due to the cost of back-porting existing interfaces and to the resource requirements for hosting high-quality, Web-accessible services.

iPLANT CYBERINFRASTRUCTURE ARCHITECTURE

Researchers created iPlant to leverage the enormous existing investments in biological data collection and bioinformatics tools, and to provide the CI necessary to enable the plant biology community to address grand challenge questions. The “iPlant Cyberinfrastructure Origins” sidebar describes the CI’s evolution.

The iPlant CI is community-driven, with informaticians and scientists collaborating in both its design and implementation. The project’s staff consists of a multidisciplinary team of computing, information science, and biology professionals at the University of Arizona (UA), the University of Texas at Austin, and Cold Spring Harbor Laboratory in New York, as well as dozens of working group participants from institutions around the world.

Figure 1 shows the architecture of iPlant’s CI, which was designed to meet the requirements of the first two grand challenges: iPlant Tree of Life (iPTOL) and iPlant Genotype to Phenotype (iPG2P). It became clear at the outset that because the problems, data, and methods were not static, the CI had to be able to change and evolve. To meet many users’ needs, accommodate large volumes of data, and quickly execute complex algorithms, the CI is built on the

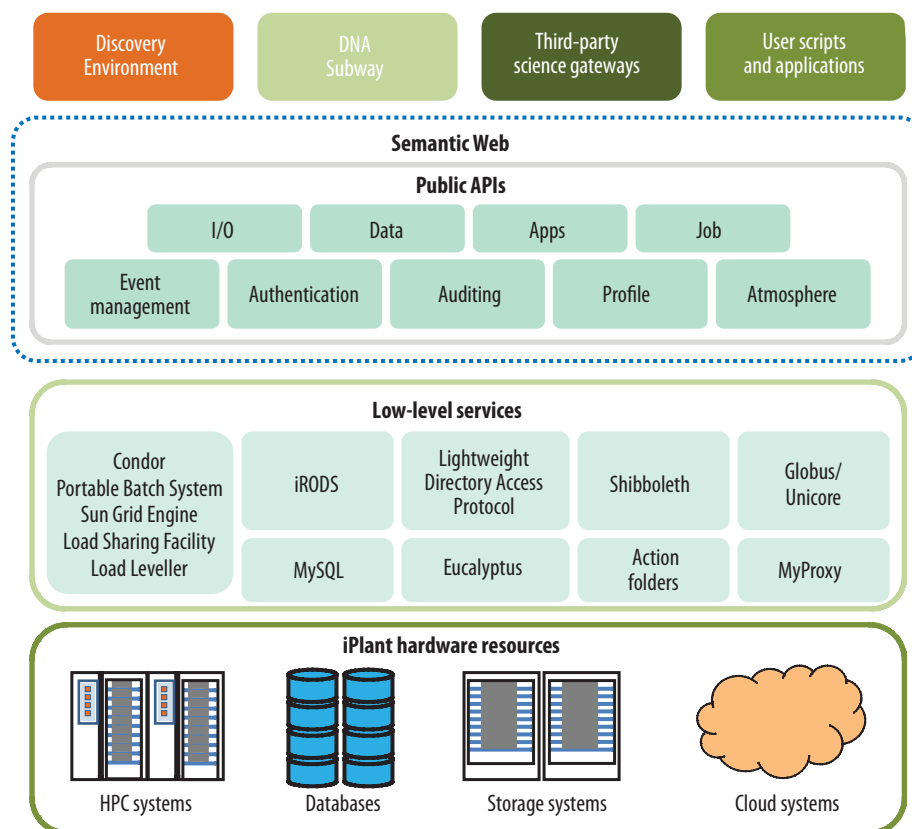


Figure 1. iPlant cyberinfrastructure architecture.

fundamental architectural principles of modularity, flexibility, openness, adaptability, and scalability.

Hardware

As with all cyberinfrastructure, hardware is at the core of iPlant's CI. The project uses not only systems deployed specifically for iPlant, but also the hardware resources of XSEDE and the NSF's open CI. Data replicated between UT Austin's Texas Advanced Computing Center (TACC) and UA via the integrated Rule-Oriented Data System (iRODS)⁷ provides a robust, reliable repository. While TACC supplies large-scale cluster computing capability, UA offers a separate cluster with virtual machine (VM) hosting and management. This cluster underlies iPlant's Atmosphere cloud computing platform, providing virtual hosting for a range of bioinformatics Web applications and user portals, as well as hosting the iPlant Web environment itself.

Middleware

Residing above the hardware layer is a rich and growing middleware layer that provides powerful abstractions through which researchers can access iPlant's compute and data resources. The application programming interface (API) is the foundation upon which iPlant's interfaces are

constructed, and bioinformaticians can use it to embed iPlant resources in their own scripts and tools.

Applications

Atop the middleware layer is the application layer, which includes a range of interfaces. The primary graphical interface for users is the iPlant Discovery Environment. The DE builds upon the API layer to ensure a rich user experience with a consistent graphical metaphor and provides additional intelligence through comprehensive tracking of provenance for reproducibility of experiments, support for collaboration among users, and extensibility through the incorporation of new tools, datasets, and workflows. An alternate interface to the iPlant CI is the DNA Subway, which is designed to serve educators and students by presenting a series of useful, teachable genomics workflows in a colorful, metaphor-oriented environment.

User access across layers

User access to the iPlant CI exists at multiple layers in the architecture. The project supports a wide range of modes for users to access the system, from moving a complete workflow to the rich Web clients to simply using iPlant as a file repository. Expert users can directly access the com-

COMPUTING PRACTICES

pute resources, which could mean command-line access to supercomputers, log-in access to VM images hosted on Atmosphere, or FUSE (Filesystem in Userspace)-mountable access to file systems. Programmers or script authors can access resources through the APIs by embedding Web requests into their code. Many large labs have bioinformaticians, who can embed API commands to retrieve data or perform large-scale computations through the iPlant CI. More typical end users can access their data, tools, or workflows through the DE or an alternate Web portal.

APPLICATION PROGRAMMING INTERFACES

The iPlant project has taken a layered approach to deploying APIs for interacting with the underlying CI. Foundational APIs present a generic set of representational state transfer (REST) interfaces to basic actions like file and data operations, authentication, application integration, compute job invocation, and event monitoring. Atop these RESTful APIs is the semantic services interface as well as interfaces for metadata-driven workflow construction and orchestration, extraction and management of metadata, and interactions with federated data sources.

The iPlant project has taken a layered approach to deploying APIs for interacting with the underlying CI.

Foundational APIs

With the RESTful services that comprise iPlant's foundational APIs, properly credentialed users can perform atomic series of operations to construct complex workflows. These APIs are all exposed via HTTP and can be consumed by rich Internet clients such as the iPlant DE, which serves as a flagship demonstration of the project's underlying technologies; workflow management applications such as Taverna and Kepler; third-party Web applications; other RESTful Web services; or user scripts. The DE exposes its own APIs for developing visualization and editor plug-ins, data type extensibility, and automated interaction.

I/O. Similar in implementation and intent to the Dropbox API (www.dropbox.com), iPlant's I/O API presents the underlying iRODS distributed file system as a \$HOME directory for users that is available across all interfaces. The interface lets users directly import, export, and organize their files as well as manage file and directory permissions. The interface also allows fetching remote resources, staging them into user storage space, and pushing user files to remote Internet-accessible locations. Behind the scenes, the I/O API optimizes access to stored data to minimize data movement during execution of computational jobs.

Data. The data API translates file formats, describing them via simple metadata that includes information about their semantic context. Format developers provide computer code for performing pairwise translation to other formats and versions. This design allows iPlant to support the state of the art, wherein file types lack a common semantics, while working toward providing a set of translators in the future that is based on a unified semantic model instead of pairwise translation. The API can perform data operations in-place, reducing the need to move files around just to translate them to other forms.

Apps and Job. Bioinformatics application developers and consumers alike can take advantage of the Apps and Job APIs, which provide a consistent, rational, REST-based interface for describing an application's properties and invocation parameters, identifying applications with specific properties or capabilities, and running instances of those applications on high-performance computing (HPC) resources. Under the iPlant model, developers build applications using existing code, wrap them in a thin arbitration shell, and deploy them on the iPlant shared-file system.

Developers use a simple JSON (JavaScript Object Notation)-based metadata language to describe parameters for invoking a particular program. The language also includes hooks for annotating each parameter according to its semantic type. They submit this metadata via HTTP POST to the Apps RESTful endpoint, where the application it describes becomes discoverable to other users. Developers can also apply fine-grained permissions, based on access control lists, to applications to describe their discoverability and invocation characteristics, an essential feature for creating sharable scientific tools and workflows.

Users access the Apps endpoint to search for programs to run and receive detailed, programmatically interpretable information about how to invoke them via the Job API. The state of a running job, as well as its outputs and submission metadata, are accessible via unique Job service URLs. The resulting files can be automatically staged back to the user's \$HOME directory, where they will become available via the I/O API.

The Apps and Job APIs interact with the higher-level semantic API to allow automatic creation of resource description and resource invocation graphs, so that all applications developed under this schema are semantically discoverable and usable. The upshot is that researchers can discover and use myriad bioinformatic analysis applications via a single, easy-to-learn interface that is compatible with today's advanced Web-based application technologies.

Other APIs. Other public APIs include an event management system that permits both users and consumer applications to publish and subscribe to notifications about the status of various activities in the iPlant CI, thus orchestrating activities without explicit polling; an authentication service that serves up tokens allowing federated access

to services without explicitly transmitting credentials; an auditing service to allow tracking of resource consumption and access patterns; and a profile service that delivers computer-readable summaries of user profile data.

Another low-level interface is the Atmosphere virtualization service. Via this API, users and developers can provision, instantiate, interact with, and snapshot VMs running within the iPlant computing cluster environment. These VMs have native access to the iPlant authentication subsystem, storage layer, and high-performance local networking environment.

Semantic API

Developers use the Simple Semantic Web Architecture and Protocol to describe iPlant's low-level services.⁸ SSWAP offers a data-driven, service-discovery technology that employs on-demand, transaction-time reasoning for data discovery and semantic match-making with suitable semantic services exposed through the other API components, such as matching the output of an assembly run with suitable tools for assessing quality metrics. SSWAP utilizes both optimized knowledge bases and peer-to-peer interactions.

Developers use JSON to implement SSWAP via an HTTP API. Although this technology, which is widely known and adopted by Web masters and programmers internationally, is easy to learn, it is not semantically aware. The HTTP JSON API thus uses another, deeper Java API, written in 2011, that allows finer control for expert developers. The Java API wraps the SSWAP and OWL (Web Ontology Language) specifics, combining semantic power with the ease of a JSON interface.

In addition to representing a clear example of migrating research into production, to our knowledge, iPlant is the first project to wrap a complete semantic Web services capability in an easy-to-access, Web-based API. The code is available at <http://sswap.svn.sourceforge.net/viewvc/sswap/API-SDK/release>.

iPLANT IN ACTION: GRAND CHALLENGE EXAMPLES

A complete description of the problems, tools, and datasets surrounding iPTOL and iPG2P is beyond the scope of this article. However, a few examples demonstrate how the iPlant CI in development is already beginning to influence plant biologists.

iPTOL

Plant scientists want to know the evolutionary relationships of all green plant species on Earth so that they can use this knowledge to understand and isolate the origins of critical traits in plants. Two areas where the iPlant CI can have an impact in meeting this challenge are in scaling up phylogenetic inference methods and in data visualization.

Phylogenetic analysis of up to half a million plant species represents a scalability challenge that iPlant is addressing on two parallel tracks.

The general approach is to optimize existing analytical methods such as maximum likelihood estimation with RAxML (<http://sco.h-its.org/exelixis/software.html>) and neighbor joining with NINJA and WINDJAMMER.⁹

Employing very large data matrices and measuring uncertainty using bootstrap replicate analysis make the ultimate goal of building phylogenetic trees for so many species a formidable challenge that demands using HPC methods to perform and update the analyses in tractable amounts of time. Improvements in existing code range from implementation of checkpointing, to parallelization, to refactoring in HPC-friendly languages. The initial runtime estimates were hundreds of days, with requirements for hundreds of gigabytes of RAM, but distributed memory approaches now allow for arbitrarily large problem sizes and are showing speedups of over 100×.

iPlant is the first project to wrap a complete semantic Web services capability in an easy-to-access, Web-based API.

Phylogenetic visualization involves the viewing, manipulation, and annotation of very large tree data structures. iPlant has developed a viewer that inputs descriptions of up to 500,000 leaf nodes, their labels, and edge lengths and displays an interactive tree. Users can browse, zoom, select, search, and annotate, all while preserving the input information. The browsing function captures the phylogeny's overall size while keeping the displayed tree, labels, and lengths readable. Zooming reveals additional information, not simply changes in a static image's scale. Users can select single nodes or groups of nodes and add annotations such as text labels, colors, or images either manually or from a file. Figure 2 shows sample screenshots from the viewer.

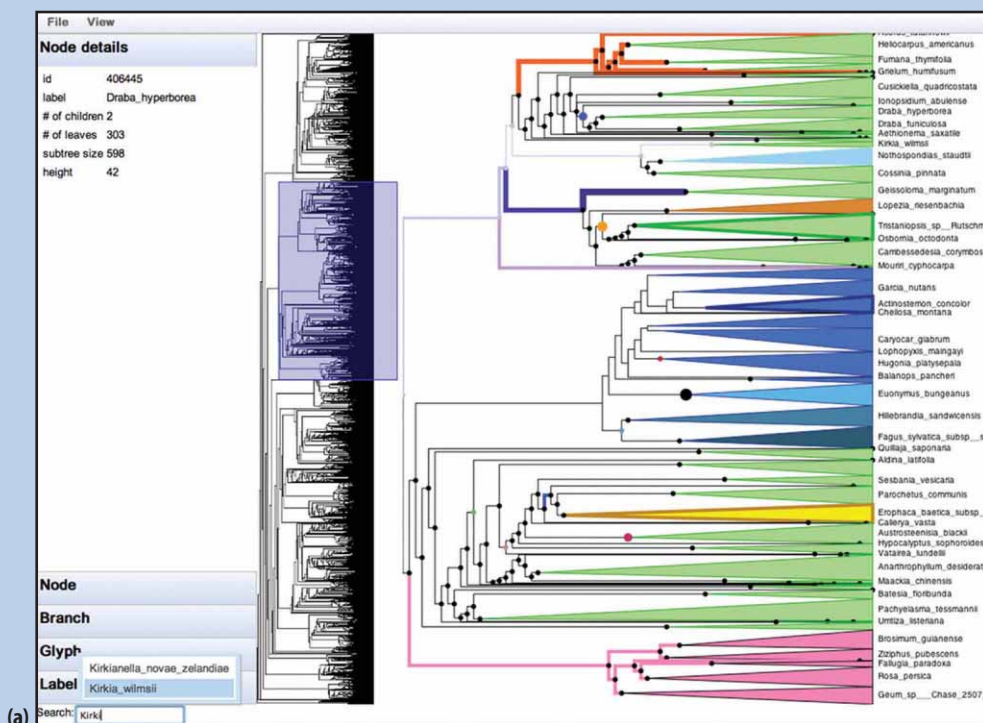
In addition to studying gene and plant functions, the iPTOL CI enables use of data on plant relatedness and other aspects of plants such as drought tolerance to help engineer plants for future world needs.

iPG2P

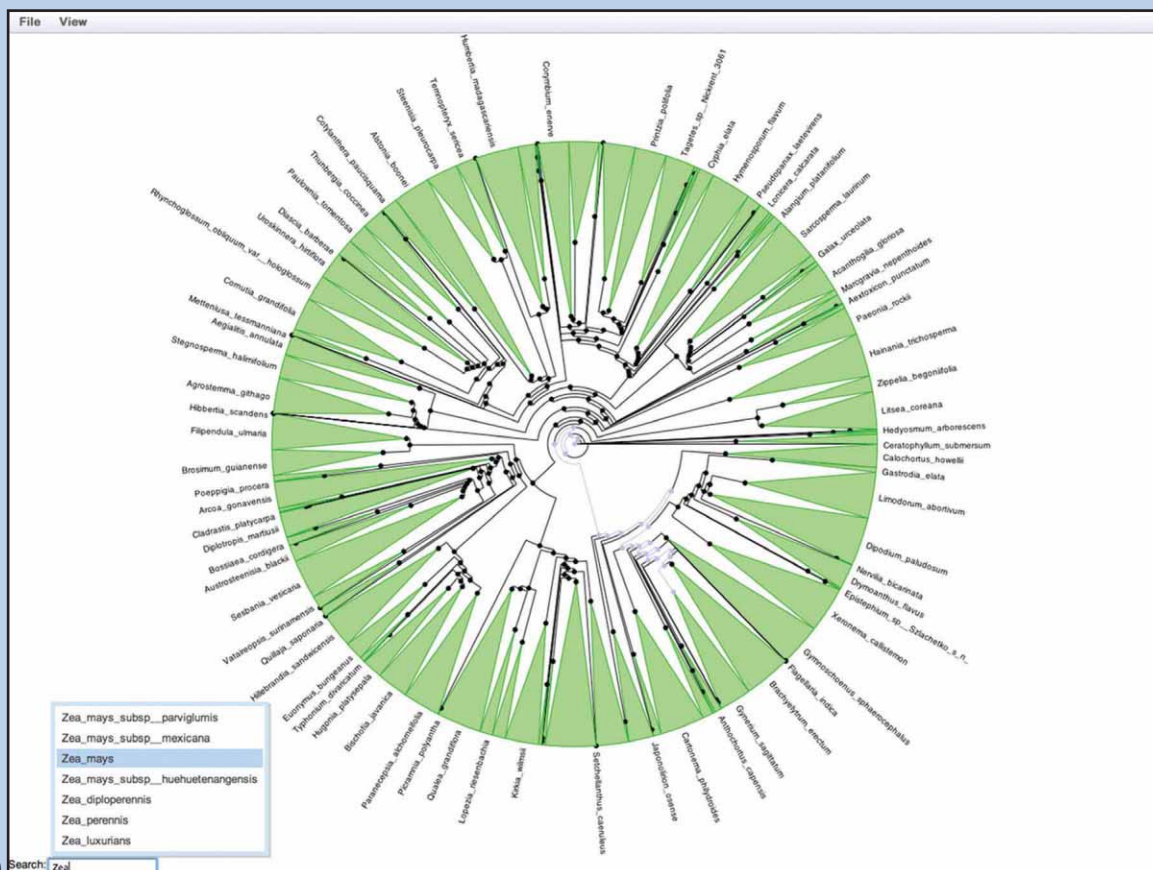
Successfully linking genotype to phenotype, which enables plant scientists to predict the outcomes of specific genomes in nonconstant physical environments, is a complex challenge.

DNA sequence analysis. Through the DE, iPlant delivers sophisticated DNA sequencing-analysis applications as user-friendly Web applications. Users can run these

COMPUTING PRACTICES



(a)



(b)

Figure 2. Prototype tree viewer showing a subset of 100,000 species from the National Center for Biotechnology Information (NCBI) taxonomy. The two different views illustrate (a) semantic zooming and (b) metadata-driven visual effects.

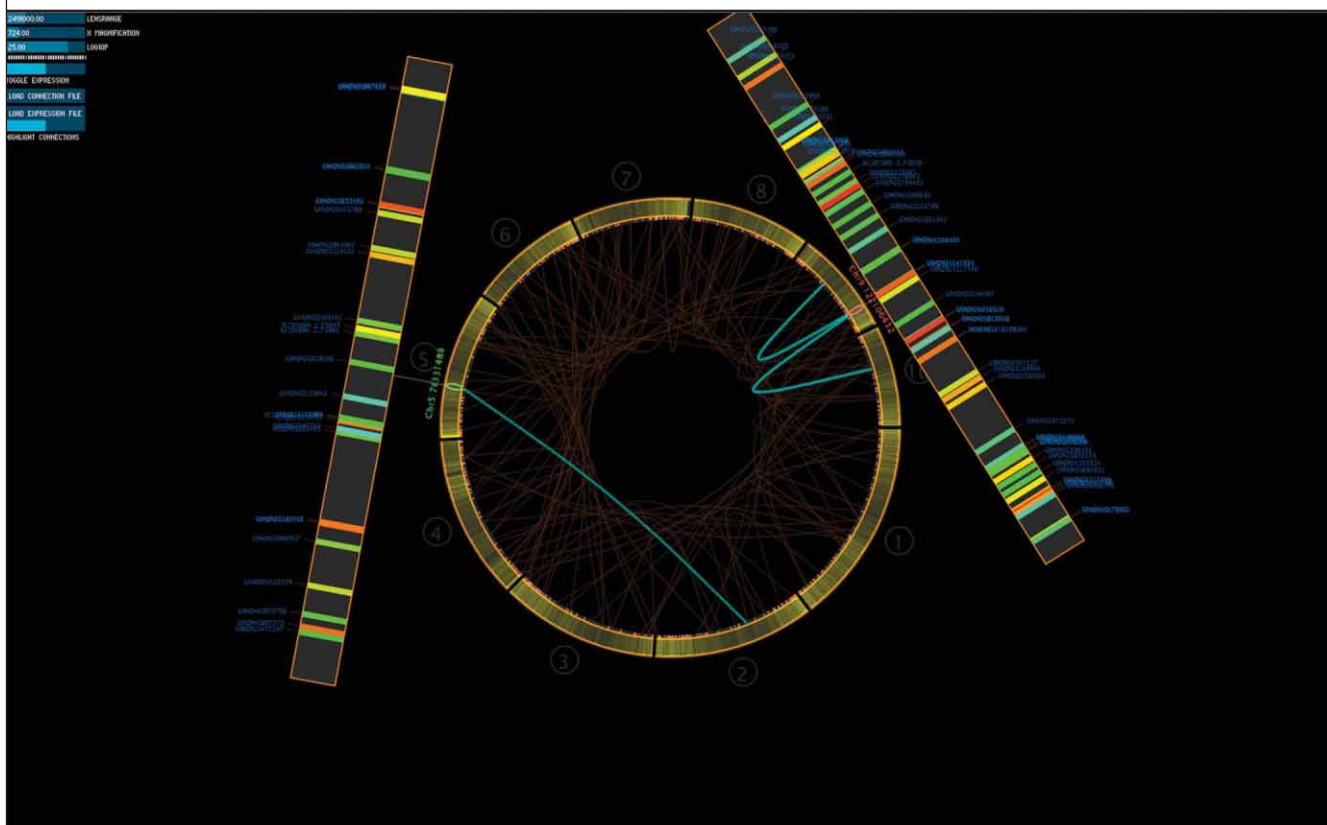


Figure 3. Interactive map of potential epistatic interactions among genetic loci in the corn genome.

analyses on any species for which a genome sequence is available, as all higher-plant genome sequences are federated into a unified Genome Services available to all iPlant-hosted applications. In fall 2011, iPlant will provide the ability to perform de novo genome assembly and run basic genome annotations on the resulting DNA sequences from within the DE, thereby allowing users to build and define their own genomes.

Much of the complexity of these applications is hidden from the user: iPlant ameliorates data-size issues via high-bandwidth intersite networking and high-performance parallel file systems; addresses the computational requirements by parallelizing common algorithms and running them on large-scale supercomputing resources; and handles integration, in part, by exposing the applications in a common API framework and graphical representation, freeing users to focus on scientific objectives. The project is actively soliciting developers of new sequencing-oriented applications to deploy usable and scalable instances of their tools, with the objective of providing a best-of-class experience for sequencing-based science applications.

G2P mapping. Independent teams of iPlant researchers are developing optimized implementations of algorithms for statistically linking genotypes and phenotypes. The formulation of the general linear model that can be applied to G2P mapping contains significant amounts of linear algebra,

and thus is amenable to parallelization. The teams are using fixed-effect linear models to develop a GPU-enabled version for use on mid- to large-size mapping datasets to model genotype-phenotype interactions, offering a several-fold speedup over a traditional $\times 86$ implementation. Future iterations of this application will take advantage of the now commonplace GPGPU (general-purpose computation on graphics processing units) clusters.

Gene interaction. Another area of interest is in determining cooperative interactions between genes to create phenotypes. This requires pairwise interaction tests among all genes, making such analyses computationally intensive. For example, testing all pairwise interactions of 32,768 genes at a rate of a gene per second would require $5.4E + 8$ seconds (17 years).

Fortunately, this problem is manageable using multi-threading and coarse-grained parallelism. In collaboration with Iowa State University geneticists and statisticians, iPlant has developed an application that runs such analyses in only a few hours on a portion of one of TACC's clusters. Figure 3 shows a screenshot from this application—in this case, an interactive map of potential epistatic interactions among genetic loci in the corn genome.

Image management. A major remaining bottleneck in the G2P space is the production of phenotype measurements. This often labor-intensive process involves

COMPUTING PRACTICES

collecting, storing, and interpreting large quantities of image or video data. Ideally, the burden on plant science researchers could be eased by building a robust platform for image data warehousing and search and coupling this with a plug-in architecture for automated algorithmic feature extraction.

Building on existing work by iPlant collaborators Edgar Spalding and B.S. Manjunath, the project is developing the PhytoBisque Image Analysis Environment. This standalone 5D image management system and analysis workflow platform is designed to take advantage of iPlant's virtualization and capacious data storage systems. Hooks between PhytoBisque and the DE permit image data and results to flow between the two applications, and PhytoBisque leverages the local iPlant execution environment to offer scalable execution of image processing code written in Matlab, Java, Python, and C. A plug-in API facilitates development of new processing algorithms, and a Web-based templating system enables complex, multipane image analysis workflows. With this platform in place, iPlant is working to match plant scientists with image processing needs to computer scientists and machine vision specialists interested in the types of images under consideration.

As the iPlant CI matures, it could become the premier gathering place for the plant science community, providing numerous unique research opportunities. Biologists and computational scientists will be able to collaborate in developing better tools, workflows, algorithms, and ontologies, and the CI itself will provide the necessary support for data conversion and output handling that researchers normally would need to build into a standalone tool.

Plant science is a fantastic incubator for biomedical applications. Most data is freely available, and the challenges of phylogenetics or genome assembly are largely the same as those for the study of human disease. The iPlant CI could become a proving ground for novel applications that one day translate to many other life sciences.

Despite iPlant's promising start, much work remains to be done. The project is eager for developers to integrate their products with the DE and through the iPlant APIs. In addition to online resources, numerous workshops, seminars, "hack-a-thons," and other outreach efforts provide hands-on training at research and academic institutions around the US to facilitate user and developer contributions (www.iplantcollaborative.org/learn/eot).

Plant biology also is in need of much better computing techniques as it transforms into a truly data-driven science. Data mining, scalable algorithms, data integration, and information visualization are just a few of the areas where computing professionals can make significant contributions to the state of the art—and in the process help to feed the world. **□**

Acknowledgments

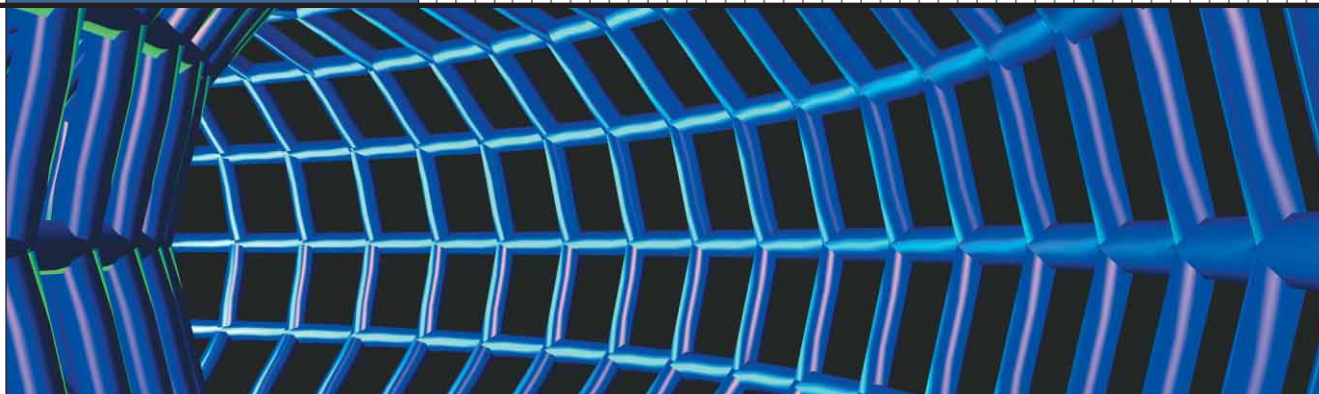
The iPlant Collaborative is funded by grant no. DBI-0735191 from the National Science Foundation Plant Cyberinfrastructure Program. This article represents the collective efforts of the entire iPlant Collaborative staff. More than CI, iPlant is about people, and without this extremely talented and dedicated group of individuals, there would be nothing to write about. This article especially benefitted from the contributions of Stephen Goff, Matthew Vaughn, Sheldon McKay, Eric Lyons, Nirav Merchant, and Naim Matasci.

References

1. NSF Cyberinfrastructure Council, "NSF's Cyberinfrastructure Vision for 21st Century Discovery," Nat'l Science Foundation, 20 Jan. 2006; www.nsf.gov/od/oci/ci_v5.pdf.
2. Food and Agriculture Organization of the United Nations, *The State of Food and Agriculture: Livestock in the Balance*, 2009; www.fao.org/fileadmin/user_upload/animalwelfare/SOFAe.pdf.
3. K.A. Wetterstrand, "DNA Sequencing Costs: Data from the NHGRI Large-Scale Genome Sequencing Program," 4 Feb. 2011, Nat'l Human Genome Research Inst.; www.genome.gov/sequencingcosts.
4. M.Y. Galperin and G.R. Cochrane, "The 2011 Nucleic Acids Research Database Issue and the Online Molecular Biology Database Collection," *Nucleic Acids Research*, Jan. 2011, pp. D1-D6.
5. D.R. Zerbinio and E. Birney, "Velvet: Algorithms for De Novo Short Read Assembly Using de Bruijn Graphs," *Genome Research*, May 2008, pp. 821-829.
6. L. Koersterke et al., "An Efficient and Scalable Implementation of SNP-Pair Interaction Testing for Genetic Association Studies," *Proc. 2011 IEEE Int'l Parallel and Distributed Processing Symp. Workshops and PhD Forum* (IPDPSW 11), IEEE, 2011, pp. 523-530.
7. Arcot Rajasekar et al., *iRODS Primer: Integrated Rule-Oriented Data System*, Morgan & Claypool, 2010.
8. D.D.G. Gessler et al., "SSWAP: A Simple Semantic Web Architecture and Protocol for Semantic Web Services," *BMC Bioinformatics*, 23 Sept. 2009; www.biomedcentral.com/1471-2105/10/309.
9. T.J. Wheeler, "Large-Scale Neighbor-Joining with NINJA," *Proc. 9th Workshop Algorithms Bioinformatics* (WABI 09), LNBI 5724, Springer, 2009, pp. 375-389.

Dan Stanzone is codirector of the iPlant Collaborative and deputy director of the Texas Advanced Computing Center at the University of Texas at Austin. His research focuses on parallel programming, scientific computing, bioinformatics, and system software for large-scale systems. Stanzone received a PhD in computer engineering from Clemson University. He is a member of IEEE, the IEEE Computer Society, and the American Association for the Advancement of Science. Contact him at dan@tacc.utexas.edu.

PERSPECTIVES



Defending against Buffer-Overflow Vulnerabilities

Bindu Madhavi Padmanabhuni and Hee Beng Kuan Tan
Nanyang Technological University, Singapore

A survey of techniques ranging from static analysis to hardware modification describes how various defensive approaches protect against buffer overflow, a vulnerability that represents a severe security threat.

In 2003, an analysis of buffer overflow pronounced it the vulnerability of the decade.¹ The following year, the Open Web Application Security Project (OWASP) listed it as the fifth most serious Web application weakness. In the first five months of 2010, the National Vulnerability Database (<http://nvd.nist.gov>) recorded 176 buffer overflow vulnerabilities, of which 136 had a high severity rating. Buffer overflow remains a major security hole today, ranking third on the Common Weakness Enumeration/SANS list of Top 25 Most Dangerous Software Errors (<http://cwe.mitre.org/top25>).

Buffer overflow occurs during program execution when an application writes beyond the bounds of a preallocated fixed-size buffer. This data overwrites adjacent memory locations and, depending on what it overwrites, can affect program behavior. The lack of bounds-checking operations for filling the buffers permits this error. Applications written in programming languages such as C or C++ are commonly associated with buffer-overflow vulnerabilities because they allow overwriting any part of memory without checking whether the data written will overflow its allocated memory.

A review of buffer-overflow exploits and an analysis of their solutions reveals deficiencies in present defenses, providing a basis for developing modifications to protect against such exploits.

BUFFER-OVERFLOW EXPLOITS

Attackers can use buffer overflows to launch denial-of-service (DoS) attacks, spawn a root shell, gain higher-order access rights (especially root or administrator privileges), steal information, or impersonate a user. In 1998, the Morris worm, one of the first to strike the Internet, exploited a buffer overflow in the Unix finger daemon (fingerd) to propagate itself from one machine to another (http://en.wikipedia.org/wiki/Morris_worm). The 2001 Code Red worm took advantage of the same weakness in the Microsoft IIS webserver and reportedly infected 359,000 systems within 14 hours; (see [http://en.wikipedia.org/wiki/Code_Red_\(computer_worm\)](http://en.wikipedia.org/wiki/Code_Red_(computer_worm))). In 2003, SQL Slammer exploited a buffer overflow in the Microsoft SQL server, spread quickly, and launched a DoS attack on various targeted networks (http://en.wikipedia.org/wiki/SQL_slammer).

PERSPECTIVES

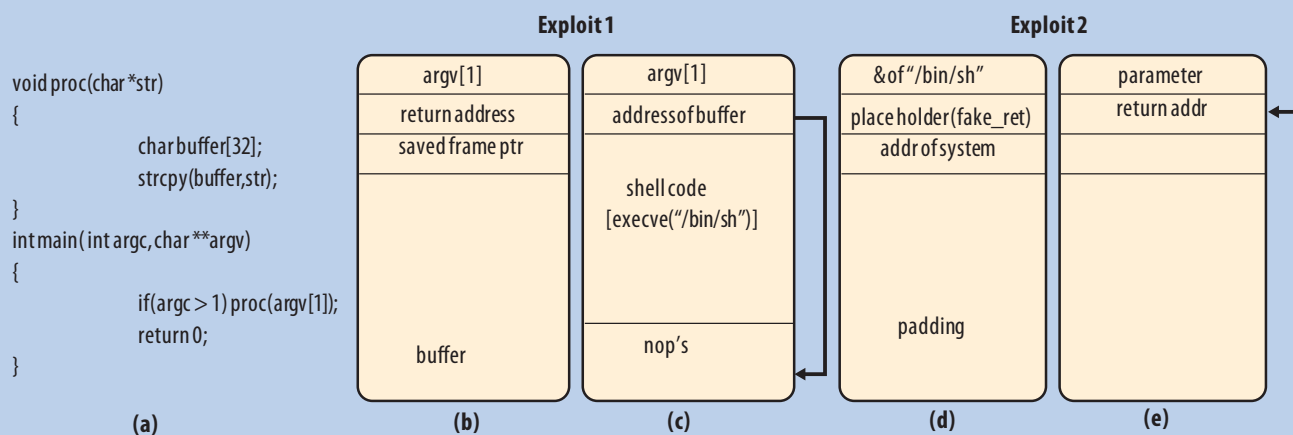


Figure 1. Examples of function activation record exploits. (a) Vulnerable code snippet. If str is longer than 31, it modifies the memory area next to it. (b) Stack frame of proc before strcpy(buffer, str). (c) Stack frame after strcpy(buffer, str). Exploit 1 uses a string consisting of shell code and the memory address to which the attack code copies the shell code to cause a buffer overflow. When the function returns, it jumps to the shell code to spawn a shell. Exploit 2 is a return-to-libc attack on the code snippet that spawns a shell by overwriting the return address with the address of system(). (d) Stack frame after buffer overflow and before program returns to system(). The attack code includes the address for system() as well as its parameters. (e) Stack frame after program returns to system().

To carry out an exploit, attackers must find suitable code to attack and make program control jump to that location with the required data in memory and registers. Attackers glean information about the vulnerable program code and its runtime behavior from the program's documentation and source code, by disassembling a binary file, or by running the program in a debugger. Buffer overflow exploits generally target function activation records, pointers, or the management data of heap-based memory blocks.

Function activation record exploits

A popular technique targets the function activation record. When program execution calls a function, stack frame is allocated with function arguments, return address, the previous frame pointer, saved registers, and local variables. In the stack frame, the return address points to the next instruction for execution after the current function returns. Attackers can overflow a buffer on the stack beyond its allocated memory and modify the return address to change program control to a location of their choice. Figure 1 shows two examples of these attacks.

Instead of supplying executable code, an attacker can supply data to a C library function, such as system(), that is already present in the program code. Such exploits are called return-to-libc attacks because they direct control to a C library function rather than to attacker-injected code; they also alter the return address. Return-to-libc attacks are ideal for exploiting programs that have memory protection mechanisms, like nonexecutable stacks, because they do not execute attacker-supplied code.

Another target is the previous frame pointer. An attacker can build a fake stack frame with a return address pointing

to the attacker's code. An overflow of the previous frame pointer will point to this fake stack frame. When the function returns, the attacker's code executes.

Pointer subterfuge exploits

Pointer subterfuge exploits involve modifying pointer values, such as function, data, or virtual pointers; they also can modify exception handlers. Consider the following code snippet, which includes a buffer that has a function pointer allocated in the data section:

```

char buf[64];
int (*pfn)(char*) = NULL;
void main(int argc, char **argv)
{
    ...
    strcpy(buf, argv[1]);
    iResult = (*pfn)(argv[2]);
    ...
}
    
```

An attack can use strcpy() to overflow the buffer and overwrite the function pointer. The overwritten pointer can point to the address of shell code or system(). The attack takes place when the program calls the function pointer.

Attackers can use pointer subterfuge in overruns of stacks, heaps, or objects containing embedded function pointers. This kind of attack is especially effective when the program uses methods for preventing return address modification because it does not change the saved return address.

An exploit can use data pointers to indirectly modify the return address. Such indirect overwriting schemes are

Example exploit

```

classVulnerable : public SomeBase
{
public:
    char buffer[100];
    virtual void Func1();
    virtual void Func2();
}
void main()
{
    ...
    Vulnerable v;
    std::cin>>v.buffer;
    v.Func1();
    ...
}

```

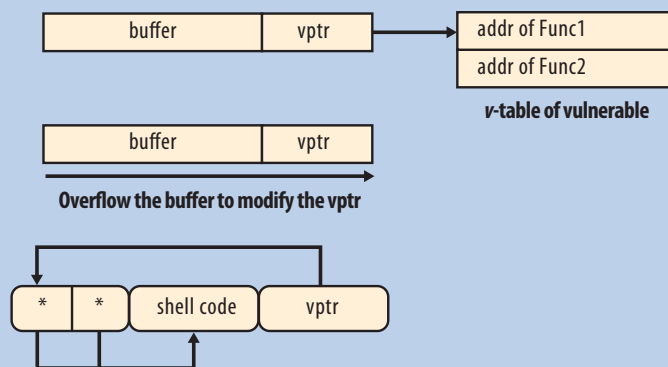


Figure 2. Sample code vulnerable to virtual-function pointer smashing. An attack can overflow the object's member variable `buffer` to modify the `vp`tr, making it point to an attacker-supplied virtual table with injected code. Control will then transfer to this code with the next call to the virtual function.

useful if the program uses a protection mechanism like StackGuard¹ because they alter the return address without changing the canary—a value placed in the stack. When a buffer in the stack overflows, it will corrupt the canary. A program can use the canary as a check against buffer overflow.

Another method for hijacking program control uses `longjmp` buffers. The C standard library provides `setjmp/longjmp` to perform nonlocal jumps. Function `setjmp` saves the calling function's environment into the `jmp_buf` type variable (which is an array type) for later use by `longjmp`, which restores the environment from the most recent invocation of the `setjmp` call. An attacker can overflow the `jmp_buf` with the address of the attacker's code; when the program calls `longjmp`, it will jump to the attacker's code.

Although not widely used, virtual-function pointer smashing is a threat even when a program uses antistack-smashing protection, because such protection does not defend against overflow in the heap. C++ compilers use tables to implement virtual functions. These tables have an array of function pointers that the program uses at runtime to implement dynamic binding. Each instantiated object has a virtual pointer pointing to its virtual table as part of an object's header. By making the virtual pointer point to an attacker-supplied virtual table with injected code, an attacker can transfer control to this code at the virtual function's next call.

Figure 2 shows sample code vulnerable to virtual-function pointer smashing. This attack overflows the object's member variable `buffer` to modify the `vp`tr, making it point to an attacker-supplied virtual table with injected

code. Control will then transfer to this code with the next call to the virtual function.

The Microsoft Windows Structured Exception Handling (SEH) mechanism is also an exploit target. When the program generates an exception, Windows SEH will catch it if the program has no handler or if the provided handler cannot process the exception. The function pointer for the exception handler is on the stack. By overflowing the stack buffer, an attacker can modify it to transfer control to another location.

Heap-based exploits

Dynamic memory allocators such as `malloc` allocate memory on the heap dynamically during runtime. Linked lists manage memory blocks that are allocated and deallocated dynamically using `malloc` and `free`. The management data for each memory block, such as its size and pointers to other memory chunks, is stored in a linked-list-like data structure. The user data and management data are adjacent in a chunk similar to local variables and the return address on a stack. By overflowing user data in the memory block, an attack can corrupt the management data. However, modifying such data does not change program control because this data is not a pointer. Heap-based exploits corrupt the metadata of heap-allocated memory blocks and use it to change other pointers.

DEFENSIVE TECHNIQUES

Researchers have proposed various approaches to address buffer overflow problems, ranging from best practices in development to automated frameworks for recovering from attacks. The five basic methods include

PERSPECTIVES

defensive coding practices, runtime instrumentation, static code analysis, combined static and dynamic code analysis, and network-based instrumentation.

Defensive coding practices

Writing secure code by applying suitable defensive coding practices is the best solution for eliminating vulnerabilities. C and C++ provide no built-in protection for detecting out-of-bound memory accesses. Choosing programming languages like Java or environments like .NET that perform runtime bounds checking eliminates the problem. Standard C library functions including `strcpy`, `strcat`, and `gets` are unsafe because they do not perform bounds checking. Using safer versions of these functions such as `strcpy_s` and `strcat_s` is another good defensive coding practice.

Although defensive coding practices are helpful in developing more secure programs, they are not always feasible in practice.

Performing bounds checking on all arrays solves the problem but decreases performance. Programmers must employ heuristics to identify security-critical buffers and then apply bounds checking to those buffers. Although defensive coding practices are helpful in developing more secure programs, they are not always feasible in practice and are applicable only when actually writing (or planning to write) code. Defensive coding will not work when the source code is unavailable or must remain the same.

Runtime instrumentation

Many runtime techniques for defending against attacks use return address modification to detect buffer overflows. Some proposals include obtaining information about buffer bounds estimates and instrumenting the code for runtime bounds checking.

Compile-time techniques like StackGuard¹ and Return Address Defender (RAD)² insert code to check for return address modification. StackGuard places a canary before the return address and checks the canary's value when the function returns. RAD creates a Return Address Repository global array and copies the return address to it in the function prologue. It then checks for modifications in the function epilogue. These approaches are not completely foolproof because attackers can alter the return address indirectly by using a pointer. StackShield stores the return address in the function prologue and transfers program control to the saved return address on function return.²

This requires source code recompilation and only defends against return-address-based attacks.

Libverify copies functions to heap memory and executes the functions from copied versions.² It uses wrapper functions to store the return address on function entry and verifies it on function return. Libverify uses the return address itself as a canary, reducing the binary instrumentation procedure because the offsets remain the same, unlike StackGuard. Libverify does not need source code but leaves the canary stack itself unprotected.

An alternative approach to Libverify's load-time code instrumentation is to insert instrument code in the executable code itself.³ Because it adds instrumented code before entry and after every function exit address, this approach requires modification of all memory references in the binary. Although it does not require source code, it cannot protect against attacks that target data structures other than the return address.

Split Stack and Secure Return Address (SAS) use two separate stacks, one for data and the other for information control.⁴ However, they do not detect buffer overflow, which corrupt a buffer's neighboring locations. SmashGuard proposes hardware modifications using modified microcoded instructions for CALL and RET opcodes.⁵ These modified instructions store and compare the return address. These two approaches do not require source code recompilation, but they only prevent exploits of frame activation records.

Hardware-supported instruction-level runtime taint analysis addresses noncontrol data attacks.⁶ This approach does not need any changes to the memory system or the processor pipeline dealing with program data. However, it requires classifying instructions as tainted and taintless.

Solar Designer² and Pax⁵ use a nonexecutable stack to combat buffer overflow. However, nonexecutable-stack methods cannot defend against return-to-libc attacks and attacks on data segments; some instances also need an executable stack. OpenWall maps the shared libraries' address space so that their addresses always contain zero bytes to defend against return-to-libc attacks. Pax uses address space layout randomization (ASLR) and a page-based mechanism to protect the heap and stack.

ProPolice places a canary before the return address and also places pointers before the local buffers.² Doing so prevents exploitation of activation records but cannot prevent heap- and data-segment-based attacks involving the function pointer, long-jump buffer variables, and members of structures because ProPolice cannot rearrange pointer variables of structures. PointGuard uses encryption to protect against code and data pointer attacks, but the encryption and decryption might significantly decrease runtime performance.¹

Libsafe intercepts all calls to unsafe standard C library functions.² It substitutes similar functions that limit any overflows within the current stack frame. Because a buffer

cannot extend beyond a stack frame, these overflow-limiting functions prevent overwriting of the return address. This method only protects those C library functions for which it has substitutes; even for these cases, an attacker can overwrite anything up to the frame pointer. Thus, this method cannot protect against attacks targeting function pointers and heap-based overflows. Some proposals extend Libsafe to intercepting calls to malloc, thereby preventing heap-based exploits as well.

Some solutions transform static buffers to dynamically allocated heap-based buffers. Any overflow to these buffers leads to a segmentation fault, which flags the attempted exploit.⁷ When such accesses occur, instead of letting the overflow occur and corrupt the memory, their compiler stores the out-of-bounds write value in a hash table. Whenever program execution references this value, the hash table information provides this stored value based on the read address, which allows the program to continue executing instead of crashing or halting. However, this approach can incur significant performance overhead and might be unsuitable for many applications.

Dytan is a taint-analysis framework in which the user specifies the taint sources, sinks, and taint level.⁶ It accesses the user-supplied library, control flow graph (CFG), and post-dominator-tree information to identify the sources and sinks, and applies the taint level to operands. It then uses the Pin tool (www.pintool.org) to produce an instrumented executable. To detect buffer overflow, users can specify the source as data read from the network and the sink as program control instructions like jump, ret, and so on. The taint markings associated with each byte or memory range, the storing of CFGs, PDOM trees for the program, and the program's related libraries result in high space and time overhead.

When these approaches detect attacks, they halt program operation—in effect, resulting in DoS—yet provide no mechanism for self-healing. Dira, a tool that can repair itself from a detected attack, maintains a log that records memory updates to track data dependencies.⁸ When an attack occurs, it uses this log to restore the program's memory to its preattack state. Dira incurs high runtime overhead because it must log each memory update and track each data dependency.

Exterminator is a runtime system for detecting, isolating, and correcting heap-based memory errors.⁹ Each object has metadata, which Exterminator uses for error isolation and detection before memory allocation. Based on the information from the error isolation algorithm, Exterminator generates a runtime patch for each error. For a buffer overflow, it pads the buffer with the maximum value encountered for this error. This approach does not need source code, and it is useful for testing or automatically correcting a deployed system. However, isolating and detecting the heap-based errors requires additional runs and increased memory consumption.

Using more secure versions of C like Cyclone helps prevent buffer-overflow attacks but would be practical only for yet-to-be-developed projects—porting legacy code to Cyclone would necessitate prohibitively costly code transformation or modification.² CCured translates C programs into CCured and establishes all pointers as either safe, sequenced, or dynamic through a constraint-based type-inference algorithm.² It uses runtime checks when static analysis is not enough to determine safety. CCured requires source code changes, but fewer than Cyclone.

Using more secure versions of C like Cyclone helps prevent buffer-overflow attacks but would be practical only for yet-to-be-developed projects.

Static code analysis

Static analysis of the program source code or disassembled binary code can identify buffer-overflow vulnerabilities. Although these techniques do not incur runtime overhead, they generate many false positives because they lack runtime information.

LCLint is an annotation-assisted static analysis tool¹⁰ that programmers can use to set preconditions and postconditions for state functions. Constraints used to describe buffer ranges include minSet, maxSet, minRead, and maxRead. When the program calls an annotated function, it checks pre- and postconditions to ensure safe access to buffers using these buffer range constraints. LCLint requires programmers to provide annotations and protects such annotated functions. Buffer integer range analysis protects only those library functions with annotations.

Vinod Ganapathy and colleagues model pointers to character buffers by four constraint variables to denote the maximum and minimum number of bytes the buffer allocates and uses.¹¹ They model integer variables using the variable's maximum and minimum values. This technique detects buffer overruns using solver and taint analysis when the maximum used value is greater than the allocated minimum or allocated maximum value for the buffer. However, it generates many false positives because of the flow-insensitive nature of the analysis.

A method that uses maximum length and used-length attributes models statements as constraints and functions as integer transfer functions.¹² Doing so converts the buffer-overflow problem to an error-checking problem by asserting the constraints and finding the reaching paths to this constraint error. However, this tool cannot perform function pointer analysis, nor can it handle arrays of pointers or user-defined structure arrays.

PERSPECTIVES

Marple identifies infeasible paths, examines buffers, and classifies paths that lead to buffer access as safe, vulnerable, overflow-input independent, or unknown.¹³ It does so by raising queries and propagating them backward along the control flow. Marple updates control flow at nodes where it can collect information and whenever it encounters a potential buffer overflow statement. The query terminates when it reaches program entry or an infeasible segment, or when information gathered during propagation resolves the query. Marple takes application source code as input and returns vulnerable path segments to the user, who can then develop patches. However, it uses a conservative analysis that might generate many false positives.

A method that traverses feasible program execution paths and uses the extracted information to perform context-sensitive taint analysis can detect vulnerabilities in x86 executables.¹⁴ The analysis identifies unsecure functions and classifies them as tainted sources or sensitive sinks. Taint analysis checks whether these functions pass tainted data from sources to sensitive sinks, and, if so, raises an alert. Because loops do not execute as many times as in concrete execution, this technique misses some feasible paths, causing false negatives.

Both source and binary code analysis tools and network tools should be part of a programmer's arsenal for protecting against buffer-overflow attacks.

Combined static and dynamic code analysis

Other solutions use both static and dynamic analysis to detect buffer-overflow vulnerabilities.

Researchers have proposed algorithms for selecting susceptible buffers, creating buffer overruns, and, based on the result, analyzing the application for susceptibility.¹⁵ This technique identifies locations that call unsafe library functions on local buffers and nonlibrary functions that read or copy user input. It then calculates the return address that attackers would overwrite to insert an attack string. This approach targets only exploits of the return address and cannot assess an application's susceptibility to other exploits.

Loop-extended symbolic execution (LESE) introduces new variables for representing trip counts for each loop and links them to variables representing program input.¹⁶ It combines these symbolic constraints with conditions for security policy violation and uses the results for vulnerability checking. LESE identifies buffer overflows in real-world programs by sending an initial benign input and uses that execution trace with grammar to discover vulnerabilities. Although this approach is suitable for discovering vulner-

abilities based on security predicate violation and input processing using loops, it might not be applicable for other purposes.

Network-based instrumentation

Network-based instrumentation techniques compare network data with vulnerability signatures from previous attacks, use dynamic taint analysis on network data, or inspect payloads for shell code.

TaintCheck identifies user input data from the network and performs runtime binary rewriting to track the propagation of tainted data.¹⁷ If the program uses tainted data as a jump target or as an argument for a system call, TaintCheck identifies an attack. It then generates an exploit signature by applying backward slicing to the tainted data propagation in the memory. TaintCheck also identifies the parts of the payload used in these attacks by monitoring how the vulnerable program uses each byte of payload at the processor instruction level. It can use this information to generate an attack signature or for hints to use in pattern extraction techniques. However, TaintCheck suffers from slow performance because it runs in Valgrind's emulation environment.

The Pasan prototype instruments source code to record information about the size of static and dynamically allocated buffers and to produce a memory update log.¹⁸ It uses RAD to detect buffer-overflow vulnerabilities. After detecting an attack, Pasan uses runtime information and the memory update log to perform a dependency analysis on the corrupted target address and identify the vulnerable code. Based on the type of code, Pasan either uses a safe library function or instruments the vulnerable code with bounds-checking code to generate a patch. However, the logging and bounds checking incurs a throughput penalty of 10 to 23 percent.

Vigilante traces network data dynamically by tracking the dataflow and generating a security trap when the program uses data unsafely, such as when it loads the data into the program counter or passes it as an argument to security-critical functions.¹⁹ However, Vigilante cannot detect attacks that overwrite security-sensitive information with values indirectly controlled by a worm.

SigFree checks for code in the request packet.²⁰ The idea is that buffer-overflow attacks need executable code to launch an exploit, but client requests to a server do not contain executables. Because SigFree is not based on vulnerability signature comparison, it can detect and block new attacks. On the other hand, it is not effective against DoS and return-into-libc attacks.

DETECTION TOOLS

Both source and binary code analysis tools and network tools should be part of a programmer's arsenal for protecting against buffer-overflow attacks.

Source code analysis tools

ITS4 (www.cigital.com/its4) scans C/C++ source code to identify dangerous standard library functions and uses a handler to perform risk evaluation based on the initial stored information such as checking for race conditions and parameters of unsafe string functions. Instead of parsing source code from a single build, it scans several files to look for vulnerabilities in multiple builds, thereby reducing false negatives. Programmers can use ITS4 in an integrated development environment to highlight errors within that editor. The tool is rudimentary, but it is better than the grep tool. However, it generates many false positives. Because ITS4 relies on a database of vulnerable functions, calling a vulnerable function not present in the database leads to false negatives.

The Rough Auditing Tool for Security (RATS; www.fortify.com/security-resources/rats.jsp) and Flawfinder (www.dwheeler.com/flawfinder) also employ a database to identify and flag security vulnerabilities. Both generate many false positives and false negatives because they perform only a rough analysis.

The Buffer Overrun Detection (www.cs.berkeley.edu/~daw/boon) tool converts the buffer-overflow detection problem to an integer constraint problem by modeling strings (based on their size and usage) and library functions. BOON identifies which buffer has been overrun, but because it is flow insensitive, it does not always reliably identify which statement has the fault or the path that leads to the fault.

Modelchecking Programs for Security (www.cs.berkeley.edu/~daw/mops) finds vulnerabilities by detecting violations of temporal safety properties. MOPS builds models of the program and of the security property, then identifies whether the program model satisfies the security property. It can detect buffer overflows and user privilege issues, but requires building rules to express temporal safety properties. MOPS also requires users to specify the properties to check.

Binary code analysis tools

Binary analysis tools cannot by themselves identify vulnerabilities. Rather, they evaluate performance and gather statistics about programs, thereby greatly aiding in the reverse engineering of binaries for vulnerability and malware detection.


Valgrind (<http://valgrind.org>) is a Linux-based instrumentation framework for building dynamic analysis tools. Programmers can use Valgrind to disassemble code into an intermediate representation, instrument it with analysis code, and convert the instrumented code back into binary code. Valgrind is useful for memory leak detection, memory debugging, program profiling, and thread error detection. It is suitable for both source and binary code analysis, but incurs performance penalties because of the code transformations.

Pin is a dynamic binary instrumentation tool that a programmer can use for binary rewriting to inject arbitrary code at selected locations during runtime. It also includes the source code for instrumentation tools such as basic block profilers, cache simulators, and instruction trace generators. The DynamoRio (www.dynamorio.org) platform can perform program analysis of a running application, profiling, and binary rewriting or instrumentation. Both Pin and DynamoRio share the execution environment with the running application, and neither can handle applications involving multiple processes.

Network tools

Snort (www.snort.org) is a network intrusion protection system (NIPS) or network intrusion detection system (NIDS) that can detect buffer overflows, as well as attacks and probes such as stealth port scans, by performing content searching or matching and protocol analysis of real-time traffic. It is a solely signature-based system and can detect only attacks for which signatures are available. Bro (www.bro-ids.org) is a Unix-based NIDS that is not limited to identifying attacks based on signatures because it works at a higher level of abstraction. It uses vulnerability signatures and events to detect known attacks as well as patterns of failed connection attempts and connection service requests. Snort and Bro rely on manually generated rules and signature databases.

Although solutions and tools exist for flagging potential buffer-overflow vulnerabilities, they are inadequate because of the wide scope of the problem and each approach's inherent limitations. Extending and improving the existing defense methods for buffer-overflow exploits is imperative, especially the ability to handle new types of exploits.

Due to the diverse nature of the attacks, it is extremely difficult—if not impossible—to prefabricate methods for defending against them. Therefore, we suggest exploring methods that can defend against buffer exploits by acquiring knowledge from various sources in a dynamic and extensible way. Such sources might include expert specifications, analysis of code involved in new attacks, and trend analysis. We further suggest the integrated exploration of program analysis, pattern recognition, and data mining to establish such methods. 

Acknowledgment

This work is funded by the Centre for Strategic Infocomm Technologies, MINDEF Singapore.

References

1. C. Cowan et al., "Buffer Overflows: Attacks and Defenses for the Vulnerability of the Decade," *Proc. Foundations*

PERSPECTIVES

- Intrusion Tolerant Systems* [Organically Assured and Survivable Information Systems] (OASIS 03), IEEE CS, 2003, pp. 227-237.
2. J. Wilander and M. Kamkar, "A Comparison of Publicly Available Tools for Dynamic Buffer Overflow Prevention," *Proc. 10th Network and Distributed System Security Symp.* (NDSS 03), Usenix, 2003, pp. 149-162.
 3. D. Nebenzahl, M. Sagiv, and A. Wool, "Install-Time Vaccination of Windows Executables to Defend against Stack Smashing Attacks," *IEEE Trans. Dependable and Secure Computing*, July-Sept. 2006, pp. 78-90.
 4. J. Xu et al., "Architecture Support for Defending against Buffer Overflow Attacks," *Proc. 2nd Workshop on Evaluating and Architecting System Dependability* (EASY 02), 2002; <http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.7372>.
 5. H. Ozdoganoglu et al., "SmashGuard: A Hardware Solution to Prevent Security Attacks on the Function Return Address," *IEEE Trans. Computers*, Oct. 2006, pp. 1271-1285.
 6. J. Clause, W. Li, and A. Orso, "Dytan: A Generic Dynamic Taint Analysis Framework," *Proc. 2007 Int'l Symp. Software Testing and Analysis* (ISSTA 07), ACM, 2007, pp. 196-206.
 7. M. Rinard et al., "A Dynamic Technique for Eliminating Buffer Overflow Vulnerabilities (and Other Memory Errors)," *Proc. 20th Ann. Computer Security Applications Conf.* (ACSAC 04), IEEE CS, 2004, pp. 82-90.
 8. A. Smirnov and T. Chiueh, "DIRA: Automatic Detection, Identification, and Repair of Control-Hijacking Attacks," *Proc. 12th Ann. Network and Distributed System Security Symp.* (NDSS 05), Internet Soc., 2005; www.isoc.org/isoc/conferences/ndss/05/proceedings/papers/dira.pdf.
 9. G. Novark, E.D. Berger, and B.G. Zorn, "Exterminator: Automatically Correcting Memory Errors with High Probability," *Proc. 2007 ACM SIGPLAN Conf. Programming Language Design and Implementation* (PLDI 07), ACM, 2007, pp. 1-11.
 10. D. Larochelle and D. Evans, "Statically Detecting Likely Buffer Overflow Vulnerabilities," *Proc. 10th Usenix Security Symp.*, Usenix, 2001; www.usenix.org/events/sec01/full_papers/larochelle/larochelle.pdf.
 11. V. Ganapathy et al., "Buffer Overrun Detection Using Linear Programming and Static Analysis," *Proc. 10th ACM Conf. Computer and Comm. Security* (CCS 03), ACM, 2003, pp. 345-354.
 12. L. Wang, Q. Zhang, and P. Zhao, "Automated Detection of Code Vulnerabilities Based on Program Analysis and Model Checking," *Proc. 2008 8th IEEE Int'l Working Conf. Source Code Analysis and Manipulations* (SCAM 08), IEEE, 2008, pp. 165-173.
 13. W. Le and M.L. Soffa, "Marple: A Demand-Driven Path-Sensitive Buffer Overflow Detector," *Proc. 16th ACM SIGSOFT Int'l Symp. Foundations of Software Eng.* (SIGSOFT 08/FSE-16), ACM, 2008, pp. 272-282.
 14. M. Cova et al., "Static Detection of Vulnerabilities in x86 Executables," *Proc. 22nd Ann. Computer Security Applications Conf.* (ACSAC 06), IEEE CS, 2006, pp. 269-278.
 15. A.K. Ghosh and T. O'Connor, "Analyzing Programs for Vulnerability to Buffer Overrun Attacks," *Proc. 21st Nat'l Information Systems Security Conf.* (NISS 98), 1998; www.ouah.org/ghosh98analyzing.pdf.
 16. P. Saxena et al., "Loop-Extended Symbolic Execution on Binary Programs," *Proc. 18th Int'l Symp. Software Testing and Analysis* (ISSTA 09), ACM, 2009, pp. 225-236.
 17. J. Newsome and D. Song, "Dynamic Taint Analysis for Automatic Detection, Analysis, and Signal Generation of Exploits on Commodity Software," *Proc. 12th Ann. Network and Distributed System Security Symp.* (NDSS 05), Internet Soc., 2005; www.isoc.org/isoc/conferences/ndss/05/proceedings/papers/taintcheck.pdf.
 18. A. Smirnov, R. Lin, and T. Chiueh, "Automatic Patch Generation for Buffer Overflow Attacks," *Proc. 3rd Int'l Symp. Information Assurance and Security* (IAS 07), IEEE CS, 2007, pp. 165-170.
 19. M. Costa et al., "Vigilante: End-to-End Containment of Internet Worms," *Proc. 20th ACM Symp. Operating Systems Principles* (SOSP 05), ACM, 2005, pp. 133-147.
 20. X. Wang et al., "SigFree: A Signature-Free Buffer Overflow Attack Blocker," *Proc. 15th Usenix Security Symp.*, Usenix, 2006, pp. 225-240.

Bindu Madhavi Padmanabhuni is a PhD student in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. Her research interests include software security, software analysis, and testing. Contact her at padm0010@ntu.edu.sg.

Hee Beng Kuan Tan is an associate professor in the Information Engineering Division at the School of Electrical and Electronic Engineering, Nanyang Technological University. His research interests include software security, software analysis, and testing. Tan received a PhD in computer science from the National University of Singapore. Contact him at ibktan@ntu.edu.sg.



Selected CS articles and columns are available for free at <http://ComputingNow.computer.org>.



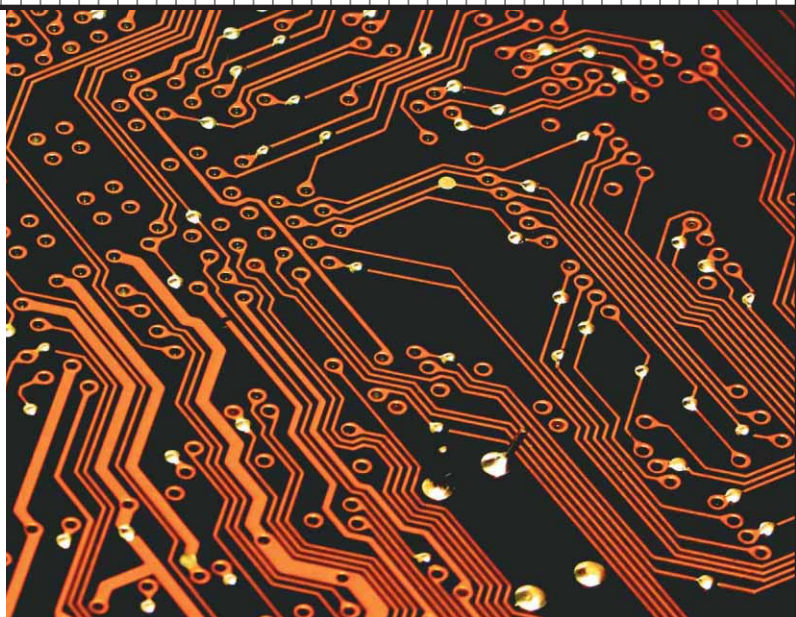
COMPUTING THEN

Learn about computing history
and the people who shaped it.

<http://computingnow.computer.org/ct>

RESEARCH FEATURE

Algorithmic Trading



Giuseppe Nuti, Mahnoosh Mirghaemi, Philip Treleaven, and Chaiyakorn Yingsaeree

UK Centre in Financial Computing, London

Traders increasingly use automated systems for one or more stages of the trading process, yet the secrecy and complexity of the algorithms prompt providing an overview of how these systems work.

Advances in telecommunications and computer technologies during the past decade have created increasingly global, dynamic, and complex financial markets, which in turn have stimulated trading by computer programs and the rise of systems for *algorithmic trading*—also known as AT, algo, or black-box—to automate one or more stages of the trading process.

These systems seek to capture fleeting anomalies in market prices, profit from statistical patterns within or across financial markets, optimally execute orders, disguise a trader's intentions, or detect and exploit rivals' strategies.¹ Ultimately, profits drive any algorithmic trading system—whether in the form of cost savings, client commissions, or proprietary trading.

As the “Electronic Trading” sidebar describes, institutional traders and managers of pension funds, mutual funds, and hedge funds increasingly deploy algorithmic trading systems. These systems currently handle approximately 50 to 60 percent of all stocks traded in the US and EU.² High-frequency algorithmic trading accounted for 60 percent of US equity volumes in 2009, and it is a major driver for computing and analytics innovation,³ especially machine learning and grid/GPU computing.

However, algorithmic trading is also of major concern to regulators, as the 6 May 2010 Flash Crash clearly illustrated.³ In this instance, the Dow Jones Industrial Average plunged about 600 points in 5 minutes, causing a loss of \$600 billion in the market value of US corporate stocks. This event revealed the lack of knowledge about high-frequency algorithmic trading and exposed its potential vulnerability. Protecting against such events requires an in-depth understanding of the trading process.

MARKET MICROSTRUCTURE

To understand algorithmic trading, it is useful to consider the different types of trading, explore how a trade is executed in an exchange, and review the objectives and challenges.

Trade execution

Dealers generally execute their orders through a shared centralized order book that lists the buy and sell orders for a specific security ranked by price and order arrival time (generally on a first-in, first-out basis). This centralized order-driven trading system continuously tries to match buy and sell orders.

RESEARCH FEATURE

ELECTRONIC TRADING

Within algorithmic trading, several closely related terms are used that are often confused. These include electronic trading, order-management systems, automated trading, systematic trading, and algorithmic trading.

Broadly, electronic trading is any method of exchanging securities, stocks, bonds, foreign exchange (currency, Forex—or FX) and derivatives (options, futures, and so on). Within electronic trading, specialized programs bring together buyers and sellers through electronic media to create an exchange (such as Nasdaq). Order-management systems facilitate and manage order execution, generally connecting to one or more electronic exchanges. Automated trading systems usually refer to trade execution programs that automatically submit trades to an exchange.

The distinguishing feature of algorithmic (referred to by some people as systematic) trading systems is the sophistication of their analysis and decision making. Broadly, these systems are deployed for highly liquid markets and high-frequency trading, such as equities, futures, derivatives, bonds (US Treasuries), and foreign exchange (currencies). The essential characteristic of a highly liquid market is that there are ready and willing buyers and sellers at all times.

Central to these systems' operation are financial protocols, such as the Financial Information Exchange (FIX) protocol, a series of messaging specifications for the electronic communication of trade-related messages. FIX messages are formed from several fields, and each field is a tag-value pairing, separated from the next field by a delimiter (similar to XML). The TAG is a string representation of an integer that indicates the field's meaning. The value is an array of bytes holding a specific value for that TAG. The FIX protocol also defines sets of fields comprising a particular message; within the sets of fields, some sets will be mandatory and others optional. There are various extensions to FIX, including FIXatdl, the FIX algorithmic trading definition language.

As Figure 1 shows, the order book is divided into two parts: *buy* orders on the left ranked with the highest price at the top, and *sell* orders on the right ranked with the lowest price at the top. Orders are generally listed by price and time priority, which means most exchanges prioritize orders based on the best price, and, if two or more orders are inserted at the same price, priority is given to the first order to be inserted. Notable exceptions include the UK three-month interest rates contracts (known as Short Sterling) on the London International Financial Futures and

Options Exchange (LIFFE). These contracts prioritize both the arrival time and order size, such that a large order can have priority over a smaller trade, even if the larger order was inserted after the smaller one.

Buy orders that a sell order can fully or partially match are automatically traded. Many variants to the order book model exist. Different exchanges will accept different order types, including limit orders, as per our example, market orders, stop-loss orders, and so on. In developing an algorithmic trading system, knowledge of the market microstructure—the detailed process governing how trades occur and orders interact in a specific market—is of paramount importance.

Recently, both media and financial regulators have focused their attention on algorithmic trading in a bid to explain some unusual phenomena in the market. The abnormal drop in US equity markets in the Flash Crash, and subsequent reversal of the move, is a good example of how market microstructure can have a drastic effect—and possibly an undesirable one from a regulator's viewpoint—on price dynamics.⁴ In this case, interconnected exchanges had different market microstructures for the securities in question, and the main exchange (by volume) had specific rules designed to interrupt trading during periods of high volatility and volume. Smaller satellite exchanges did not have the same safety features. Automated order-routing programs directed trades to the smaller exchanges which, because of a lack of liquidity, caused several share prices to plummet.

However, there is still uncertainty about the underlying causes of the Flash Crash, especially as one of the theories explored by the US Securities and Exchange Commission's investigation attributed the cause of the crash to an abnormally large order erroneously inserted by a human trader. Although the jury is still out on the crash's ultimate cause, algorithmic trading systems had an amplifying effect on the severity and breadth of that day's equity markets' decline.

Trading objectives

Banks are usually thought of as intermediaries, acting as brokers or market makers or servicing clients by buying

Order Book – ABC Inc.			
Buy		Sell	
Quantity	Price	Price	Quantity
5,000	99	99	4,000
8,000	98	100	10,000
10,000	97	101	1,000
15,000	95	103	15,000

(a)

Order Book – ABC Inc.			
Buy		Sell	
Quantity	Price	Price	Quantity
1,000	99	100	10,000
8,000	98	101	1,000
10,000	97	103	15,000
15,000	95	104	3,000

(b)

Figure 1. An example trade order book (a) before matching a trade and (b) after matching a trade. Buy orders, which are generally listed by price and time priority, are ranked with the highest price at the top, while sell orders are ranked with the lowest price at the top.

or selling stocks or bonds. However, most banks and funds also engage in proprietary trading on their own accounts. Proprietary trading occurs when an institution actively trades stocks, bonds, options, commodities, derivatives, or other financial instruments with its own money as opposed to its customers' money.

The type of trading—broker or proprietary—will shape the algorithmic trading strategy's design; typically, broker algorithmic trading systems seek to minimize the cost of trading by optimizing the execution strategy—that is, minimize market impact cost or time to execution, optimize the price, and so on—whereas proprietary algorithmic trading systems seek to maximize profits against some measure of financial risk. In practice, all algorithms target profits, either in the form of cost savings or trading profit and loss; the difference is in who participates in the profits—the clients versus the trading firm—and who takes on the trading risk. Generally, the institution that is taking the risk also takes the lion's share of the profits, whereas an institution acting only as intermediary collects a brokerage fee.

Current proprietary trading strategies include index arbitrage, statistical arbitrage, merger arbitrage, fundamental analysis, volatility arbitrage, and macrotrading.

Trading process

An intuitive way to classify algorithmic trading is through the separate processes being automated within a trade's life cycle. As Figure 2 illustrates, algorithmic trading can be used at any stage of the trading process and for

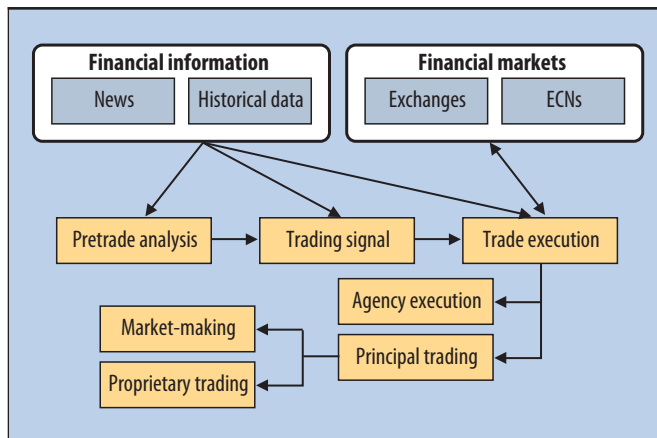


Figure 2. Algorithmic trading systems. The figure shows the three stages of algorithmic trading—pretrade analysis, trading signal, and trade execution—and the two major firm types: agency execution and principal trading.

various purposes, including market making, spread trading (also known as relative value or basis trading), arbitrage, and macrotrading. Algorithmic trading therefore covers a wide variety of systems. In trade-execution programs, for example, the algorithm might decide aspects such as timing, price, and the order's quantity splits. Other systems might automate the complete trading process.

As the “Algorithmic Trading System Components” sidebar describes, the trading process can be split into four stylized steps: pretrade analysis, trading signal generation, trade execution, and post-trade analysis.

ALGORITHMIC TRADING SYSTEM COMPONENTS

Figure A shows the major components of an algorithmic trading system and the steps at which they occur. Pretrade analysis includes three mathematical models:

- The *alpha* model predicts the future behavior of the financial instruments to trade.
- The *risk* model evaluates the levels of exposure/risk associated with the financial instruments.
- The *transaction cost* model calculates the (potential) costs associated with trading the financial instruments.

Trading signal generation consists of the portfolio construction model. This model takes as its inputs the results of the alpha, risk, and transaction cost models and decides what portfolio of financial instruments should be owned going forward and in what quantities.

At trade execution, the execution model executes the trades, making several decisions with constraints on (actual) transaction costs and trading duration. The most general decision is the trading strategy followed by the venue and order type.

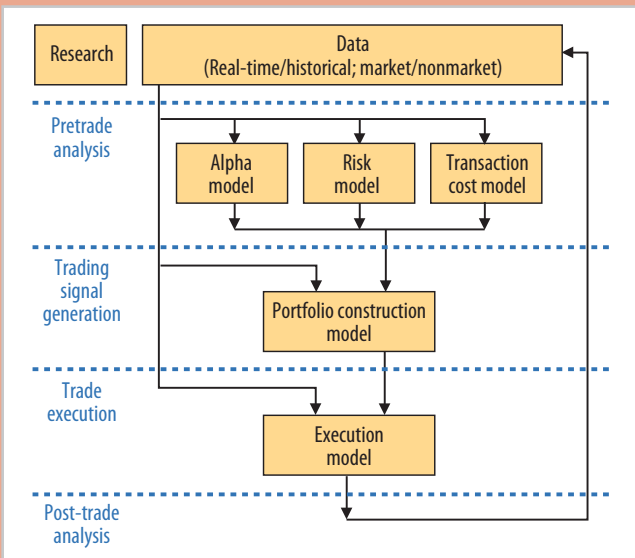


Figure A. Steps at which the algorithmic trading system components occur.

RESEARCH FEATURE

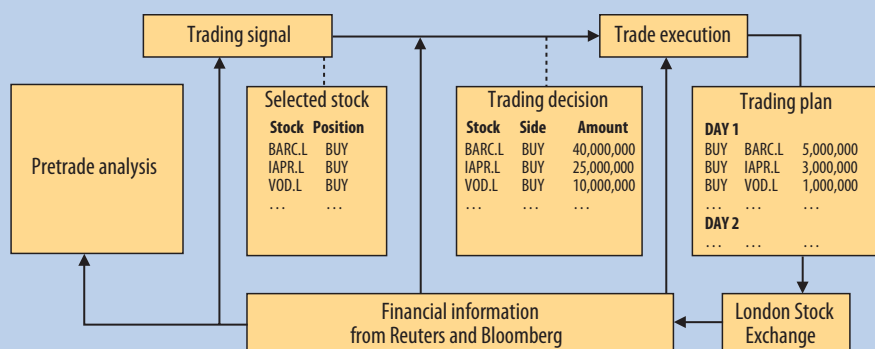


Figure 3. Example algorithmic trading system. The pretrade analysis selects a group of stocks to buy and passes this information to the trading signal component, which determines the quantity of each stock to buy. Finally, the trade execution determines the trading plan by selecting the exchanges and associated quantities.

Pretrade analysis is the most common use of algorithms within a trading environment. It encompasses any system that uses financial data or news to analyze certain properties of an asset. It can be as simple as a method to value a company, or it can involve state-of-the-art algorithms that use artificial intelligence techniques to scan news or Twitter feeds to forecast asset price volatility. Pretrade analysis, as a stand-alone algorithmic trading system, stops short of generating a trade signal. Human traders use the output to make trading decisions that are most likely based on a selection of trading signals and some discretionary input.

The next step in automating the trading process is *trading signal generation*. Systematic asset managers and trading institutions often use this level of automation. Human traders can execute the generated signal if they require further discretionary input or—more often than not—the trade is not executable electronically because of the order size versus the market liquidity. This level of automation is generally applicable to all but high-frequency trading, where complete automation is a prerequisite.

The third step is *trade order execution*. Algorithmic trading can execute trades and place orders in one or more exchanges. A human trader can make the actual trading decision, in which case the algorithm only optimizes the execution (this is often associated with agency trading). If the trading decision is generated algorithmically, more often than not the trade is proprietary.

ALGORITHMIC TRADING SYSTEM EXAMPLE

The structure and operation of an actual algorithmic trading system depends on which stages of the trading process are being automated, whether the system is supporting broker or proprietary trading, and what type of securities are being traded—for example, equities, bonds, or currencies.

To illustrate an algorithmic trading system's operation, we use a simple example of a fully automated system for

the equity market—specifically, from the FTSE (*Financial Times* and London Stock Exchange) 100 index. In this example, the objective is to replicate the FTSE 100 index by trading a subset of stocks (that is, not the entire index) that provides a similar performance to the index while minimizing the transaction costs naturally incurred when rebalancing any replicating portfolio.

In pretrade analysis, the system compares and contrasts the historical performances of several index-tracking strategies to help the user select the strategy that best suits current market conditions. The other components work toward implementing the chosen replicating strategy. Specifically, the trading signal component selects the stocks for constructing the benchmark portfolio and determines the weight of each stock in the portfolio.

After all of the portfolio composition decisions have been made, the trade-execution component determines how best to execute the trade to minimize market impact and timing risk. In the example in Figure 3, strategies in trading signal generation and trade execution are formulated as an optimization problem with multiobjective constraints. Algorithms for solving these problems might use quadratic programming, genetic programming, or particle swarm optimization.

PRETRADE ANALYSIS

Pretrade analysis in an algorithmic trading system generally involves both analyzing financial data or news with the aim of forecasting future price movement or volatility and generating trading signals when a trading opportunity occurs. Broadly, the three categories of techniques used to perform this include fundamental analysis, technical analysis, and quantitative analysis.

Fundamental analysis

Fundamental analysis involves a detailed study of related information that might affect asset prices with the aim

of determining the asset's fair value (or its potential future price movements). Relevant information might include the overall state of two countries' economies (such as unemployment figures), interest rates, gross domestic product, or national policies. The idea that the current market price for an asset is not equivalent to its fair value contradicts the somewhat controversial efficient-market hypothesis, which suggests that the current price is a reflection of all the available information.⁵

Fundamental analysis typically generates trading signals when the current asset price differs from the fair value obtained using discount models for the analysis; ratios, such as the price-to-earnings ratio; or certain fundamental properties, such as earnings yielding more than twice the AAA-rated bond yield. In the past few decades, analysts and traders have used advanced mathematical and statistical models from machine learning and computational statistics to determine the relationship between the stock price's future value and its fundamental quantities with the aim of identifying stocks with a potential to appreciate (or depreciate) significantly.

Technical analysis

Technical analysis aims to predict future price movements based on asset price history and, sometimes, related trading information such as trade volume. By assuming that the market's price reflects all relevant information, technical analysis seeks to identify and exploit price movement patterns rather than examine the underlying factors affecting asset prices.

Many popular technical analysis techniques are based on the premise that asset prices move in trends, hence technical trading systems typically generate entry signals when a new trend is identified and generate exit signals when a trend ends. Traditionally, they determine trends by analyzing the continuation patterns (such as the ascending symmetrical triangle pattern) and reversal patterns in a chart, as well as trend lines, support, and resistance areas. Technical traders also use indicators such as moving averages, advance-decline lines and ratios, the relative strength index, and the stochastic oscillator, which indicates market momentum. Recently, more complex modeling techniques have included trading rules formulated using genetic programming, or statistical time series forecasting methods such as autoregressive fractionally integrated moving-average models and neural networks.

Quantitative analysis

Quantitative analysis treats asset prices as random and uses mathematical and statistical analysis to find a suitable model for describing this randomness. This type of analysis has dominated the financial industry in recent decades, forming a solid foundation for portfolio theory, derivatives pricing, and risk management.

Although fundamental analysis and technical analysis also use mathematical and statistical methods, they are primarily concerned with the deterministic relationship between the asset price and related information. In contrast, quantitative analysis focuses on an asset price's stochastic behavior. Consequently, quantitative analysis is generally related to the pricing of derivative products, such as options and swaps, whose fair value relies on the underlying asset's stochastic property as well as the analysis of the temporal convergence and divergence of price movements of pairs and baskets of assets.

When used within an algorithmic trading system, quantitative analysis typically generates trading signals when the current asset price differs from the asset's fair value, such as in statistical arbitrage, which attempts to profit from pricing inefficiencies. The most commonly used and simplest case of statistical arbitrage is pairs trading, which tries to identify divergence of the correlated prices of two stocks.

Assessing the complexity of the trading rules is useful because it highlights the predictability of future returns using only historical prices.

Pretrade analysis example

This example uses genetic programming to identify profit-making opportunities.⁶ Traders use genetic programming to find technical trading rules for a composite stock index. The algorithm aims to find decision rules that divide days into two disjoint categories: in the market (earning the market rate of return) or out of the market (earning the risk-free rate of return).

Each genetic structure represents a particular technical trading rule. A trading rule returns either a buy or a sell signal for any given price history. Building blocks for trading rules include a simple function of past price data, numerical and logical constants, and logical functions that allow the combination of low-level building blocks to more complicated expressions.

The root node of each genetic structure corresponds to a Boolean function to ensure that the trading strategy is well-defined. Assessing the complexity of the trading rules is useful because it highlights the predictability of future returns using only historical prices. If the algorithm reveals complicated rules, the results would be consistent with a view that there is some kind of hidden structure that could be discovered from past prices. If it finds only relatively simple rules, the results would be more consistent with a view that past prices have limited value in predicting future returns.

RESEARCH FEATURE

TRADING SIGNAL GENERATION

The difference between pretrade analysis and trading signal generation is often blurred because there is generally substantial overlap between the two. The major difference between them is that an actual trading signal generated by an algorithm will come with a specific price (and possibly a quantity) and might even include risk management recommendations, such as specific stop-loss values.

The distinguishing features are that pretrade analysis only offers recommendations to buy or sell, which are purposely left vague (at what price, with what trading horizon, and so on), whereas trading signal generation is augmented with specific values that can be translated into actual trades. In other words, a pretrade analysis recommendation ought to be seen as being only part of a possibly complex decision process that might ultimately be translated into a trading strategy. Conversely, a trading signal can be directly translated into a trading strategy and—most importantly—can therefore be

The difference between pretrade analysis and trading signal generation is often blurred because there is generally substantial overlap between the two.

replayed against historical market data to simulate its performance. The ability to simulate the performance of a trading signal is useful in assessing its value, yet it adds a great deal of complexity in that the trading signal must be able to correctly analyze diverse trading environments, such as the volatile few minutes following an important economic data release or quiet and illiquid market times during Asian trading hours.

Pretrade analysis overlaps trading signal generation in providing a recommendation on when (if not at what price) to enter into a trade; a trading signal must also provide a strategy to close out the trade. Exit strategies, risk management, and cash management are popular strategies to complete the round-trip of a trade.

Entry strategies

To generate a signal, an algorithmic trading system must generate an entry strategy. This can be as simple as a fixed expected profit, such as a predetermined difference between the current price and the fair value, if using a fundamental analysis system. Conversely, a technical analyst might define the start of a trend as a simple entry point for a trade.

Unfortunately, simple rules often have two main deficiencies:

- The trading signal might oscillate (from buy to sell) for prolonged periods, thus incurring large transactional costs.
- The rule might be unable to detect a regime shift that would invalidate the model's assumptions and thus the trading signal.

Clearly, to improve an entry strategy we might need to add further complexity to the trading rules. For example, we might want to impose a minimum number of consecutive identical trade recommendations to improve a trading signal that oscillates between opposing trades. The set of possible additional rules is large and varied; furthermore, the effect of combining more than one rule can be complex. Although undeniably a requirement, forming complex and profitable entry strategies is still somewhat of an art, at least in the selection of homogeneous rules and rule parameters to search through.

Exit strategies

Developing an exit strategy requires specifying when to take profits and when to exit a trade at a loss. For example, we might want to stay in the trade as long as our assumption about the market remains valid to maximize profits, and exit the trade as soon as our assumption is proven wrong.

In practice, this determination largely depends on the system's trading strategy. For example, a fundamental trading strategy that opens a buy position when an asset's value is less than its fundamental value might keep the position open until the asset price moves closer to its fair value, while exiting the position (possibly with a loss) when changes in the underlying factors reduce the asset's fundamental value.

A technical trading strategy might keep the position open until it reaches a desired target, but will exit with a loss if the pattern or the trend used to enter the trade is proved wrong. For example, a strategy that trades a resistance level (a specific price level that we expect not to be surpassed in the near future) might close the position if the price moves decidedly above this level. These exit strategies are generally implemented by a stop-loss order—that is, an order to buy (or sell) an asset once its price has climbed above (or fallen below) a specified price.

Risk management

For each trading opportunity, a trading system must calculate how to manage its market exposure. This is perhaps more important than making entry and exit decisions as poor risk management can easily turn a profitable trading system into a loss-making one. This is, in its basic form, a method to decide how much (that is, what quantity) to trade for each signal.

Several approaches address this problem. The simplest and most traditional method is a fixed-amount system, which places an arbitrary fixed amount in every trade. The problem with this method is that it does not distinguish between periods of high and low volatility. Intuitively, we might want to trade larger quantities during low volatile periods and smaller ones during volatile periods. In other words, we might look to target a constant risk-adjusted return, though we must ultimately accept that our view of future volatility is but an estimation of what will be the realized market volatility. A further option is to use a fixed fractional system that simply risks f percent of the capital on every trade. The optimal fixed fractional system (sometimes referred to as optimal- f strategy) can maximize the geometric growth rate for a series of trades.

Trading signal example

To illustrate trading signal generation, we use a trading signal technique based on alpha generation⁷ for an exchange traded fund (ETF), a fund that can be traded on exchange like regular stock. For a trading signal technique based on alpha generation, alpha is defined as excess risk-adjusted returns measured above a benchmark. In this case, the alpha generator—the trading signal—will determine if a security can generate excess returns or returns higher than a preselected benchmark with a controlled risk, when added to an existing portfolio of assets. Analysts use these signals to develop mathematical and statistical models that help determine whether a specific investment might be profitable.

A mean-reverting ETF strategy example assumes that the returns of each asset within the ETF will, in the long run, converge to the overall ETF's return. Therefore, an asset that underperforms its peers will be expected to catch up, and vice versa. Given a specific level of under- or overperformance, and satisfying a predetermined entry strategy rule set, we systematically buy underperforming assets and sell outperforming ones in the hope that our mean-reverting assumptions are correct.

TRADE EXECUTION

After generating the trading signal, an algorithmic trading system must make several decisions regarding constraints on transaction costs and trading duration. To execute a trade, an order must be submitted to a trading venue, with the choice depending on several factors including order size, trading mechanism, and degree of trader's anonymity. If the trade is too large to execute in a single order in an open market, the system must either break it down into several smaller orders, which it submits to the market over a period of time to minimize market impact, or execute it in alternative markets such as crossing networks or dark pools that do not publicly reveal the current order book.

The system must also determine whether to execute the trade immediately by submitting market orders, or trade

patiently to get a better price by submitting limit orders. As a result, selecting an appropriate execution strategy for a particular trade is not a trivial task; rather, it involves several decisions, each a challenge in itself.

Trading venues

Many financial instruments can be traded in more than one financial market. Thus, trading systems must determine which market to submit the order to. Some of the most important characteristics the trader considers are liquidity, trading mechanism, degree of trader anonymity, and differential execution costs.

Trading systems generally choose to submit orders to the market with the most liquidity because a highly liquid market is usually associated with fast trade execution and low transaction costs. For example, when immediacy is required, a system will typically trade in a continuous double-auction market; otherwise, it will likely trade in a periodic auction, which has lower price volatility.

Generating an optimal trade schedule involves achieving a desired balance between price impact and opportunity cost.

Trading schedules

A system might break a large order into several smaller orders to minimize the trade's impact on the market, because a small order is more likely to flow under the market's radar than a large order. However, delayed execution of smaller orders can expose a trader to potential adverse price movements as well as to an opportunity cost. Generating an optimal trade schedule involves achieving a desired balance between price impact and opportunity cost.

In the past decade, interest in optimal trade schedule models has increased. Typically, there are two main steps in specifying the trading objective:

1. Determine execution cost by defining the specification of transactional cost and choosing the desired benchmark price (for example, previous close, opening price, arrival price, volume-weighted average price [VWAP], time-weighted average price [TWAP], and future close).
2. Specify the degree of risk aversion (that is, how much to penalize variance relative to expected cost), which indicates the level of trading aggressiveness or passiveness.

Aggressive trading is associated with higher cost and less risk, whereas passive trading is associated with lower market impact and higher risk. An arithmetic random walk

RESEARCH FEATURE

is often the most popular model for specifying the dynamics of future market prices. Given specifications of all the factors, an optimal trading strategy for a specific trading objective might be obtained by solving the corresponding stochastic dynamic optimization problem.

Order type

The two main types of trades are market orders and limit orders. When market conditions permit (that is, there is enough liquidity), a market order provides immediate execution, but the execution price is uncertain. A limit order guarantees the execution price, but it can sometimes be executed only partially or not at all.

Traditionally, the decision whether to submit a market order or a limit order to execute the trade is examined in the context of the tradeoff between the payoff associated with limit orders and the risk of nonexecution. Placing an order far from the best bid or ask price will increase the payoff, but the larger the distance from the best price, the larger the chance that the order will not be executed. The key is finding the right tradeoff. Undoubtedly, one of the most important factors in valuing such tradeoffs is having a model of limit order execution times and the associated execution probability. This is because the expected profit of traders who decide to trade via limit orders is an increasing function of the execution probability.

Trade execution example

A trade execution system's main objective is to reduce the hidden costs of trading by balancing the tradeoff between market impact and timing risk over the trading horizon.⁸ The system provides the necessary transparency and flexibility to develop a customized algorithmic strategy to ensure that the algorithm's parameters are consistent with the overall investment goal.

To use the system, investors must specify the benchmark price—that is, the price they are trying to achieve (for example, previous close price, day's closing price, VWAP, TWAP, or day closing price), trading style (for example, aggressive, normal, or patient), and preferred adaptation tactic. This tactic describes how they want the algorithm to adapt to changing market conditions, such as becoming more aggressive in times of favorable prices and more passive in times of adverse price movement. Combining these requirements with a model of market impact and the dynamics of future market prices, and solving the corresponding optimization problem, generates the optimal trading strategy. Traders approximate the market impact model by fitting some parametric function with the historical data, while assuming that the dynamics of future market prices will follow an arithmetic random walk model (that is, the market randomly moves up or down with a given probability that is proportional to the asset's expected volatility).

FUTURE INFLUENCES

Three areas are influencing the future of algorithmic trading systems.

Dark pools are a type of platform that allows the trading of large blocks of shares without revealing quantities or prices publicly (to other traders) until after trades are completed. Dark pools are similar to standard markets (with similar order types, pricing rules, and prioritization rules), but they do not publicly display the order book. These dark liquidity pools offer institutional investors many of the efficiencies associated with trading on the exchanges' public limit order books but without showing their hands to others. However, they are criticized for their lack of transparency and because they could lead to less efficient pricing than traditional open exchanges.

Ultrahigh-frequency trading refers to the buying and selling of stocks at extremely fast speeds with the help of powerful computers. Using algorithms, these computers can scan dozens of public and private markets simultaneously, execute thousands of orders a second, and alter strategies in a matter of milliseconds. In the US, ultrahigh-frequency trading firms represent 2 percent of the approximately 20,000 firms operating today, but they account for 73 percent of all equity trading volume. In ultrahigh-frequency trading, a trader or algorithmic trading system measures its holding period in seconds, sometimes even in hundreds of milliseconds.

Exchange traded funds combine the valuation feature of a mutual fund or unit investment trust (which can be bought or sold at the end of each trading day for its net asset value) with the tradability feature of a closed-end fund (which trades throughout the trading day at prices that might be more or less than its net asset value). An ETF holds assets such as stocks or bonds and trades at approximately the same price as the net asset value of its underlying assets over the course of the trading day. ETFs provide fertile grounds for algorithmic trading systems because they are a new multiasset instrument that makes their trading more complex, hence they offer a greater technical challenge.

Algorithmic trading might be described as an arms race drawing on the skills of top computing professionals. Traditionally, investment banks and funds hired economists for trading positions and computing professionals for technology posts. Now, given the increasing importance of algorithmic trading in financial markets, firms are seeking trader-programmers skilled in C-based languages and analytics (such as computational statistics and machine learning). For computing professionals interested in finance, it is a stimulating and certainly well-paid career. **□**

References

1. A. Chaboud et al., "Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market," *Int'l Finance Discussion Papers*, Board of Governors of the Federal Reserve System, Oct. 2009; www.federalreserve.gov/pubs/ifdp/2009/980/ifdp980.pdf.
2. T. Hendershott, C.M. Jones, and A.J. Menkveld, "Does Algorithmic Trading Improve Liquidity?" *J. Finance*, Feb. 2011, pp. 1-33.
3. R.K. Narang, *Inside the Black Box: The Simple Truth About Quantitative Trading*, Wiley Finance, 2009.
4. Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues, "Findings Regarding the Market Events of May 6, 2010"; www.sec.gov/news/studies/2010/marketevents-report.pdf.
5. M.H. Pesaran, "Predictability of Asset Returns and the Efficient Market Hypothesis," A. Ullah and D.E. Giles, eds., *Handbook of Empirical Economics and Finance*, Taylor & Francis, 2010, pp. 281-312.
6. M.A. Kaboudan, "Genetic Programming Prediction of Stock Prices," *J. Computational Economics*, Dec. 2000, pp. 207-236.
7. J.M. Hill, "Alpha as a Net Zero-Sum Game," *J. Portfolio Management*, vol. 32, no. 4, 2006, pp. 24-32.
8. R. Kissell and R. Malamut, "Algorithm Decision Making Framework," *J. Trading*, vol. 1, no. 1, 2006, p. 10.

Giuseppe Nuti is an honorary senior research fellow in the Department of Computer Science at University College London and has worked at both Deutsche Bank and Citadel Securities developing trading algorithms within the fixed-income asset class. His research interests include nonparametric asset distribution estimation, optimal execution strategies, and multivariable Bayesian statistics. Nuti received a PhD in computer science from University College London. Contact him at giuseppenuti@gmail.com.

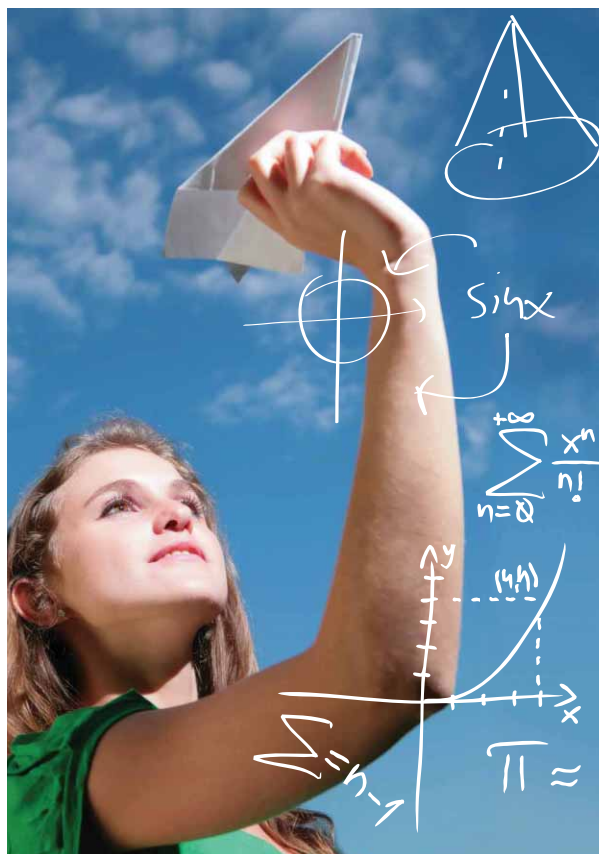
Mahnoosh Mirghaemi is a PhD student in the Department of Computer Science at University College London. His research interests include the econometrics of high-frequency fixed-income data and its application in algorithmic trading. Mirghaemi received an MSc in financial mathematics from Kings College London. Contact him at m.mirghaemi@cs.ucl.ac.uk.

Philip Treleven is professor of computing at University College London and director of the UK Centre in Financial Computing, a partnership of University College London, the London School of Economics, and the London Business School. His research interests include computational finance and artificial intelligence. Treleven received a PhD in computer science from the University of Manchester, UK. He is a member of IEEE and the IEEE Computer Society. Contact him at p.treleven@ucl.ac.uk.

Chaiyakorn Yingsaeree is a PhD student in the Department of Computer Science at University College London, where his research is supported by a scholarship from the National Electronics and Computer Technology Center, Thailand. His research interests include the modeling of limit-order execution and its application in algorithmic trading. Yingsaeree received an MSc in engineering from Kasetsart University, Thailand. Contact him at c.yingsaeree@cs.ucl.ac.uk.



Selected CS articles and columns are available for free at <http://ComputingNow.computer.org>.



Imagine a teenager excited about technology

Every innovative, life-changing idea comes from someone's imagination.

The IEEE Foundation provides resources to advance education, innovation and preservation.

Make a gift—and imagine the difference you can make.

Donate today at
www.ieeefoundation.org



IEEE FOUNDATION

COMPUTER SOCIETY CONNECTION

Chuck Seitz Wins Cray Award



Charles L. (Chuck) Seitz is an architect and designer of innovative computing and communication systems, many of which exploit asynchrony and concurrency to achieve performance at the limits of available technology. Known for creating new disciplines of digital design, Seitz will receive the IEEE Computer Society's Seymour Cray Award in November at the annual SC conference in Seattle. His award citation reads "For innovations in high-performance message passing architectures and networks."

SEYMOUR CRAY AWARD

Established in 1997, the Seymour Cray Award recognizes innovative contributions to high-performance computing systems that best exemplify the creative spirit demonstrated by supercomputer pioneer Seymour Cray. Winners receive a crystal memento, an illuminated certificate, and a \$10,000 honorarium.

CONTRIBUTIONS TO COMPUTING

Seitz became fascinated with digital design during the 1960s at MIT, where he earned a BS, MS, and PhD in electrical engineering. While a graduate student, he taught courses in switching and automata theory,



Chuck Seitz has made strides in message-passing architectures and networks.

developed MIT's digital-system project-laboratory course, and received the MIT Goodwin Medal "for conspicuously effective teaching." Seitz's PhD thesis on asynchronous logic helped to expose and explain the fundamental problems of mutual exclusion and of synchronizing asynchronous signals to a free-running clock.

Seitz later became an assistant professor of computer science at the University of Utah and worked at the Evans & Sutherland Computer Corporation, designing high-performance graphics engines. He then moved to California to perform research for Burroughs on aperture filtering digital-video techniques for character and geometric display.

In 1977, Seitz joined the computer science faculty at Caltech, where his research and teaching focused on VLSI design and concurrent computing. Under DARPA sponsorship, Seitz


and his students developed the first multicomputer, the Cosmic Cube; devised the key programming and packet-switching techniques for the second-generation multicomputers; and transferred these technologies to industry. The Intel Paragon, ASCI Red, and Cray T3D/E employ message-passing techniques licensed from his Caltech patents.

Seitz's 1992 election to the National Academy of Engineering carried the citation "for pioneering contributions to the design of asynchronous and concurrent computing systems."

COMPUTER SOCIETY AWARDS

The IEEE Computer Society recognizes outstanding work by computer professionals who advance the field through exceptional technical achievement and service to the profession and to society.

In the technical area, awards recognize pioneering and significant contributions to the field of computer science and engineering. Service awards honor both volunteers and staff for well-defined and highly valued contributions to the Society. In most cases there are no eligibility restrictions on the nominee or nominator.

Nomination forms are available via the Society's website at www.computer.org/awards. 

Cleve Moler Wins Fernbach Award

Cleve Moler, founder, chairman and chief scientist of MathWorks, was recently honored with the IEEE Computer Society Sidney Fernbach Award for high-performance computing.

Moler was a professor of mathematics and computer science for almost 20 years at the University of Michigan, Stanford University, and the University of New Mexico. At New Mexico, he was a professor in the mathematics department in the late 1970s and then chair of the computer science department in the early 1980s. During this time, he developed several packages of mathematical software for computational science and engineering that eventually formed the basis for MATLAB, a high-level technical computing environment.

MATHWORKS

In 1984, Moler and Jack Little founded MathWorks to commercial-




Cleve Moler coauthored the LINPACK and EISPACK subroutine libraries.

ize and continue the development of MATLAB.

Before joining MathWorks full time in 1989, Moler spent five years with two computer hardware manufacturers, Intel Hypercube and Ardent Computer. At MathWorks, Moler has served as chief scientist, overseeing the mathematical aspects of the company's products. Moler is the one of the authors of the LINPACK

and EISPACK scientific subroutine libraries, as well as author or coauthor of five textbooks on numerical analysis and computational science. He is a member of the National Academy of Engineering and a past president of the Society for Industrial and Applied Mathematics. Today, Moler works from his home office in Santa Fe, New Mexico, writing books, articles, and MATLAB programs.

FERNBACH AWARD

The IEEE Computer Society Sidney Fernbach Award was established in 1992 in memory of one of the pioneers in the development and application of high-performance computers to the solution of large computational problems. Winners receive a certificate and \$2,000 honorarium in recognition of outstanding contributions in the application of high-performance computers using innovative approaches. 

FIFTIETH ANNIVERSARY OF MIT'S COMPATIBLE TIME-SHARING SYSTEM

David Walden, *Chair, IEEE Computer Society History Committee*

Time sharing was in the air around MIT and Cambridge in the years circa 1961. MIT faculty, staff, and students who had worked directly with the Whirlwind or TX-0 computers wanted more of that interactive access. Traditional computer system batch-processing approaches were very slow for program debugging and were challenged by machine overloading as digital computing became more popular.

In the spring of 1961, Professor Fernando Corbató, then associate director of MIT's Computation Center, began to design the Compatible Time-Sharing System (CTSS) for MIT's IBM 709 computer. Corbató initially worked with two of his Comp Center staff members, Robert Daley and Marjorie Merwin. They arranged for IBM to provide an interrupt capability for the 709 that allowed them to take control of the machine. They created a special version of the operating system that set aside 5 kilowords of memory (of 32 kw total) for the time-sharing supervisor (and for buffering typewriter terminal input and output). They used tape drives to store the programs and files of the users of the four terminals. It was crude, but that original configuration allowed a November 1961 demonstration of interactive

computer use. Thus, 2011 is the 50th anniversary of the conception and initial demonstration of CTSS.

By 1963, CTSS was a stable, large-scale system, operating with a large disk drive for file storage, more memory for a refined time-sharing supervisor, and a controller that handled a significant number of local and remote terminals. That summer, CTSS supported experimentation with time-shared computing at MIT by a stream of visiting computing pioneers. CTSS was proof positive of the feasibility of time sharing.

Corbató's substantive, solid, early implementation of time sharing in CTSS at MIT was an important stimulus for the era in computing that continues to this day—the era in which users themselves have direct contact with a computer, telling the computer what they want it to do from moment to moment.

In honor of the 50th anniversary of CTSS, the Computer Society's History Committee has prepared a commemorative brochure that is publicly available as article number 5 on the Society's website at www.computer.org/portal/web/volunteercenter/history.

COMPUTER SOCIETY CONNECTION

COMPUTER SOCIETY HISTORY ACTIVITIES

Most IEEE Computer Society publications and activities highlight the state of the art of computing technology and the computing profession, with occasional retrospective articles in the Society's journals and magazines. Founded in 1987, the Society's History Committee focuses exclusively on computing history. Members of the History Committee work on various projects relating to Society history and to computing history more generally.

The Computer Society's longest-running explicit history activity is *IEEE Annals of the History of Computing*, a print journal that is also now available in epub format. Founded by computing pioneers in 1979 (and operated by the Society since 1992), *Annals* has long pub-

lished first-person accounts from participants in the history of computing. More recently, *Annals* has become a preeminent journal of scholarly writing by professional historians on computing history. *Annals* encourages submissions from both computing practitioners and computing historians.

In recent years, the Computing Then department of the Society's Computing Now portal has highlighted and posted online historical content from *Annals* and other Computer Society sources. Read articles from Computing Then at www.computer.org/portal/web/computingthen.

Susan Graham Receives Kennedy Award

Susan L. Graham, a professor of computer science at the University of California, Berkeley, was recently honored with the IEEE Computer Society's 2011 Ken Kennedy Award winner for her contributions to computer programming tools that have significantly advanced software development. Her award citation reads, "For foundational compilation algorithms and programming tools; research and discipline leadership; and exceptional mentoring."

RESEARCH AND PROJECTS

Graham's research covers human-computer interaction, programming systems, and high-performance computing. Her work has led to the development of interactive tools that



Susan L. Graham is the Pehong Chen Distinguished Professor of Electrical Engineering and Computer Science Emerita at UC Berkeley.

enhance programmer productivity as well as new implementation methods for programming languages that improve software performance.

As a participant in the Berkeley Unix project, Graham and her

students built the Berkeley Pascal system and the widely used gprof program profiling tool. Her most recent projects include Harmonia, a language-based framework for interactive software development, and Titanium, a Java-based parallel programming language, compiler, and runtime system that supports high performance scientific computing on large-scale multiprocessors.

Graham currently serves as vice-chair of the Council of the Computing Community Consortium, which is sponsored by the National Science Foundation.

KEN KENNEDY AWARD

The Kennedy Award was established in 2009 to recognize substantial contributions to programmability and productivity in computing as well as significant community service and mentoring activities. The award was named for high-performance computing expert Ken Kennedy, founder of Rice University's computer science program. Previous recipients of the Kennedy Award include David Kuck of the University of Illinois, Urbana-Champaign and Francine Berman of Rensselaer Polytechnic Institute. Winners receive a \$5,000 honorarium. **C**

Innovative Technology for Computer Professionals
Computer

NEXT ISSUE

COMPUTING AND THE ARTS

CALL AND CALENDAR

CALLS FOR ARTICLES
FOR COMPUTER

Computer seeks submissions for a September 2012 special issue on modeling and simulation of smart and green computing systems.

Sustainable and efficient utilization of available energy resources is perhaps the fundamental challenge of the current century. Academic and industrial communities have invested significant efforts in developing new solutions to address energy-efficiency challenges in several areas, including IT and telecommunications, green buildings and cities, and the smart grid.

Modeling and simulation methodologies are necessary for the comprehensive performance evaluation that precedes costly prototyping activities for such complex, large-scale systems. This special issue aims to disseminate the latest advances in modeling and simulation of smart and green computing systems, which are critical from the perspective of sustainable economic growth and environmental conservation.

Topics of interest include modeling and simulations of energy-efficient computing systems, green communications systems, and smart grid applications.

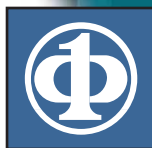
For author guidelines and information on how to submit a manuscript electronically, visit www.computer.org/portal/web/peerreviewmagazines/computer.

Articles are due by **1 March 2012**. Visit www.computer.org/portal/web/

SUBMISSION
INSTRUCTIONS

The Call and Calendar section lists conferences, symposia, and workshops that the IEEE Computer Society sponsors or cooperates in presenting.

Visit www.computer.org/conferences for instructions on how to submit conference or call listings as well as a more complete listing of upcoming computing-related conferences.



computingnow/cocfp9 to view the complete call for papers.

Computer seeks submissions for a September 2012 special issue on the move toward electronic health records.

The US Patient Protection and Affordable Care Act of 2010 embraces the notion that electronic health information is the bedrock of modern healthcare.

A multitude of projects are under way that support the transition to electronic health records (EHRs), which enable information exchange among various healthcare-related parties while maintaining patient privacy and security protection. This aggressive endeavor and its timeline of milestones poses new and interesting problems pertaining to the sharing of private information among government agencies, physicians, institutions, and individuals. While supported by incentive programs, the goal of having 80 percent of physicians using EHRs by 2014 seems unattainable given that currently less than 20 percent use this technology. Fine-tuning certification criteria and establishing best practices appear to be keys to this initiative's overall success.

Topics of interest include modeling and simulations of energy-efficient computing systems, green communications systems, and smart grid applications.

Articles are due by **15 May 2012**. Visit www.computer.org/portal/web/

computingnow/cocfp11 to view the complete call for papers.

CALLS FOR ARTICLES FOR
IEEE CS PUBLICATIONS

IEEE Computer Graphics and Applications plans a September/October 2012 special issue titled "Biomedical Applications: From Data Capture to Modeling."

Today's broad array of image and data-capture tools is dramatically changing the understanding of biological processes. Imaging modalities like computed tomography and magnetic resonance imaging let users visualize and track complex biological processes. Motion capture can help researchers to understand how animals move.

Just as calculus helped physicists understand and model the mechanical world, computers can help model complex biological systems for researchers to use in reasoning and making predictions about them. Computer graphics techniques and algorithms—from modeling to animation—make this possible.

This special issue is dedicated to multidisciplinary efforts in building, verifying, and understanding biological models. The guest editors seek contributions that address biology problems ranging from biochemistry through computational anatomy.

Articles are due by **14 January 2012**. Visit www.computer.org/portal/web/computingnow/cgacfp5 to view the complete call for papers.

CALL AND CALENDAR

EVENTS IN 2011-2012

December

5-8E-Science 2011
7-9ICPADS 2011
11-14ICDM 2011
18-21HiPC 2011

January 2012

4-7HICSS 2012
9-11WACV 2012

February 2012

1-3ICOIN 2012
2-3dMEMS 2012
25-29HPCA 2012

HPCA 2012

The 18th International Symposium on High-Performance Computer Architecture is a leading forum for scientists and engineers to present their latest research findings in this rapidly changing field. Authors are invited to submit papers on all aspects of high-performance computer architecture. In particular, organizers have solicited papers on topics that include processor, cache, and memory architectures; high-performance I/O systems; architectures for cloud-based HPC; and innovative hardware-software tradeoffs.

The conference is organized into paper presentations, workshops and tutorials, and industry sessions. A special session at HPCA-18 will highlight best papers from IEEE Computer Architecture Letters.

HPCA 2012 takes place **25-29 February** in New Orleans. Visit www.hpcaconf.org/hpca18 for complete conference details.

IEEE Internet Computing plans a November/December 2012 special issue on future Internet protocols.

The Internet is based on a set of layered protocols, their servers, and architectures that support them. The Internet has now evolved far beyond the original TCP/IP protocol and architecture. Application protocols have similarly evolved, as evidenced by HTML 5.0. Future applications and basic usage will require significant changes; however, a clean-slate approach is not likely to be adopted, making a feasible migration path a must for any new proposal.

Topics of interest include high-latency-tolerant TCP (also for mobile environments); TCP replacement protocols, requirements for app management, smart routers, and policies and protocols for cloud computing and enterprise management.

Articles are due by **1 March 2012**. Visit www.computer.org/portal/web/computingnow/iccfp6 to view the complete call for papers.

IEEE Security & Privacy plans a November/December 2012 special issue titled "Lost Treasures of Computer Security & Privacy."

This special issue of S&P will address key lessons from the past 50 years—not merely to recapitulate them, but to learn from them. The editors solicit articles from individu-

als and organizations about lessons learned from successful and unsuccessful attempts to define standards for measuring security. Also welcome are summaries of solid computer security science lost because of building with unpopular metric definitions as well as business failures.

Articles are due by **1 March 2012**. Visit www.computer.org/portal/web/computingnow/spcfp6 to view the complete call for papers.

CALENDAR

DECEMBER 2011

5-8 Dec: E-Science 2011, 7th Int'l Conf. on e-Science, Stockholm; www.escience2011.org

7-9 Dec: ICPADS 2011, IEEE Int'l Conf. on Parallel and Distributed Systems, Tainan, Taiwan; <http://conf.ncku.edu.tw/icpads>

11-14 Dec: ICDM 2011, IEEE Int'l Conf. on Data Mining, Vancouver, Canada; <http://webdocs.cs.ualberta.ca/~icdm2011/index.php>

18-21 Dec: HiPC 2011, IEEE Int'l Conf. on High-Performance Computing, Bangalore, India; www.hipc.org

JANUARY 2012

4-7 Jan: HICSS 2012, 36th Hawai'i Int'l Conf. on System Sciences, Manoa,

Hawai'i; www.hicss.hawaii.edu

9-11 Dec: WACV 2011, Workshop on the Applications of Computer Vision, Breckenridge, Colorado; www.wacv2012.org

FEBRUARY 2012

1-3 Feb: ICOIN 2012, Int'l Conf. on Information Networking, Bangkok; www.icoin.org

2-3 Feb: dMEMS 2012, 2nd Workshop on Design, Control and Software Implementation for Distributed MEMS, Besançon, France; <http://dmems.univ-fcomte.fr>

25-29 Feb: HPCA 2012, Int'l Symp. on High-Performance Computer Architecture, New Orleans; www.ece.lsu.edu/hpca-18

MARCH 2012

13-15 Mar: CAMP 2012, Int'l Conf. on Information Retrieval & Knowledge Management, Kuala Lumpur, Malaysia; <http://pgcs.upm.edu.my/camp12>

26-29 Mar: AINA 2012, 26th Int'l Conf. on Advanced Information Networking and Applications, Fukuoka, Japan; www.aina-conference.org

B. Ward, Editor; bnward@computer.org

GREEN IT

End-to-End Energy Management

Yung-Hsiang Lu, *Purdue University*

Qinru Qiu, *Syracuse University*

Ali R. Butt and Kirk W. Cameron, *Virginia Tech*



To improve energy efficiency, we must consider the end-to-end energy use of a task involving multiple computer systems.

The green IT community has been aflutter in recent months, with numerous workshops, summits, and meetings devoted to identifying and meeting the challenges of sustainable computing. Achieving consensus among such a growing, vibrant group of researchers is a challenge in itself. However, the stakes are high, with limited funding opportunities available.

An analysis of the literature in IEEE Xplore highlights a glaring gap in our understanding of energy efficiency in computer systems. Since 2005, researchers around the world have published more than 20,000 papers on energy management. Many if not most of these articles discuss techniques to improve the energy efficiency of individual components or systems: processors, memory, wireless networks, laptops, supercomputers, datacenters, handheld devices, and so on.

In reality, energy is task- not system-centric. Several disparate systems, or systems of systems, collectively use energy to accomplish a given task and satisfy service-level expectations. Consider, for example, someone who takes a photo with a smartphone and sends it to a friend.

Taking and transmitting the photo consumes energy from the smartphone, the data transfer consumes energy from the Internet routers, and the recipient's local system consumes energy to display the image.

To improve energy efficiency, we must consider the end-to-end energy use of a task involving multiple computer systems.

REDEFINING ENERGY EFFICIENCY

When systems are interconnected, defining energy efficiency isn't straightforward. Saving energy on one computer can possibly cause another to consume more energy, increasing overall carbon emissions. In the case of file sharing, for example, should a computer compress a file before sending it? Although compression reduces data size and thus saves network energy, the processor's energy consumption could outweigh this savings. The solution depends on many factors including the compression ratio, network data rate, and number of hubs between sender and receiver.

The systems in use impact these energy efficiency calculations. For a smartphone, the wireless network is a significant energy consumer. In con-

trast, a desktop's processor, memory, display, and hard drive together consume much more energy than the network. Moreover, if the desktop has a wired broadband connection, compression could result in negligible network energy savings.

Cloud computing further complicates energy management. For mobile users, battery life is the leading factor in customer satisfaction, and some studies suggest that offloading computation to cloud servers could extend battery lifetime (M. Satyanarayanan, "Mobile Computing: The Next Decade," *Proc. 1st ACM Workshop Mobile Cloud Computing and Services [MCS 10]*, ACM, 2010, article no. 5; E. Cuervo et al., "MAUI: Making Smartphones Last Longer with Code Offload," *Proc. 8th Int'l Conf. Mobile Systems, Applications, and Services [MobiSys 10]*, ACM, 2010, pp. 49-62; K. Kumar and Y-H. Lu, "Cloud Computing for Mobile Users: Can Offloading Computation Save Energy?," *Computer*, Apr. 2010, pp. 51-56).

However, this approach gives little or no consideration to network or server energy use. Extending smartphone battery life could come at the cost of more carbon emissions from the network and servers supporting the offloaded computation.

GREEN IT

For all these reasons, it's necessary to redefine energy efficiency. In high-performance systems, energy efficiency is often measured as gigaflops per watt—for example, the Green500 supercomputers are ranked on this basis. However, this metric doesn't capture many salient energy-efficiency aspects of connected systems, including mobile devices, access points, network backbones, and servers. The basic definition of energy efficiency is still the ratio between the amount of completed work and the energy used, but the numerator and denominator have become much more complex.

From an environmental impact perspective, any energy dissipation related to computation should be counted. Definitions of energy use should thus include the charging and discharging of the battery of mobile devices, the energy for cooling servers, and the energy lost in power distribution systems. New definitions of the amount of work performed should include meaningful examples such as the number of minutes videos can be watched on a mobile phone.

END-TO-END ENERGY MANAGEMENT

New energy-efficiency techniques must reflect the reality that computer systems no longer exist in isolation. *End-to-end energy management* considers the effects of energy management on the computers that comprise a system of systems.

Many researchers estimate the energy dissipation and carbon footprint of different tasks, but these estimates are based on empirical data, and validation is nearly impossible because no single entity owns the entire system of systems. Even when large companies own many system pieces, energy-efficiency details are often trade secrets. Consequently, existing estimations aren't precise enough to validate energy-efficiency management techniques targeting systems of systems (K.W. Cameron,

"The Challenges of Energy-Proportional Computing," *Computer*, May 2010, pp. 82-83).

To achieve end-to-end energy management, academia and industry must recognize that true energy efficiency can be accomplished only by considering the complex interactions within systems of systems. Given the scale of these challenges, researchers and engineers must work together to develop new benchmarks, metrics, models, and measurement techniques.

True energy efficiency can be accomplished only by considering the complex interactions within systems of systems.

Benchmarks

Benchmarks need to incorporate networks and heterogeneous computers as intrinsic elements. In addition to throughput-centric workloads, system optimizations must consider the myriad effects of energy management—for example, whether slight performance degradation that saves significant energy in a server would cause unacceptable disruptions to mobile users.

Metrics

Current metrics such as Gflops/W or energy-delay products are insufficient. In fact, any calculation that summarizes complex interactions in a single number is probably undesirable. Metrics should acknowledge that energy consumption and other factors can affect one another—for example, that there is a tradeoff between energy and reliability (G. Wang, A.R. Butt, and C. Gniady, "On the Impact of Disk Scrubbing on Energy Savings," *Proc. 2008 Workshop Power Aware Computing and Systems* [HotPower 08], Usenix, 2008, pp. article no. 16). Energy-reduction solutions should also adapt to users' preferences: a

user wants to maximize battery lifetime on a road trip but performance at the office.

Models

Researchers should investigate new techniques to model how changes in energy use at one location affect the performance and energy dissipation of the rest of the system. Such models could include a map profile of energy distribution over the network. As the application environment, network traffic, and server workload change, the causal relationships and their impact also vary; hence, the models become time-variant. This makes model development particularly challenging. Multi-agent-based distributed techniques should be considered to improve the scalability and reduce the complexity of the models (Y. Ge, Q. Wu, and Q. Qiu, "A Multi-Agent Framework for Thermal Aware Task Migration in Many-Core Systems," *IEEE Trans. VLSI Systems*, Aug. 2011; doi: 10.1109/TVLSI.2011.2162348).

Measurement techniques

New measurement techniques are urgently needed to support scientific evaluation of end-to-end energy management (K. Kant, "Toward a Science of Power Management," *Computer*, Sept. 2009, pp. 99-101). Building such facilities requires sophisticated equipment and expertise in many fields.

In the simplest sense, performance can be measured with a stopwatch. In contrast, energy consumption must be measured using equipment that might significantly alter a system's very nature—a mobile phone connected to a power meter is no longer mobile. Furthermore, the lack of "energy counters" in chips and circuit boards means that measurement must be intrusive.

End-to-end energy management is a new frontier ripe with opportunities for eager

researchers. Because we currently lack the infrastructure to study energy efficiency in systems of systems, most end-to-end techniques must rely on disparate empirical data. In addition, we have neither accepted metrics nor accepted models of performance and energy efficiency for interconnected systems.

We've been in this situation before. As we entered the current millennium, there was a dearth of infrastructure and tools to enable energy management in isolated systems. Yet, in the past six years alone, research has led to vast improvements in mobile phone battery life

and capability as well as mobile laptops that outperform workstations from just a few years ago. And as before, the necessity of addressing the challenges of end-to-end energy management is likely to drive innovation in the coming years to ensure that we meet our sustainable computing goals. **□**

Yung-Hsiang Lu is an associate professor in the School of Electrical and Computer Engineering at Purdue University. Contact him at yunglu@ecn.purdue.edu.

Qinru Qiu is an associate professor in the L.C. Smith College of Engineer-

ing and Computer Science at Syracuse University. Contact her at qiqiu@syr.edu.

Ali R. Butt is an assistant professor in the Department of Computer Science at Virginia Tech. Contact him at butta@cs.vt.edu.

Kirk W. Cameron, Green IT editor, is an associate professor in the Department of Computer Science at Virginia Tech. Contact him at greenit@computer.org.

cn Selected CS articles and columns are available for free at <http://ComputingNow.computer.org>.

New Computer Architecture Titles from Morgan Kaufmann

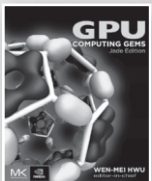


Computer Architecture, 5th Edition A Quantitative Approach

John L. Hennessy & David A. Patterson
ISBN: 9780123838728 | \$89.95

Key Features

- Updated to cover the mobile computing revolution.
- Emphasizes the two most important topics in architecture today: memory hierarchy and parallelism in all its forms.
- Develops common themes throughout each chapter: power, performance, cost, dependability, protection, programming models, and emerging trends ("What's Next").
- Includes three review appendices in the printed text. Additional reference appendices are available online.
- Includes updated Case Studies and completely new exercises.

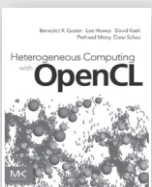


GPU Computing Gems Jade Edition

Wen-mei W. Hwu
ISBN: 9780123859631 | \$74.95

This is the second volume of Morgan Kaufmann's GPU

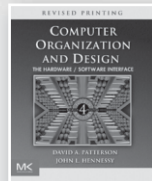
Computing Gems, offering an all-new set of insights, ideas, and practical "hands-on" skills from researchers and developers worldwide.



Heterogeneous Computing with OpenCL

Benedict Gaster, Lee Howes, David R. Kaeli, Perhaad Mistry & Dana Schaa
ISBN: 9780123877666 | \$69.95

Learn parallel programming with CPUs, GPUs, and APUs, from OpenCL community leaders.



Computer Organization and Design, Revised 4th Edition The Hardware/Software Interface

David A. Patterson & John L. Hennessy
ISBN: 9780123747501 | \$89.95

Key Features

- The Revised Fourth Edition of Computer Organization and Design has been updated with new exercises and improvements throughout suggested by instructors teaching from the book.
- The companion CD provides a toolkit of simulators and compilers along with tutorials for using them, as well as advanced content for further study and a search utility for finding content on the CD and in the printed text.

MK
MORGAN KAUFMANN

mkp.com

Scan the QR code to view all MK's Computer Architecture Titles!



Prices subject to change.

IEEE computer society

PURPOSE: The IEEE Computer Society is the world's largest association of computing professionals and is the leading provider of technical information in the field.

MEMBERSHIP: Members receive the monthly magazine *Computer*, discounts, and opportunities to serve (all activities are led by volunteer members). Membership is open to all IEEE members, affiliate society members, and others interested in the computer field.

COMPUTER SOCIETY WEBSITE: www.computer.org

OMBUDSMAN: To check membership status or report a change of address, call the IEEE Member Services toll-free number, +1 800 678 4333 (US) or +1 732 981 0060 (international). Direct all other Computer Society-related questions—magazine delivery or unresolved complaints—to help@computer.org.

CHAPTERS: Regular and student chapters worldwide provide the opportunity to interact with colleagues, hear technical experts, and serve the local professional community.

AVAILABLE INFORMATION: To obtain more information on any of the following, contact Customer Service at +1 714 821 8380 or +1 800 272 6657:

- Membership applications
- Publications catalog
- Draft standards and order forms
- Technical committee list
- Technical committee application
- Chapter start-up procedures
- Student scholarship information
- Volunteer leaders/staff directory
- IEEE senior member grade application (requires 10 years practice and significant performance in five of those 10)

PUBLICATIONS AND ACTIVITIES

Computer: The flagship publication of the IEEE Computer Society, *Computer*, publishes peer-reviewed technical content that covers all aspects of computer science, computer engineering, technology, and applications.

Periodicals: The society publishes 13 magazines, 18 transactions, and one letters. Refer to membership application or request information as noted above.

Conference Proceedings & Books: Conference Publishing Services publishes more than 175 titles every year. CS Press publishes books in partnership with John Wiley & Sons.

Standards Working Groups: More than 150 groups produce IEEE standards used throughout the world.

Technical Committees: TCs provide professional interaction in more than 45 technical areas and directly influence computer engineering conferences and publications.

Conferences/Education: The society holds about 200 conferences each year and sponsors many educational activities, including computing science accreditation.

Certifications: The society offers two software developer credentials. For more information, visit www.computer.org/certification.

NEXT BOARD MEETING

13–14 Nov., New Brunswick, NJ, USA

EXECUTIVE COMMITTEE

President: Sorel Reisman*

President-Elect: John W. Walz*

Past President: James D. Isaak*

VP, Standards Activities: Roger U. Fujii†

Secretary: Jon Rokne (2nd VP)*

VP, Educational Activities: Elizabeth L. Burd*

VP, Member & Geographic Activities: Rangachar Kasturi†

VP, Publications: David Alan Grier (1st VP)*

VP, Professional Activities: Paul K. Joannou*

VP, Technical & Conference Activities: Paul R. Croll†

Treasurer: James W. Moore, CSDP*

2011–2012 IEEE Division VIII Director: Susan K. (Kathy) Land, CSDP†

2010–2011 IEEE Division V Director: Michael R. Williams†

2011 IEEE Division Director V Director-Elect: James W. Moore, CSDP*

*voting member of the Board of Governors †nonvoting member of the Board of Governors

BOARD OF GOVERNORS

Term Expiring 2011: Elisa Bertino, Jose Castillo-Velázquez, George V. Cybenko, Ann DeMarle, David S. Ebert, Hironori Kasahara, Steven L. Tanimoto

Term Expiring 2012: Elizabeth L. Burd, Thomas M. Conte, Frank E. Ferrante, Jean-Luc Gaudiot, Paul K. Joannou, Luis Kun, James W. Moore

Term Expiring 2013: Pierre Bourque, Dennis J. Frailey, Atsuhiko Goto, André Ivanov, Dejan S. Milojicic, Jane Chu Prey, Charlene (Chuck) Walrad

EXECUTIVE STAFF

Executive Director: Angela R. Burgess

Associate Executive Director; Director, Governance: Anne Marie Kelly

Director, Finance & Accounting: John Miller

Director, Information Technology & Services: Ray Kahn

Director, Membership Development: Violet S. Doan

Director, Products & Services: Evan Butterfield

COMPUTER SOCIETY OFFICES

Washington, D.C.: 2001 L St., Ste. 700, Washington, D.C. 20036-4928
Phone: +1 202 371 0101 • Fax: +1 202 728 9614

Email: hq.ofc@computer.org

Los Alamitos: 10662 Los Vaqueros Circle, Los Alamitos, CA 90720-1314
Phone: +1 714 821 8380

Email: help@computer.org

MEMBERSHIP & PUBLICATION ORDERS

Phone: +1 800 272 6657 • Fax: +1 714 821 4641 • Email: help@computer.org

Asia/Pacific: Watanabe Building, 1-4-2 Minami-Aoyama, Minato-ku, Tokyo 107-0062, Japan

Phone: +81 3 3408 3118 • Fax: +81 3 3408 3553

Email: tokyo.ofc@computer.org

IEEE OFFICERS

President: Moshe Kam

President-Elect: Gordon W. Day

Past President: Pedro A. Ray

Secretary: Roger D. Pollard

Treasurer: Harold L. Flescher

President, Standards Association Board of Governors: Steven M. Mills

VP, Educational Activities: Tariq S. Durrani

VP, Membership & Geographic Activities: Howard E. Michel

VP, Publication Services & Products: David A. Hodges

VP, Technical Activities: Donna L. Hudson

IEEE Division V Director: Michael R. Williams

IEEE Division VIII Director: Susan K. (Kathy) Land, CSDP

President, IEEE-USA: Ronald G. Jensen



revised 24 August 2011

IDENTITY SCIENCES

Human Ear Recognition

Arun Ross

West Virginia University

Ayman Abaza

West Virginia High Tech Consortium Foundation



Ear recognition technology is a potentially valuable tool in the biometric arsenal.

In a world where social interaction is increasingly digital in nature (Facebook, Google+, Skype) and financial transactions are routinely conducted over the Internet (online banking), reliably establishing an individual's identity is of paramount importance. Several law enforcement and military applications also need a dependable method to identify people—for example, to determine if an encountered individual is a potential threat or criminal suspect.

The limitations of traditional modes of authentication based on ID cards and passwords have led to the development of sophisticated biometric systems that establish human identity using an individual's physical or behavioral attributes, such as fingerprints, face, iris, hand geometry, voice, or gait. Biometric systems are now being incorporated in various applications ranging from personal laptop access to international border control. The US-VISIT program, for example, employs fingerprint recognition to determine if a traveler to the US is on a government watch list. Similarly, the United Arab Emirates uses the Iris Expellee Tracking System to identify and apprehend deported individuals who attempt to reenter the country using false travel documents.

In spite of tremendous biometric advances, identifying noncooperative individuals in public spaces and other unconstrained environments remains a challenging problem. Only partial or corrupted biometric information might be available—for example, a surveillance video might capture only a portion of an individual's face.

To improve human recognition, biometric researchers are exploring the use of ancillary characteristics such as scars, marks, tattoos, height, and body shape in conjunction with primary features like the face. The ear is one such promising “soft” biometric.

The external ear flap, known as the pinna, has several morphological components as Figure 1a shows. While its structure is relatively simple, it varies significantly across individuals. Figure 1b shows examples of these variations, which, along with the ear's size, color, and texture can serve as a distinguishing characteristic. Changes in facial expression and age do not significantly impact the ear's appearance, although the effect of gravity and ear accessories can perturb the length of the ear lobe.

EARLY RESEARCH

The ear's potential for use in human identification was recognized as early as the 1880s by Alphonse

Bertillon, a French police officer who pioneered the use of physical measurements to identify criminals. Bertillon combined qualitative and quantitative descriptions of various body parts, including the ear, in what he called anthropometry (*Identification anthropométrique: instructions signalétiques*, 1885).

In 1906, R. Imhofer, a doctor in Prague, studied a set of 500 ears and noted that he could clearly distinguish between them based on only four features (“Die Bedeutung der Ohrmuschel für die Feststellung der Identität,” *Archiv für die Kriminologie*, vol. 26, pp. 150-163).

More than 50 years later, a team of researchers visually assessed 206 sets of ear photographs of newborn babies and concluded that the morphological constancy of the ear could be used to establish a newborn's identity (C. Fields et al., “The Ear of the Newborn as an Identification Constant,” *Obstetrics and Gynecology*, July 1960, pp. 98-102).

Between 1948 and 1962, Alfred Iannarelli collected ear photographs of thousands of individuals and extracted 12 different geometric measurements of the ear based on the crus of helix (*The Iannarelli System of Ear Identification*, Foundation Press, 1964), as Figure 2 shows. Iannarelli claimed that this set of measurements

IDENTITY SCIENCES



Figure 1. External anatomy of the ear. (a) The external flap, referred to as the pinna, has several morphological components: (1) helix rim, (2) lobule, (3) antihelix, (4) concha, (5) tragus, (6) antitragus, (7) crus of helix, (8) triangular fossa, and (9) incisure intertragica. (b) The pinna's structure varies across individuals. Examples of right (top row) and left (bottom row) ear images.

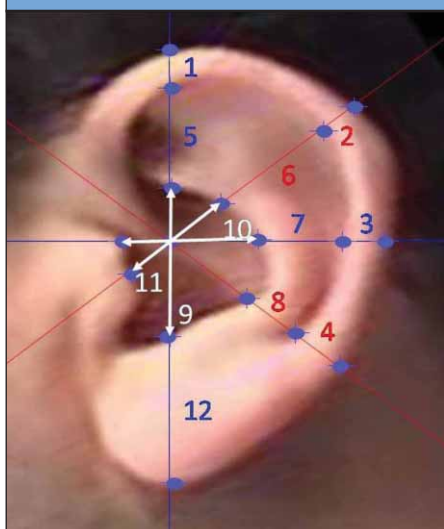


Figure 2. The Iannarelli identification system entails the extraction of 12 geometric measurements of the ear based on the crus of helix.

was reasonably unique across individuals.

EAR BIOMETRICS

An ear biometric system can be viewed as a typical pattern recognition system that reduces an input image to a set of features and then compares this against the feature sets of other images to determine its identity. Ear recognition can be accomplished using either a 2D digital image of the ear or a 3D point cloud that captures the ear's surface.

Ear recognition involves four steps. *Ear detection.* The first step is to localize the ear's position in an image. The system typically uses a rectangular boundary to indicate the ear's spatial extent in the side profile of a face image. Ear detection is critical because errors at this stage can undermine the system's utility.

Feature extraction. While the system can directly use the segmented ear during the matching stage, most systems extract a salient set of features to represent the ear. Feature extraction reduces the segmented ear to a mathematical model—for example, a feature vector—that summarizes the discriminatory information present in the ear image.

Matching. The system compares the features extracted from the input ear image to those stored in the database to establish the ear's identity. In its simplest form, matching generates scores indicating the similarity to other ear images.

Decision. The system uses the match scores to render a final decision. In verification mode, a “yes” indicates a genuine match and a “no” an impostor. In identification mode, the output is a list of potential matching identities ranked by match score.

AUTOMATED EAR RECOGNITION

Mark Burge and Wilhelm Burger reported the first attempt to automate the ear recognition process in 1997 (“Ear Biometrics for Computer Vision,” *Proc. 21st Workshop Austrian Assoc. for Pattern Recognition*, 1997, pp. 275-282). They used a mathematical graph model to represent and match the curves and edges in a 2D ear image.

Two years later, Belén Moreno, Ángel Sanchez, and José Vélez described a fully automated ear recognition system based on various features such as ear shape and wrinkles (“On the Use of Outer Ear Images for Personal Identification in Security Applications,” *Proc. 33rd Ann. Int'l Carnahan Conf. Security Technology*, IEEE, 1999, pp. 469-476).

Since then, researchers have proposed numerous feature extraction and matching schemes, based on computer vision and image processing algorithms, for ear recognition. These range from simple appearance-based methods such as principal component analysis and independent component analysis to more sophisticated techniques based on scale-invariant feature transforms, local binary patterns, wavelet transforms, and force fields. (D.J. Hurley, M.S. Nixon, and J.N. Carter, “Force Field Feature Extraction



Figure 3. Occlusion due to accessories and hair can lower or inhibit ear recognition system performance.

for Ear Biometrics,” *Computer Vision and Image Understanding*, June 2005, pp. 491-512).

In 2005, Hui Chen and Bir Bhanu presented a 3D ear recognition system that exploited the depth and structure of the ear’s morphological components (“Contour Matching for 3D Ear Recognition,” *Proc. 7th IEEE Workshop Applications of Computer Vision [WACV 05]*, IEEE, pp. 123-128).

IMPROVING MATCHING ACCURACY

As Figure 3 shows, occlusion due to hair and accessories can lower or inhibit ear recognition system performance. Changes in external lighting and variations in facial pose with respect to the camera can also have a negative impact.

In addition, the recognition accuracy of ear recognition algorithms has predominantly been evaluated using ear images acquired under ideal conditions, such as an indoor environment with highly controlled lighting. This has generated criticism that the matching accuracy of these algorithms, as reported in the literature, could be overly optimistic.

Nevertheless, ear recognition technology is a potentially valuable tool

in the biometric arsenal. For example, forensic examiners reviewing surveillance videotapes in the Netherlands used the ear biometric to identify suspects in gas station robberies who had covered their faces, but not their ears, (A.J. Hoogstrate, H.V.D. Heuvel, and E. Huyben, “Ear Identification Based on Surveillance Camera Images,” *Science & Justice*, July 2001, pp. 167-172).

To improve matching accuracy, researchers are exploring the possibility of combining images of the ear and the face. Even if the ear cannot be used to verify human identity in a given situation, it could *exclude* an identity from being considered as a potential match if it is sufficiently different from the input probe image.

EARPRINTS

The use of 2D or 3D ear images for human recognition differs from the use of *earprints*: marks left by secretions from the outer ear when someone presses up against a wall, door, or window. Earprints have been introduced as physical evidence in several criminal cases in the US and other countries, although some convictions that relied on earprints have been overturned. Earprints haven’t

been widely accepted in court due to a lack of scientific consensus as to their individuality.

Currently, there are no commercially available ear recognition systems. However, the future holds tremendous potential for incorporating ear images with face images in a multibiometric configuration, even as researchers continue to refine the technology. For example, assigning an ear image to one of several predefined categories could allow for rapid retrieval of candidate identities from a large database. In addition, the use of ear thermograms could help mitigate the problem of occlusion due to hair and accessories. As the technology matures, both forensic and biometric domains will benefit from this biometric. **C**

Arun Ross is an associate professor in the Lane Department of Computer Science and Electrical Engineering at West Virginia University. Contact him at arun.ross@mail.wvu.edu.

Ayman Abaza is a senior scientist in the Advanced Technology Group at the West Virginia High Technology Foundation. Contact him at aabaza@wvhtf.org.

Editor: Karl Ricanek Jr., director of the Face Aging Group at the University of North Carolina Wilmington; ricanekk@uncw.edu

cn Selected CS articles and columns are available for free at <http://ComputingNow.computer.org>.

build your career
IN COMPUTING

www.computer.org/buildyourcareer



ONLINE PLUS™
publishing evolved

A new publication model that provides subscribers with our best features and added benefits over traditional print such as:

- Lower Price
- Interactive Disk and Printed Book of Abstracts
- Full Archival Online Access to the CSDL
- Available Print on Demand

Available Transactions Titles by 2012:

• TDSC	• TPDS (NEW!)
• TMC (NEW!)	• TVCG
• TPAMI	

Renew your membership and subscription now!
www.ieee.org/renew

For more information about OnlinePlus™, please visit <http://www.computer.org/onlineplus>.

IEEE  **IEEE**

IEEE  **computer society**



NEW ISSUE ALERTS!

Stay connected with the IEEE Computer Society Transactions by signing up for our new and improved Issue Alerts. They are free and contain valuable information like:

- News about your favorite transactions,
- Contributions from the Editorial Board,
- Information about related conferences,
- Top 5 recent downloaded papers,
- And much more.

Not a subscriber? Don't worry. You can still sign up to receive news about the transactions.

Visit <http://www.computer.org/newsletter> to sign up today!



IEEE  **computer society**

INDUSTRY PERSPECTIVE

Opportunities in the Mobile Search Market

José Luis Gómez-Barroso, *National University of Distance Education, Spain*

Claudio Feijóo, *Technical University of Madrid, Spain*

Ramón Compañó, *Institute for Prospective Technological Studies, Spain*



Success in the mobile search market will come to those who provide value-added apps that exploit unique mobile functionalities, especially those related to personalized and context-based services.

If the lessons of the desktop-era fixed Internet experience hold, search engines will play a preponderant role in configuring mobile Internet markets. But as the dominant forces in the computer search market try to transfer their hegemony, they're quickly learning that mobile search must consider additional context variables unique to portable devices. Success isn't guaranteed to the field's current dominant players because mobile search opens up new options for creative applications.

A SHIFTING BATTLEGROUND

In April 2007, Google reached an agreement to acquire the online advertising company DoubleClick for US\$3.1 billion in cash. Acquiring DoubleClick expanded Google's

influence far beyond algorithm-driven ad auctions into a relationship-based business with Web publishers and advertisers. The DoubleClick buyout certainly isn't the sole factor—or even possibly the primary one—in the company's evolution, but Google's revenues in 2008 were twice those in 2006, and the move expanded the company's dominance of the search engine market. Just three years later, in May 2010, Google invested \$750 million in the buyout of AdMob, a major mobile advertising platform that claimed to serve more than 8.5 billion mobile banner and text ads per month across thousands of mobile networks, websites, and applications.

Google's purchase of AdMob indicates that the battleground of the extremely lucrative network-provided applications and services markets is

shifting. With an increasing number of users demanding ubiquity and permanent availability, the new scenario will undoubtedly become more mobile. But has the shift already happened?

Susan Wojcicki, vice president of product management at Google, gave some clues in a blog post she wrote following the AdMob acquisition: "Over the past two years, Google's mobile search volumes have grown more than fivefold, at an accelerated pace. In the first three months of 2010, people with smartphones with 'full' WebKit browsers (such as the iPhones, Android devices, and Palm Pre) searched 62 percent more than they did in the previous three months" (<http://googleblog.blogspot.com/2010/05/weve-officially-acquired-admob.html>). Moreover, Eric Schmidt,

INDUSTRY PERSPECTIVE

former chief executive of Google, has mentioned several times since 2008 that Google can make more money in mobile than it can on the desktop.

All of this activity is impressive, but what's truly stunning is the first official statistical information on wireless broadband. In December 2010, the Organization for Economic Cooperation and Development released statistics on broadband penetration in OECD countries as of June 2010, reporting a wireless broadband indicator for the first time (www.oecd.org/sti/ict/broadband). There were 294 million fixed broadband subscriptions but the number of mobile broadband subscriptions (those included in a mobile voice plan) and dedicated mobile data subscriptions (which require an additional data plan) totaled nearly 435 million.

In the fixed Internet, search engines serve as the access gate to all sorts of content and applications, providing those who control that access with a lot of power. For the mobile Internet, the strategy seems obvious: learn from history and repeat it. However, a series of circumstances makes a literal translation impossible. Success will come to those who provide value-added apps that exploit unique mobile functionalities, especially those related to personalized and context-based services.

MOBILE SEARCH TECHNOLOGY

At first glance, it might appear that mobile search is merely search on a mobile gadget. Adding some "mobility" enhancement function, such as refining results by taking into account the user's location or adapting them to the type of display, would still involve a search-as-usual strategy that extends the same desktop-based approach, systems, and algorithms to a new platform with specific features and limitations.

But in reality, mobile search should exploit contextual infor-

mation, such as relevant data embedded in the mobile device, information in the surrounding environment, and the user's profiles or behavioral patterns, to improve search result relevance or to provide a more valuable and entertaining user experience. Such specificity could flourish in an environment in which hardware becomes "senseware," information coats objects and people, and ubiquitous location-aware social networks enhance the available information, sorting it on the user's behalf.

Search functionalities will be tightly embedded into the value chain of wider mobile services, which themselves can be numerous and complex.

In this scenario, mobile devices become the entry point to a networked environment in which "intelligence" is distributed across different elements. To make this vision a reality, three groups of technologies will likely have a direct impact on mobile search.

The first group comprises generic search technologies for retrieving accurate and enriched content. Such technologies could include semantic approaches, cognitive approaches, and multimedia retrieval.

The second group comprises specific mobile search technologies that would render mobile data acquisition—both its processing and its matching—more context aware or introduce augmented reality technologies to enrich context awareness.

Finally, the third group would include any technology components that can enable mobile applications. These would include wireless networks (broadband access ubiquity and dynamic spectrum management),

sensor networks (RFID and Internet of Things), devices (multimedia capabilities, location, interoperability, and openness), and cloud computing (Web browsers, connectivity, security, and data protection).

MONETIZING MOBILE SEARCH

Interestingly, no significant bottlenecks seemingly stand in the way of introducing new mobile search applications from a technological viewpoint. Most of the building blocks are either already available or in an advanced prototype stage. The main difficulty lies in how to better integrate existing technologies.

That said, the real challenge will be how to monetize new mobile search applications. Advertising seems like a natural choice, particularly for search-as-usual applications; obviously, the strategies and formats must be adapted to the mobile environment.

Other sources of revenue are possible, but two basic factors will influence a new scheme's business model: the feasibility of monetizing the added value that mobile search provides within a given application, and the economic value of the search functionality.

Future search-based applications will neither be simple nor autonomous. Rather, search functionalities will be tightly embedded into the value chain of wider mobile services, which themselves can be numerous and complex. In technoeconomic terms, search functionality is a key constituent in an ecosystem in which industrial players compete or collaborate to generate successful and scalable business strategies in a highly dynamic and still emerging market landscape.

Additional factors must be considered when determining the sustainability of new types of mobile search ventures. First, the stakeholders are diverse and heterogeneous—device manufacturers,

mobile network operators, infrastructure providers, mobile OS providers, Web search players, and mobile-specific search players all feature prominently. The variety of players, technologies, and approaches also complicates interoperability and increases transaction costs.

Second, this ecosystem is embedded in an institutional framework, so the success of search-based applications depends on regulatory environment factors, ranging from international data roaming costs to spectrum allocation issues to privacy regulations involved with personal data collection.

Data roaming in particular is relevant for search-based mobile applications because of the usefulness and innovative proposals it can offer to users on the move. International data roaming is slowly being solved regionally, with larger initiatives (some imposed by regulation, some originating in market forces) being introduced in Europe, Africa, and Asia.

Privacy regulation is a notoriously immature yet controversial issue. On one hand, privacy by design gives users control over their personal data through technology and contractual provisions; on the other, privacy protection requires a minimum set of mandatory rules to defend consumer interests. Depending on the country, users have very different privacy controls. US legislators have taken a utilitarian approach to data protection, whereas European legislators tend to define privacy as a fundamental right.

WHAT IS THE MARKET SAYING?

As mobile search expands rapidly and steadily, established players are taking up as much of the market as they can to help them evolve smoothly into offering new and smarter search technology as needed. Is it worth it for newer or smaller competitors to enter the fray? The field is certainly

large enough: user demand for opt-in, highly personalized, location- and social-aware search services isn't yet satisfied.

Services aren't yet fully interoperable, don't link multiple dynamic databases, and don't morph according to context. Furthermore, current interfaces don't allow dynamic usage situations. Voice-, touch- and movement-based interfaces should seamlessly support users in accessing information in situations that change based not only on location but also on interactions with other

Services aren't yet fully interoperable, don't link multiple dynamic databases, and don't morph according to context.

devices, users, and available services as well as needs, activities, and preferences. Most of these possibilities remain commercially unexplored.

From the mobile business perspective, a secret war of uncertain result is under way—mobile browsers versus mobile apps—in which the users are the unaware army. If the browser wins, mobile devices will become a convenient wireless extension of the fixed Internet, with advertising as the main financial model. If apps dominate, we'll see more value-added innovations, but at the cost of fragmented solutions available for users—which could be either device- or OS-dependent.

The two approaches have other differences as well. Application development appears to be more agile, but it's very dependent on the technoeconomic evolution of current platforms. Browser standards seem more stable, but browsers are better suited to cloud computing, which is much less defined.

The challenge remains in bridging data and information needs and offering useful services that entice people to pay for them. Innovations are the key to fulfilling these expectations, and they depend on a conjunction of technological, economic, social, and regulatory aspects, along with a bit of luck—many of the most successful mobile industry apps have evolved from an initial user base in ways totally unforeseen by the apps' original designers.

In the case of mobile search, all these aspects are relevant, encompassing both hurdles and uncertainties. This is the daily scenario faced by innovators wanting to open up a space in the marketplace. **■**

José Luis Gómez-Barroso is a professor at the National University of Distance Education, Spain, and a PURC senior research associate, University of Florida. Contact him at jlgoomez@cee.uned.es.

Claudio Feijóo is a professor at the Technical University of Madrid, Spain. Contact him at cfeijoo@cedint.upm.es.

Ramón Compañó is the program manager at the Institute for Prospective Technological Studies—Joint Research Centre of the European Commission. Contact him at ramon.compano@ec.europa.eu.

Editor: Sumi Helal, Department of Computer and Information Science and Engineering, University of Florida; helal@cise.ufl.edu

Readers are encouraged to use the message board at www.computer.org/industry_perspective to post comments, offer feedback, or ask questions.

cn Selected CS articles and columns are available for free at <http://ComputingNow.computer.org>.



Keep up with the latest IEEE Computer Society publications and activities wherever you are.
Follow us on Twitter, Facebook, Linked In, and YouTube.



@Computer Society, @ComputingNow



facebook.com/IEEEComputerSociety
facebook.com/ComputingNow



IEEE Computer Society, Computing Now



youtube.com/ieeecomersociety

IEEE  computer society

HARD ISSUES

If Anything in This Life Is Certain, It's That You Can Kill Any ISA

Shubu Mukherjee, *Cavium*



Which ISA among the four popular ones—x86, PowerPC, ARM, and MIPS—will survive and which markets will they win?

In *The Godfather Part II*, Michael Corleone says, “If anything in this life is certain, if history has taught us anything, it’s that you can kill anyone.” This thought appropriately applies to instruction set architectures.

An ISA is the hardware interface a processor offers to software. Typically, the processor compiles programs written in a high-level language, such as C or C++, down to machine or binary code that conforms to a particular ISA and runs on the corresponding processor supporting the ISA.

Example ISAs include the ubiquitous x86 from Intel and AMD and the ARM ISA from ARM Holdings. Interestingly, with the emergence of new technological requirements such as lower power dissipation, new domains such as high-speed networking and smartphones, and new applications such as iPhone apps, the ISA debate has resurfaced. The question is which ISA among the four popular ones—x86, PowerPC, ARM, and MIPS—will survive and which markets will they win?

THE RISE AND FALL OF ISAS

We’ve seen the rise and fall of ISAs in the past. The VAX ISA from DEC failed to survive, except in a niche

market. Intel’s Itanium ISA is arguably on its way out. In the late 1980s and early 1990s, there was a raging debate on which ISA was the best, largely fueled by the RISC (reduced ISA) versus CISC (complex ISA) controversy. However, that debate withered soon after Intel demonstrated that a RISC ISA could be implemented underneath a CISC architecture.

One of the pillars of Intel’s success so far has been the x86 instruction set, which continues to evolve as it provides binary compatibility with existing software. Because consumers don’t want to pay for new software every time they buy new hardware, binary compatibility for old software was a big factor in x86’s success.

DEC, on the other hand, made the fundamental mistake of moving to a new ISA—from VAX to Alpha—without providing binary compatibility in hardware between the legacy VAX binaries and the Alpha ISA. However, Intel not only provides binary compatibility for legacy software, but also continues to evolve the x86 ISA, filing patents on the extensions. It’s the patent protection on the x86 ISA extensions that prevents others from legally manufacturing an x86 processor without a license from Intel.

Economic forces that dictated the success of the x86 ISA through binary compatibility are, however, changing rapidly. There’s now significant economic incentive to not be on the x86 ISA for certain rapidly evolving domains. The laptop and server x86 processor sockets typically have a power footprint of 60-80 watts and 130 watts, respectively. Compare this to smartphone or tablet processors that need to be in the 0.1 to a few watts range. With the proliferation of smartphones and the emergence of iPads, the demand for low-power processors has been increasing exponentially.

Regrettably, what prompted x86’s success as an ISA has now become its Achilles’ heel. To maintain binary compatibility, x86 has been burdened with legacy instructions and features. Further, the x86 CISC ISA’s instruction decoding logic is extremely complex and less streamlined than other RISC-like ISAs. Although Intel offers the Atom line of x86 processors that run in the 1-2 watt range, the performance/watt and performance/area characteristics of comparable RISC processors, such as ARM or MIPS, are arguably better, and they’re gaining rapid adoption in the low-power iPhone, iPad, tablet, netbook, and Android markets.

HARD ISSUES

This new genre of tablets and smartphones and the emergence of multicore processors for networking during the past decade have opened up new avenues for ISA innovation. For example, to improve code density and power efficiency, ARM introduced the Thumb instruction set (with enhancements to Thumb in later versions), a 16-bit subset of the ARM instruction set. Similarly, MIPS has introduced the microMIPS instruction set.

In multicore processors targeted for networking, the thirst for increased router performance has been growing exponentially due to smartphone data consumption, video streaming from companies like Netflix, the emergence of cloud services such as those offered by Microsoft, and continued demands for data-intensive applications. Popularly known as “data plane” processors, these multicore processors must soon handle 100-Gbps wired connections (with 100 Gigabit Ethernet) and more than 100 Mbps with 4G wireless technology.

To meet the bandwidth demands for the seven-layer network protocol stack, these multicore processors now offer aggressive hardware support for what was done in software in the past. For example, Cavium has several cryptographic engines in its processor core to support network processing, such as IPsec, that require packet encryption and decryption.

These new domains are a direct threat to the traditional “control plane” processors, such as x86, which typically provide better performance for single-thread applications than throughput-oriented multicore applications. As the Android and iPhone markets grow, so does the software base. This is also the case for multicore processors used for networking.

Because limited legacy software was available, these new domains are prompting the development of what's known as “virgin” code, particularly on multicore processors. Further, because of the thirst for low power

and higher performance, investing in new software in these domains is cost-effective. This is slowly eroding the dominance of x86 in these markets and prompting the development of binaries for other ISAs. Eventually, the rapid development of software for these new domains will slow down, but the software certainly won't be based solely on x86.

ISA DOMINANCE

The three ISAs contending for future dominance are ARM, MIPS, and PowerPC. All three are open RISC ISAs licensed by individual companies that charge a fee for their use and even offer processor cores as an IP block that you can drop into your design. This open ISA model has benefitted all three ISAs enormously.

ARM is the dominant ISA in most smartphones today. MIPS is trying to get into the smartphone business through its adoption for the Android operating system, but it's a bigger player in the world of multicore processors used for networking, where companies such as Cavium and NetLogic use the MIPS ISA in their cores.

PowerPC is the leading ISA in the networking market, followed by the x86 and MIPS. PowerPC's adoption in the networking market is mostly fueled by Freescale, but MIPS-based products from Cavium and NetLogic (soon to be acquired by Broadcom) are rapidly gaining a significant market share.

Fueling the growth of these ISAs are embedded software vendors such as Wind River (now owned by Intel), Monta Vista (an independent subsidiary of Cavium), ENEA, and Green Hills. Even traditional software vendors are starting to develop software for these ISAs. For example, Microsoft has announced the development of Windows for ARM.

THE INNOVATOR'S DILEMMA

So, now what? How will the overall processor market be carved up among

different ISAs? Will x86 move into a niche, such as datacenters, where it's gaining more rapid adoption?

This is the classic case of the innovator's dilemma. The emergence of these three ISAs and their rapid adoption in markets other than the laptop, desktop, and server markets can in the end provide the “destructive” innovation that can move the incumbent ISA into a niche.

A question to ponder is what three dominant ISAs should do to accelerate the destructive innovation of the incumbent and what the x86 should do to prevent being marginalized by these three ISAs. In addition to having the technological advantage of power efficiency, the incumbent ISAs have the first mover advantage in the new domains as well as the momentum of ongoing rapid software development.

X86 faces an even bigger challenge. Should it try to create a reduced x86 ISA, similar to ARM's Thumb and the microMIPS from MIPS? That would eliminate x86's advantage of binary compatibility. Or, should x86 open up the ISA to allow broader hardware development? This again would create direct competition among the vendors themselves.

It will be interesting to observe how these four ISAs carve out different portions of the market during the next decade. The question in many people's minds is whether x86 will continue its dominance, or will it be pushed aside by ARM, MIPS, or PowerPC? Or will we almost certainly kill off one or more of these ISAs?

Shubu Mukherjee is a Distinguished Engineer at Cavium Networks in Marlborough, Massachusetts. Contact him at shubu.mukherjee@caviumnetworks.com.

cn Selected CS articles and columns are available for free at <http://ComputingNow.computer.org>.

Super Computing Special Offer for You!

Apply now and get the CSDP Bundle for \$495 (regularly \$745)

Enter the following code for the discount: SC2011

www.computer.org/csdp



Distinguish Yourself From the Crowd Earn Your CSDP

Earning the Certified Software Development Professional (CSDP) credential is the best way to prove your abilities, skills, and knowledge.

By adding the CSDP credential to your resume, you will demonstrate you are:

- **Current** with best software practices
- **Connected** with industry's brightest minds
- **Career-minded** and ready for that next promotion
- **Committed** to advancing the software engineering profession



To read how the CSDP credential has helped employers and employees, go to:

www.computer.org/getcertified

SOCIAL COMPUTING

Crowdsourcing Maps

Mikhail Masli
University of Minnesota



Geowikis combine user-contributed content with map-based interactive Web applications.

As an avid cyclist living in Minneapolis, Minnesota, I regularly ride the bike trail that runs by the Hiawatha Light Rail Transit line near downtown. To reach this trail during rush hour, I use a parking lot to cut diagonally across a rectangular city block instead of using the bounding streets, which at that time are often bustling with motor traffic.

In the past, I often found myself unable to describe this route to friends because a parking lot isn't technically a street and hence isn't indicated on popular Web-based resources like Google Maps (even on their bicycling layer) and MapQuest. Fortunately, a new technology has emerged—the *geowiki*—that lets cyclists themselves mark and maintain such informal pathways. The concept of crowdsourcing maps can be applied to many other domains as well.

GEOWIKIS

Geowikis are at the confluence of two social computing trends: user-contributed content and map-based interactive Web applications. Geowikis offer

- simple, WYSIWYG editing of geographic features like roads and landmarks,

- versioning that works with a network of tightly coupled objects rather than independent documents, and
- spatial monitoring tools that make it easier for users to “watch” a geographic area for possibly malicious edits and interpret map changes visually.

One prominent geowiki is OpenStreetMap (www.openstreetmap.org), a project started at University College London in 2004 by Steve Coast that aims to build a free street map of the entire world from scratch with user-provider content. Volunteers gather this information in various ways—for example, by traveling with GPS devices that track their paths or by consulting aerial photography—or simply supply firsthand local knowledge. OSM data has been extended to niche domains like cycling (OpenCycleMap) and public transit (OpenBusMap).

Another geowiki that specifically addresses my needs is Cyclopath (<http://cyclopath.org>), a University of Minnesota research effort founded in 2008 by Reid Priedhorsky while a graduate student and Professor Loren Terveen that serves as a resource for cyclists in the Minneapolis-St. Paul metropolitan area. Cyclopath computes biking routes that “match

the way you ride” based on content provided by the cycling community and the requestor's personal “bikeability ratings.” Unlike OSM, Cyclopath allows anonymous access.

In the context of cycling, geowikis have two distinct advantages over current mapping solutions (R. Priedhorsky, B. Jordan, and L. Terveen, “How a Personalized Geowiki Can Help Bicyclists Share Information More Effectively,” *Proc. 2007 Int'l Symp. Wikis [WikiSym 07]*, ACM, 2007, pp. 93-98). First, commercial maps often don't include bike paths because cycling is less popular than driving and thus such paths are unattractive to advertisers. Second, cyclists require detailed and continually changing information to plan their routes—something that's hard for hobbyists as well as centrally managed mapping providers to sustain. Crowdsourcing changes this equation by distributing the task across numerous, motivated users.

Another application of geowikis is crisis management. During political disturbances, like those associated with Arab Spring, and in the wake of natural disasters, such as the tsunami that devastated Japan in March 2011, order often breaks down and obtaining up-to-date information can be difficult. By aggregating data from many geographically scattered individuals,

crowdsourcing enables officials and news monitors to keep abreast of such crises. The “Ushahidi: Crowdsourced Crisis Response” sidebar describes one map-based Web application that uses content provided by volunteers to help organize relief efforts and monitor political events around the world.

OPEN COLLABORATION

Geowikis typically support multiple contribution modes and, like Wikipedia and other traditional text-based counterparts, follow a revision-based paradigm. In Cyclopath, for example, users can add, modify, or delete road and trail segments; points of interest; parks, neighborhoods, and other bounded regions; and notes and tags about any of these map features.

Editing is simple and interactive. As Figure 1 shows, a user simply drags and drops objects on a Flash front end within the browser and then clicks “Save Changes.” This sends all edits to the server as a single revision to the map. All revisions are public for transparency. To help the user community moderate itself, the GUI includes a Recent Changes list and My Watch Regions—areas that individuals volunteer to monitor—to identify and undo malicious changes.

Since Cyclopath’s release, the user community has made more than 13,000 revisions—including the addition of my parking lot shortcut in downtown Minneapolis. Deviating slightly from the pure wiki model, users can also record some private information in the form of bikeability ratings, on a five-level scale (“impassable” to “excellent”), for each road or trail segment; these are private to the rater and used publicly only in the aggregate by Cyclopath’s route finder.

Unlike standard wikis, volunteer work in geowikis can have a significant offline social element. For example, the OSM community organizes “mapping parties” to create content for local geographic

USHAHIDI: CROWDSOURCED CRISIS RESPONSE

Adapting the Swahili word for testimony, Ushahidi (<http://ushahidi.com>) is a map-based Web application for crisis management that relies on user-contributed content. First proposed by Kenyan lawyer Ory Okolloh on her blog in the aftermath of the country’s violent 2008 elections as a means to collect eyewitness accounts, it also served as a platform for organizing rescue and rehabilitation efforts after the recent earthquakes in Japan, New Zealand, Chile, and Haiti and is being used to monitor ongoing conflicts in the Middle East and Africa.

Ushahidi crowdsources information through various means including e-mail, Twitter, the Short Message Service, and the Web. It then plots and updates this data in real time on an interactive map to make finding patterns easy. People can subscribe to the latest citizen-reported news via RSS. A separate, dedicated analytics platform called SwiftRiver can be used in tandem with Ushahidi to more effectively understand crowdsourced content.

Deployment of Ushahidi has faced several challenges such as the lack of reliable Internet connectivity, language barriers, cultural resistance to information sharing, and the difficulty of filing reports on the run in dangerous situations. Despite these challenges, the platform has been quite effective. In Kenya, for example, “it has changed how elections are monitored,” said Philip Thigo, an adviser with the Nairobi-based nongovernmental organization SODNET, which is using Ushahidi to gather reports on election violence across Africa. “It is working in real-time to impact elections as they take place, creating pressure on officials to act” (J. Wakefield, “Africa’s Quiet Digital Revolution,” 29 Sept. 2011, BBC News; www.bbc.co.uk/news/technology-14986314).

areas (M. Haklay and P. Weber, “OpenStreetMap: User-Generated Street Maps,” *Pervasive Computing*, Oct.-Dec. 2008, pp. 12-18). At one of the first such parties in May 2006 on the Isle of Wight, off the south coast of England, more than 30 volunteers spent two days traveling the island’s

streets on bikes and in cars tracking their paths using GPS. “A big aspect of getting OSM off the ground was the mapping parties: getting drunk and arguing with people,” Coast said. These local events are common in countries like India where OSM is gaining popularity.

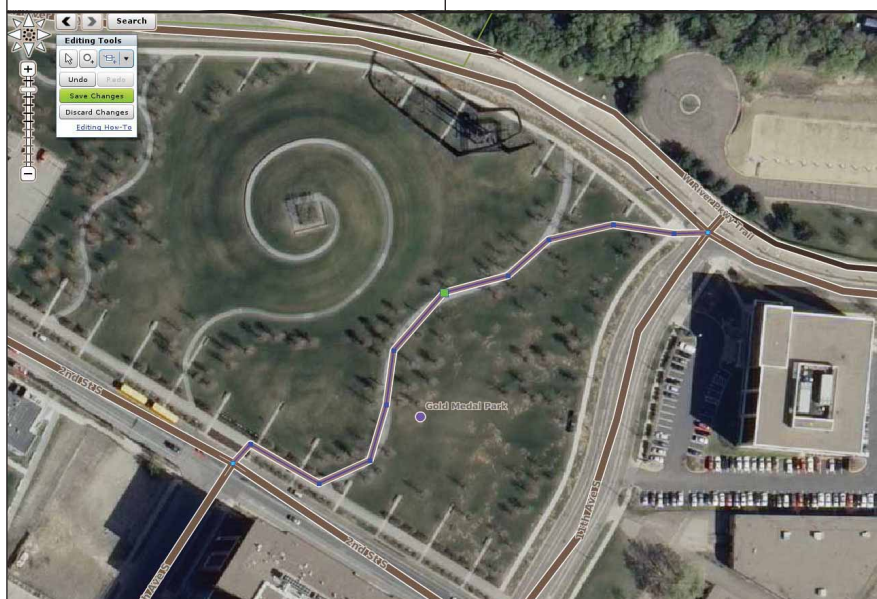


Figure 1. Editing in Cyclopath is a simple matter of dragging and dropping. In this case, a user has shaped a bike path to match the underlying aerial imagery.

SOCIAL COMPUTING

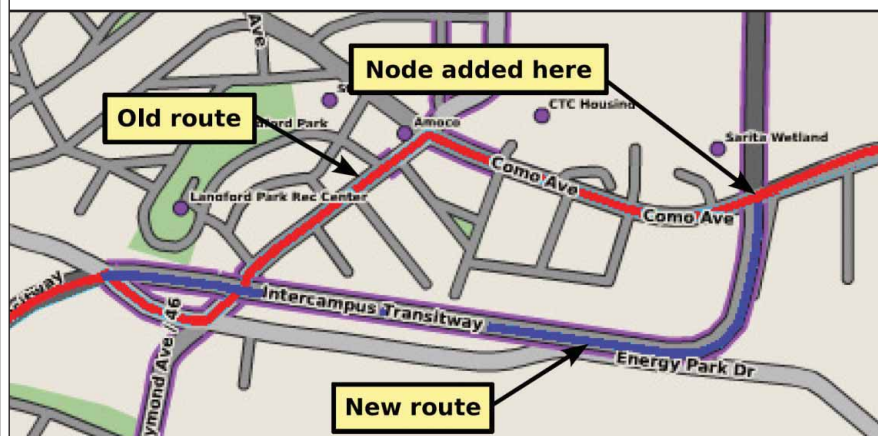


Figure 2. Example of how a user-contributed change improved the quality of a Cyclopath route. Connecting the bike trail to the road at the intersection of Como Avenue and Intercampus Transitway shortened the route from 15.6 km to 15 km.

INTERFACE CHALLENGES

Geowikis carry spatial as well as nonspatial data. Consequently, presenting this information in a visually clean manner can be difficult.

To meet cyclists' complex needs, the Cyclopath map interface is rich with information. Each segment of road between two intersections has a set of objective attributes such as number of lanes, lane width, width of shoulder, and whether it is one-way or two-way. It also has a bikeability rating: users logged in to the system see their own rating, if recorded, or a predicted rating based on other users' ratings and the road segment's properties. In addition, the segment could have several one- or two-word tags, such as "scenic" or "very bumpy," and descriptive notes.

When there's a lot of information, visualization of the map itself becomes a challenge. To mitigate this problem, several OSM sister geowikis present domain-specific views—for example, OpenCycleMap highlights cycling-specific attributes when rendering its map, while OpenBusMap highlights public transit stops and routes.

Even within a single domain, usage patterns can provide insights into simplifying the GUI.

Cyclopath's editing interface doesn't exercise any restrictions:

users can edit multiple roads, points of interest, or annotations within a single revision. Further, while OSM has separate viewing and editing modes, Cyclopath makes no such distinction. Nevertheless, most Cyclopath revisions are to a single feature—say, a road or point of interest—and those that do involve multiple features often follow a pattern such as roads and notes or roads and tags (M. Masli, R. Priedhorsky, and L. Terveen, "Task Specialization in Social Production Communities: The Case of Geographic Volunteer Work," *Proc. 5th Int'l AAAI Conf. Weblogs and Social Media*, [ICWSM 11], AAAI, 2011, pp. 217-224).

This observation suggests the usefulness of creating modal interfaces—for example, different modes for editing roads and points of interest.

GEOWIKI UTILITY

What benefits does a geowiki offer by virtue of it being a wiki? Two recent studies highlight some measurable benefits of open collaboration.

An analysis of 800 randomly chosen routes requested by Cyclopath users from August 2008 through April 2009 found that the nearly 8,500 revisions made during that period shortened the average route length by 1 km (R. Priedhorsky, M. Masli, and

L. Terveen, "Eliciting and Focusing Geographic Volunteer Work," *Proc. 2010 ACM Conf. Computer Supported Cooperative Work [CSCW 10]*, ACM, 2010, pp. 61-70). Figure 2 shows an example of a user-contributed change. Shortcuts via parking lots and connections of bike trails to nearby streets are examples of contributions responsible for this effect.

Similarly, a comparison of OSM's data with that of the Ordnance Survey, England's official national mapping agency, found that within the London area, the volunteer-gathered data was on average within 6 meters of the positions recorded by the agency and that about 80 percent of motorway objects between the two datasets overlapped (M. Haklay, "How Good Is Volunteered Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets," *Environment and Planning B: Planning and Design*, vol. 37, no. 4, 2010, pp. 682-703).

However, not all benefits are quantitative. A survey of 290 Cyclopath users found that 30 percent believed that their participation resulted in better navigation, and 26 percent felt that it increased knowledge about specific areas (K. Panciera, "User Lifecycles in Cyclopath: A Survey of Users," *Proc. 2011 iConf.*, ACM, 2011, pp. 741-741). "I imagine there are other cyclists who travel to the landmarks I've tagged, who use the routes I've marked," one contributor said. "I'd like to think I've made their rides just a little easier." Eighty-four percent of users indicated that Cyclopath improved the cycling experience by providing, for example, "tips on routes to take through neighborhoods I'm not familiar with ... what roads to avoid, what streets allow for faster commute, etc."

MOTIVATING GEOGRAPHIC VOLUNTEERS


A geowiki has little value without the contributions of its various users. Like an open collaboration system,

contribution to Cyclopath is highly skewed: a tiny fraction of users are responsible for a majority of revisions to the map (K. Panciera et al., "Lurking? Cyclopaths? A Quantitative Lifecycle Analysis of User Behavior in a Geowiki," *Proc. 28th Int'l Conf. Human Factors in Computing Systems [CHI 10]*, ACM, 2010, pp. 1917-1926).

The problem of motivating contributions from wiki volunteers has been widely researched in several domains. Simply highlighting areas on the map requiring user input, as Figure 3 shows, can be a powerful incentive in geowikis.

A field study revealed that Cyclopath users request and contribute significantly more work when given visual cues (R. Priedhorsky, M. Masli, and L. Terveen, "Eliciting and Focusing Geographic Volunteer Work"). The study also indicated that directing users' attention to areas familiar to them elicits more informative contributions for types of work that require local knowledge: merely studying the aerial imagery of an area is sufficient to confirm if two roads intersect, but to rate a road as "highly bikeable," a person must know the area well.

The proliferation of location-based services like Foursquare and Gowalla, coupled with the incorporation of location information in social networking sites like Facebook, Google+, and Twitter, point to a growing geographic dimension in online volunteer activity that has the potential to make the world even more interconnected.

Geowikis like OSM and Cyclopath, as well as other types of crowdsourced maps, can put considerable power in the hands of the communities they're designed to support. The benefits extend far beyond discovering cycling shortcuts. By drawing on the volunteer efforts of millions of mobile-phone-wielding citizens, this emerging technology could facilitate improvements around the world in areas ranging from emergency response to protecting human rights. 

Mikhail Masli is a PhD student in the Department of Computer Science and Engineering at the University of Minnesota. Contact him at masli@cs.umn.edu.

Editor: John Riedl, Department of Computer Science and Engineering, University of Minnesota; riedl@cs.umn.edu

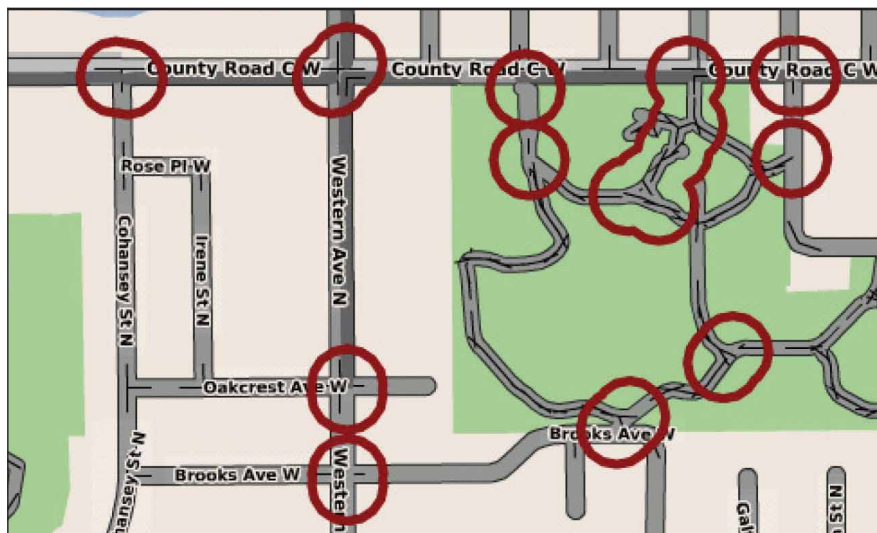


Figure 3. Visually highlighting areas on a map requiring human input can motivate geowiki volunteers. In this case, circling roads that appear to cross prompts some knowledgeable users to indicate whether these are intersections or bridges.

Nokia Siemens Networks US LLC (NSN)

has the following positions in

Irving, TX:

Carrier Ethernet Care Engineer

Review customers telecom networks; set up & maintain server & client network configuration; work with hiD switches, mobile switching, & interoperability between layer 3 switches & layer 2 microwave products; & other duties/skills required. CCIE certification required. Traveling required at least 30% (domestic & international). [Job ID: NSN-TXII-CECE]

Optical Care Engineer

Work with optical transmission technologies such as Dense Wavelength Division Multiplexing, Synchronous Digital Hierarchy, Synchronous Optical Networking, Network Management Systems & IP; & other duties/skills required. [Job ID: NSN-TXIIJ-OP]

Mail resume to:
NSN Recruiter
MS 4C-1-1580
6000 Connection Dr
Irving, TX 75039
& note specific Job ID#.

CAREER OPPORTUNITIES

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, Faculty Positions.

The Department of Electrical Engineering and Computer Science (EECS) seeks candidates for faculty positions starting in September 2012. Appointment will be at the assistant or untenured associate professor level. In special cases, a senior faculty appointment may be possible. Faculty duties include teaching at the graduate and undergraduate levels, research, and supervision of student research. We will consider candidates with backgrounds and interests in any area of electrical engineering and computer science. Faculty appointments will commence after completion of a doctoral degree. Candidates must register with the EECS search website at <https://eecs-search.eecs.mit.edu>, and must submit application materials electronically to this website. Candidate applications should include a description of professional interests and goals in both teaching and research. Each application should include a curriculum vita and the names and addresses of three or more individuals who will provide letters of recommendation. Letter writers should submit their letters directly to MIT, preferably on the website or by mailing to the address below. Please submit a complete application by December 15, 2011.

Send all materials not submitted on the website to: Professor Anantha Chandrakasan, Department Head, Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Room 38-401, 77 Massachusetts Avenue, Cambridge, MA 02139. M.I.T. is an equal opportunity/affirmative action employer.

PURDUE UNIVERSITY, Computer Science Department. The Department of Computer Science at Purdue University invites applications for a tenure-track position at the assistant professor level beginning August 2012. Outstanding candidates in all areas of Computer Science and with a multi-disciplinary focus are encouraged to apply. Specific needs that have been identified include theory and software engineering. The Department of Computer Science offers a stimulating and nurturing academic environment. Forty-eight faculty members direct research programs in analysis of algorithms, bioinformatics, databases, distributed and parallel computing, graphics and visualization, information security, machine learning, networking, programming languages and compilers, scientific computing, and software engineering. Information about the

department and a detailed description of the open position are available at <http://www.cs.purdue.edu>. All applicants should hold a PhD in Computer Science, or a closely related discipline, be committed to excellence in teaching, and have demonstrated potential for excellence in research. The successful candidate will be expected to teach courses in computer science, conduct research in field of expertise and participate in other department and university activities. Salary and benefits are highly competitive. Applicants are strongly encouraged to apply online at <https://hiring.science.purdue.edu>. Hard copy applications can be sent to: Faculty Search Chair, Department of Computer Science, 305 N. University Street, Purdue University, West Lafayette, IN 47907. Review of applications will begin on November 10, 2011, and will continue until the position is filled. A background check will be required for employment in this position. Purdue University is an Equal Opportunity/Equal Access/Affirmative Action employer fully committed to achieving a diverse workforce.

SENIOR SYSTEMS & SOFTWARE DESIGNER: M-F/8-5/40hr.wk. Design & develop next generation of Promptlink

Baylor University

Assistant or Associate Professor of Computer Science

Chartered in 1845 by the Republic of Texas, Baylor University is the oldest university in Texas and the world's largest Baptist University. Baylor's mission is to educate men and women for worldwide leadership and service by integrating academic excellence and Christian commitment within a caring community. Baylor is actively recruiting new faculty with a strong commitment to the classroom and an equally strong commitment to discovering new knowledge as Baylor aspires to become a top tier research university while reaffirming and strengthening its distinctive Christian mission as described in Baylor 2012 (www.baylor.edu/vision/). The combination of teaching, research and service has made Baylor one of the best universities for faculty, according to the Chronicle of Higher Education <http://chronicle.com/article/Great-Colleges-to-Work-For/128312/>.

The Department of Computer Science seeks a productive scholar and dedicated teacher for a tenure-track position beginning August, 2012. All specializations will be considered. Game/simulated environments, mobile computing, and graphics are of particular interest. The successful candidate will hold a terminal degree in Computer Science or a closely related field, demonstrate scholarly capability in his or her area of specialization, and exhibit a passion for teaching and mentoring at the graduate and undergraduate level. For position details and application information please visit: <http://www.ecs.baylor.edu>.

The Department: The Department offers a CSAB-accredited B.S. in Computer Science degree, a B.A. degree with a major in Computer Science, a B.S. in Informatics with a major in Bioinformatics, and a M.S. degree in Computer Science. We are currently seeking approval to offer a dual Ph.D. degree in cooperation with a well-established European institution. The Department has 15 full-time faculty, over 370 undergraduate majors and 30 master's students. The Department's greatest strength is the faculty's dedication to the success of the students and each other. Interested candidates may contact any faculty member to ask questions and/or visit the web site of the School of Engineering and Computer Science at <http://www.ecs.baylor.edu>.

The University: Baylor University, situated on a 500-acre campus next to the Brazos River. It annually enrolls more than 14,000 students in over 150 baccalaureate and 80 graduate programs through: the College of Arts and Sciences; the Schools of Business, Education, Engineering and Computer Science, Music, Nursing, Law, Social Work, and Graduate Studies; plus Truett Seminary and the Honors College. For more information see <http://www.baylor.edu>.

Application Procedure: Please submit a letter of application, current curriculum vitae, and transcripts. Include names, addresses, and phone numbers of three individuals from whom you have requested letters of recommendation to: Jeff Donahoo, Ph.D., Search Committee Chair, Baylor University, One Bear Place #97356, Waco, Texas 76798-7356, Materials may be submitted to: Jeff_Donahoo@baylor.edu

Appointment Date: Fall 2012. For full consideration, applications should be received by January 1, 2012. However, applications will be accepted until the position is filled.

Baylor is a Baptist university affiliated with the Baptist General Convention of Texas. As an Affirmative Action/Equal Employment Opportunity employer, Baylor encourages minorities, women, veterans, and persons with disabilities to apply.

Cable Network Monitoring, Cable Modem Provisioning & Cable Modems Test Platform; set-up R&D program for new products: IPTV in DOCSIS environment, DSL Modem Test Platform, provisioning system for TR-069/TR-104 devices; prepare development plan for current products to support Docsis 3.0 and IPv6; develop test procedures & support production rollout of new software applications; consult & coach development team, review & refactor new & existing code; provide Level 3 technical support; proficiency in Java EE, c/c++/c# programming & Linux, MySQL, PostgreSQL/Oracle, Camailio/SER/Asterisk administration, IPv4 & IPv6 networking, DOCSIS/Packetcable/xDSL/IPDR/TR-069/TR-104 specifications & SIP/MGCP/NCS/RTP protocols, signal processing, HFC network & metrology. Req.: M.S. in Information Science. Submit resume w/ad copy to: Foad Towfiq, Promptlink Communications, Inc., 4005 Avenida de la Plata, Oceanside, CA 92056.

THE UNIVERSITY OF COLORADO, COLORADO SPRING invites applications for tenure-track Assistant Professor positions in all areas of CS and Software Eng. The CS department offers Bachelor, Master and PhD degrees. See full job description and apply electronically at <http://www.JobsatCU.com>, refer to posting #815131. Review of applications will begin on January 15, 2012 and continue until the positions are filled.

SERVICES ARCHITECT - SECURITY (Islandia, NY, Locs throughout US). Architect, dsngn, engineer, & implmt integrated sec. solutions in client env. rldt to Identity & Access Mgmt, Info & Threat Mgmt. Config & support enterprise mgmt s/ware. Dsgn, scope, assess & deliver solutions. Confer w/clients & dev. team to plan sec. & s/ware modifications. Dsgn & implmt custom secu. policies. Provide project planning support. Review plans to ensure compatibility of planned sec. measures w/established guidelines & industry leading comp sec. sys. s/ware. Troubleshoot tech integrated solution implmtn issues for solution for sec. s/ware. Reqs: Bach's deg or for. equiv in CS, CIS, Math, Eng (any) or rel. fld + 5 yrs prog. exp in job offd &/or rel. pos. Employer will accept Master's deg or for. equiv in CS, CIS, Math, Eng (any) or rel. fld + 1 yr exp in job offd &/or rel. pos. Must have exp w/: architecting & implmtg sec. products as cross Business Unit multi-product integrated solution w/in client env.; dsngn, scoping, assessing & delivering solutions; troubleshooting tech integrated solutions implmtn issues for solutions for sec. s/ware; providing

FIU | FLORIDA INTERNATIONAL UNIVERSITY

FIU is a multi-campus public research university located in Miami, a vibrant, international city. FIU offers more than 180 baccalaureate, masters, professional and doctoral degree programs to over 42,000 students. As one of South Florida's anchor institutions, FIU is worlds ahead in its local and global engagement and is committed to finding solutions to the most challenging problems of our times.

The School of Computing and Information Sciences seeks exceptionally qualified candidates for multiple tenure-track and tenured faculty positions at all levels. Outstanding candidates are sought in areas of bio/medical/health informatics, computer architecture, computer graphics, large-scale data management, search, and visualization, human-computer interaction (HCI), networking, programming languages, robotics and game theory, and telecommunication. Exceptional candidates in other areas will be considered as well. Preference will be given to candidates who will enhance or complement our existing research.

Ideal candidates for junior positions should have a record of exceptional research in their early careers. Candidates for senior positions must have an active and proven record of excellence in funded research, publications, and professional service, as well as a demonstrated ability to develop and lead collaborative research projects. In addition to developing or expanding a high-quality research program, all successful applicants must be committed to excellence in teaching at both graduate and undergraduate levels. An earned Ph.D. in Computer Science or related disciplines is required.

Florida International University (FIU), the state university of Florida in Miami, is ranked by the Carnegie Foundation as a comprehensive doctoral research university with high research activity. The School of Computing and Information Sciences (SCIS) is a rapidly growing program of excellence at the University, with 31 faculty members and 1,400 students, including 65 Ph.D. students. SCIS offers B.S., M.S., and Ph.D. degrees in Computer Science, an M.S. degree in Telecommunications and Networking, and B.S., B.A., and M.S. degrees in Information Technology. SCIS has received approximately \$12.6M in the last three years in external research funding, has six research centers/clusters with first-class computing infrastructure and support, and enjoys broad and dynamic industry and international partnerships.

HOW TO APPLY:

Applications, including a letter of interest, contact information, curriculum vitae, and the names of at least three references, should be submitted directly to the FIU J.O.B.S Link website at <https://www.fiujobs.org>; refer to Position # 33334. The application review process will begin on January 16, 2012, and will continue until the position is filled. Further information can be obtained from the School website <http://www.cis.fiu.edu>, or by e-mail to recruit@cis.fiu.edu.

FIU is a member of the State University System of Florida and is an Equal Opportunity, Equal Access Affirmative Action Employer.

Positions at the Institute for Defense Analyses Center for Computing Sciences

The Institute for Defense Analyses Center for Computing Sciences (IDA/CCS) is looking for outstanding researchers to address difficult computing problems vital to the nation's security. IDA/CCS is an independent, applied research center sponsored by the National Security Agency (NSA). Emphasis areas for IDA/CCS technical staff include high-performance computing, cryptography, and network security. Members of the technical staff come from a diverse variety of backgrounds, including computer science, computer architecture, computer/electrical engineering, information processing, and the mathematical sciences; most have Ph.D.s. Special attention is paid to the design, prototyping, evaluation, and effective use of new computational algorithms, tools, paradigms, and hardware directly relevant to the NSA mission. Stable funding provides for a vibrant research environment, and an atmosphere of intellectual inquiry free of administrative burdens.

The center is equipped with a very large variety of hardware and software. The latest developments in high-end computing are heavily used and projects routinely challenge the capability of the most advanced algorithms and architectures. IDA/CCS research staff members have always been at the forefront of computing, as evidenced by lasting, visible contributions to areas as varied as multi-threaded architectures (e.g., Horizon), novel computing systems (e.g., FPGA-based Splash and Splash-2, Processing-In-Memory chips), design and implementation of operating systems (e.g., the Linux kernel), and programming language design and implementation for high-performance computing systems (e.g., Universal Parallel C and Cinquecento).

IDA/CCS research staff work on complex topics often engaging multidisciplinary teams; candidates should demonstrate depth in a particular field as well as a broad understanding of computational issues and technology. Because the problems of interest are continually evolving, IDA/CCS recruitment focuses on self-motivation, strength of background, and talent, rather than specific expertise.

Located in a modern research park in the Maryland suburbs of Washington, DC, IDA/CCS offers a competitive salary, an excellent benefits package, and a superior professional working environment. U.S. citizenship and a Department of Defense TSSI clearance (with polygraph) are required. IDA/CCS will sponsor this clearance for those selected. The Institute for Defense Analyses is proud to be an equal opportunity employer.

Please send responses or inquiries to:

Dawn Porter
Administrative Manager
IDA Center for Computing Sciences
17100 Science Drive
Bowie, MD 20715-4300
dawn@super.org



THE HONG KONG UNIVERSITY OF
SCIENCE AND TECHNOLOGY

Department of Computer Science and Engineering Faculty Positions

The Department of Computer Science and Engineering, HKUST (<http://www.cse.ust.hk/>) has more than 40 faculty members, recruited from major universities and research institutions around the world, and about 800 students (including about 200 postgraduate students). The medium of instruction is English. In 2011, we were ranked 21st among all Computer Science Departments worldwide according to Academic Ranking of World Universities, and 26th according to QS World University Ranking.

The Department will have at least two tenure-track faculty openings at Assistant Professor/Associate Professor/Professor levels for the 2012-2013 academic year. We are looking for faculty candidates with interests in bioinformatics, security, or cloud computing. Strong candidates in core computer science and engineering research areas will also be considered. Applicants at Assistant Professor level should have an earned PhD degree and demonstrated potential in teaching and research.

Salary is highly competitive and will be commensurate with qualifications and experience. Fringe benefits include medical/dental benefits and annual leave. Housing will also be provided where applicable. For appointment at Assistant Professor/Associate Professor level, initial appointment will normally be on a three-year contract. A gratuity will be payable upon completion of contract.

Applications should be sent through e-mail including a cover letter, curriculum vitae (including the names and contact information of at least three references), a research statement and a teaching statement (all in PDF format) to cse recruit@cse.ust.hk. Priority will be given to applications received by **29 February 2012**. Applicants will be promptly acknowledged through e-mail upon receiving the electronic application material.

(Information provided by applicants will be used for recruitment and other employment-related purposes.)

project planning support; configuring & supporting enterprise mgmt s/ware; dsngng & implmtg custom sec. policies. Travel throughout US reqd. Work from home benefit available. Send resume to: Althea Wilson, CA Technologies, One CA Plaza, Islandia, NY 11749, Requisition # 24857.

OLD DOMINION UNIVERSITY, Department of Modeling, Simulation and Visualization Engineering, Batten College of Engineering and Technology. Position Advertisement, Fall 2011.

Position FO639A: MSVE – Network/Cyber Security. Modeling and Simulation Faculty Position. The Department of Modeling, Simulation and Visualization Engineering at Old Dominion University's Batten College of Engineering and Technology invites applications for a tenure-track faculty position beginning July 25, 2012. The successful applicant will have research experience in core modeling and simulation (M&S) areas and a commitment to providing quality teaching in the department's bachelors, masters, and doctoral programs. Duties include undergraduate and graduate teaching and development of a strong, externally-funded research program. This is an opportunity to join and help shape the first M&S department. Preference will be given to applicants having experience in performing interdisciplinary research in the application of M&S for computer networking and cyber security. Of particular interest are candidates whose research spans one or more of the following areas: mobile networks, web services, network security, and information security. ODU is critically situated near major DoD organizations and other federal agencies having interest in these research areas. The department also works in partnership with the Virginia Modeling, Analysis and Simulation Center (VMASC), ODU's internationally recognized M&S research center. Applicants must have an earned Ph.D. in an engineering or science discipline closely related to M&S. Preference will be given to candidates who possess or are eligible to obtain a security clearance. Applications should include a cover letter, complete resume, statement of teaching and research interests, statement concerning security clearance eligibility, and three letters of reference. All application materials must be submitted via email as a single pdf document to Dr. Roland Mielke, Chair, MSVE Department, at the following email address: dachterf@odu.edu. Review of applications will begin December 1, 2011 and will continue until the position is filled. Old Dominion University is an equal opportunity, affirmative action institution and requires com-

pliance with the Immigration Reform and Control Act of 1986.

TSINGHUA UNIVERSITY, Institute for Interdisciplinary Information Sciences (IIIS), Assistant/Associate/Full Professor. IIIS (<http://iiis.tsinghua.edu.cn/en/>) invites applications from highly-qualified candidates in areas including (but not limited to) Computer Systems, Algorithms and Complexity, Machine Learning, Multimedia, Databases, Computer Networks, Wireless Sensor Networks, Information Security, Web Technologies, Energy-Efficient Computing, Computational Finance, Quantum Information, Computational Biology. Please send applications or nominations in the form of an application letter enclosing a current CV to iiisdean@mail.tsinghua.edu.cn.

COLORADO STATE UNIVERSITY, Assistant Professor, Department of Computer Science. Colorado State University has an opening for one or more tenure-track assistant professor positions in Computer Science, beginning fall 2012. Areas of interest include com-

putational biology, parallel computing, programming languages that focus on parallel programming models, mobile computing and HCI. For more information go to: <http://www.cs.colostate.edu>. Applications must be received by January 9, 2012 at <http://www.natsci.colostate.edu/employment/compsci/>. Complete applications of semi-finalists will be available to department faculty for review. CSU is an EO/EA/AA employer. Colorado State University conducts background checks on all final candidates.

SR. SERVICES CONSULTANT (Islandia, NY/Locs throughout US). Architect, dsgn, implmt & administer CA n/work security & Identity Mgmt Solutions (SSO); Implmt secure Enterprise SSO envts Troubleshoot implmnt. Travel to mult locs throughout US reqd. Work from home benefit available. Reqs: Bach's deg or for. equiv in CS, CIS, Math, Eng (any), or rel. tech + 5 yrs prog. exp in job &/or rel. Employer will also accept Master's deg or for. equiv in CS, Math, Eng (any) or rel. tech + 1 yr exp in job offd &/or rel. Must have exp w/ implmtg secure Enterprise Single Sign-On envts using CA SiteMind-

University of South Florida Assistant Professor Positions Instructor Position Computer Science and Engineering

Applications are invited for two **tenure-track Assistant Professor positions** and one **Instructor position** in the Department of Computer Science and Engineering. For the Assistant Professor positions the Department is hiring in all areas of computer engineering and computer security (cybersecurity). We seek candidates with a record of outstanding-quality research publications and potential for excellence in teaching. For the instructor position we seek an individual who can teach a broad range of core computer science and computer engineering courses – both software and hardware – at the undergraduate level, as well as advise students.

The Department of Computer Science and Engineering (<http://www.cse.usf.edu>) has 23 faculty members and offers B.S., M.S., and Ph.D. degrees. The research program is well supported by federal and state agencies and industry. The University of South Florida serves over 47,000 students and is one of the nation's top public research universities.

For further information and for application instructions, please see our faculty search website: <http://www.cse.usf.edu/faculty-search/>. Direct any questions to faculty-search@cse.usf.edu. Applications will be considered starting immediately until the positions are filled.

According to Florida law, applications and meetings regarding them are open to the public. The University of South Florida is an Equal Opportunity/Equal Access/Affirmative Action Institution. Women and minorities are strongly encouraged to apply.



TEMPE UNIVERSITY Non-tenure-track, Open Rank, Full-time Faculty Position Department of Computer and Information Sciences

Applications are invited for a non-tenure-track, open rank, full-time faculty position in the Department of Computer and Information Sciences at Temple University. Areas of interest include, but are not limited to, information security and assurance, enterprise computing, computer networking, databases, client-server computing, e-commerce systems, project management, and applications software development. Candidates are expected to teach advanced undergraduate and graduate masters level courses in both CS and IS&T programs.

Applications should include curriculum vitae and a statement of recent achievements and teaching goals. Candidates should also have three reference letters sent directly. Please submit applications online at <http://academicjobsonline.org>.

Review of candidates will begin on November 1, 2011 and will continue until the position is filled. For further information check <http://www.cis.temple.edu> or send email to search committee chair Dr. Justin Shi at shi@temple.edu. Temple University is an equal opportunity, equal access, affirmative action employer.

Baylor University

Assistant, Associate or Full Professor of Computer Science

The Department of Computer Science seeks a productive scholar and dedicated teacher for a tenure-track position beginning August, 2012. The ideal candidate will hold a terminal degree in Computer Science or closely related field, demonstrate scholarly capability and an established and active independent research agenda in one of several core areas of interest, including, but not limited to, game design and development, software engineering, computational biology, machine learning or large-scale data mining. A successful candidate will also exhibit a passion for teaching and mentoring at the graduate and undergraduate level. For position details and application information please visit: <http://www.baylor.edu/hr/index.php?id=81302>

The Department: The Department offers a CSAB-accredited B.S. in Computer Science degree, a B.A. degree with a major in Computer Science, a B.S. in Informatics with a major in Bioinformatics, and a M.S. degree in Computer Science. The Department has 13 full-time faculty members, over 250 undergraduate majors and approximately 30 master's students. We are currently seeking approval to offer a dual Ph.D. degree in cooperation with a well-established partner institution. Interested candidates may contact any faculty member to ask questions and/or visit the web site of the School of Engineering and Computer Science at <http://www.ecs.baylor.edu>.

The University: Chartered in 1845 by the Republic of Texas, Baylor University is the oldest university in Texas and the world's largest Baptist University. It is situated on a 500-acre campus next to the Brazos River and annually enrolls more than 14,000 students in over 150 baccalaureate and 80 graduate programs. Baylor's mission is to educate men and women for worldwide leadership and service by integrating academic excellence and Christian commitment within a caring community. Baylor is actively recruiting new faculty with a strong commitment to the classroom and an equally strong commitment to discovering new knowledge as Baylor aspires to become a top tier research university while reaffirming and strengthening its distinctive Christian mission as described in Baylor 2012 (www.baylor.edu/vision/).

Application Procedure: Applications, including detailed curriculum vitae, a statement demonstrating an active Christian faith, and contact information for three references should be sent to: Chair Search Committee, Department of Computer Science, Baylor University, One Bear Place #97356, Waco, TX 76798-7356.

Appointment Date: Fall 2012. For full consideration, applications should be received by January 1, 2012.

Baylor is a Baptist university affiliated with the Baptist General Convention of Texas. As an Affirmative Action/Equal Employment Opportunity employer, Baylor encourages minorities, women, veterans, and persons with disabilities to apply.

University of Maryland, College Park PROFESSOR and DIRECTOR Center for Bioinformatics and Computational Biology

The University of Maryland invites applications for Director of the Center for Bioinformatics and Computational Biology. Candidates are expected to be prominent scholars with publications and research experience at the interface of biological science and computing. Their primary responsibility will be to lead a nationally visible research program complementing existing strengths in computational genomics, proteomics, and molecular evolution. They will also be expected to promote the CBCB, and help build collaborative relationships, both on and off-campus. Information about the Center can be found at www.cbcb.umd.edu. Collectively, the CBCB faculty spans the fields of computer science, mathematics and statistics, biology, and biochemistry. The Center is housed in contiguous space and has access to significant high-end computing infrastructure through the University of Maryland Institute for Advanced Computer Studies. CBCB faculty members are also affiliated with at least one other campus academic unit appropriate to their interests. There is ample potential for collaboration with other organizations in the area, such as the NIH, the JCVI, and the Smithsonian Institution. For more information contact the search chair, Thomas D. Kocher (tdk@umd.edu). To apply, send a letter of application, curriculum vitae, and names of three references, following the instructions at <https://jobs.umd.edu> – Faculty Position Number 117572 AND <http://cbcb.umd.edu/hiring/>. Candidates must apply to both websites to receive consideration. Review of applications will begin November 15, 2011.

The University of Maryland is an affirmative action, equal opportunity employer. Women and minorities are encouraged to apply.

er, IdentityMinder, Identity Manager, ETrust Admin & federation Manager; Architecting, dsgng, implmtg & administering Network Security & Identity Mgmt solutions (SSO) incl user repositories, web access mgmt, provisioning & role-based access controls (RBAC); Implmtg Federated identity tech using SAML simple SSC & Liberty ID-FF; providing SiteMinder Access manager, Identity Manager & federation architecture consulting services, dsgng, implmtg & deploying integrated solutions that rely on mult. security technologies & OS incl Windows, UNIX, Linux, Sun Solaris Unix. Must have Cisco Certified Network Associate, MS Certified Systems Engineer, MS Certified Database Administrator & MS Certified Professional (MCP) certs. Send resume to: Althea Wilson, CA Technologies, One CA Plaza, Islandia, NY 11749, Requisition # 24865.

SENIOR SOFTWARE ENGINEER. Burlington, MA Develop the 3D visualization software and software user interface to map, in real time, the electrical activity of the heart for patients with cardiac arrhythmias undergoing a cath-

The University of North Texas Department of Computer Science and Engineering Assistant/Associate/Full Professors

The Department of Computer Science and Engineering at the University of North Texas (UNT) is seeking candidates for multiple tenure-track/tenured faculty positions at the Assistant, Associate or Full Professor level beginning August 15, 2012. The department plans to build on its existing strength in 3 areas: **Computer Systems**, including operating systems, runtime systems for cloud and high performance or mobile and handheld devices, software engineering of net-centric, real-time and embedded systems, and energy efficient and low power circuits and systems; **Intelligent Systems**, including data mining, machine learning, information retrieval, scientific visualization, human-computer interaction, and computational life sciences; and **Security**, including information assurance, network security and intrusion detection, and secure software systems and vulnerability analysis. Candidates should have demonstrated the potential to excel in research in one or more of these areas and in teaching. A Ph.D. in Computer Science, Computer Engineering or closely related field is required at the time of appointment. At the Assistant Professor level, the applicant's record must include high quality publications. At the Associate Professor level, the applicant must have at least 5 years of experience beyond an earned doctoral degree with a significant record of publications and extramural funding. A Full Professor would be expected to be a leader in his/her field with a record of building and maintaining a large-scale research program of international renown.

The Computer Science and Engineering department is home to 730 Bachelors students, 136 masters students and 77 Ph.D. students. Additional information about the department is available at the department's website: www.cse.unt.edu.

Application Procedure:

All applicants must apply online to: <https://facultyjobs.unt.edu/applicants/Central?quickFind=51533>. Submit nominations and questions regarding the position to Dr. Philip Sweany (sweany@cse.unt.edu).

Application Deadline:

The committee will begin its review of applications on December 1, 2011 and continue to accept and review applications until the positions are closed.

The University:

With about **36,000** students, UNT is the nation's **33rd** largest university. As the largest, most comprehensive university in Dallas-Ft. Worth, UNT drives the North Texas region. UNT offers **97** bachelor, **88** master's and **40** doctoral degree programs, many nationally and internationally recognized. A student-focused public research university, UNT is the flagship of the UNT System.

The University of North Texas is an AA/ADA/EOE committed to diversity in its educational programs.

eter procedure. Help direct and define software system testing and validation procedures to ensure specifications are met. Analyze the needs of physicians and patients with cardiac arrhythmia as well as software requirements to determine feasibility of design within time and cost restraints. Develop signal processing algorithms. Develop software in Linux and Windows based environments. Utilize C++ to develop and code 3D applications. Utilize Qt library for user interface design. A masters degree in information systems engineering, engineering, computer science, or related field is required. 1 year of experience in a commercial environment is required. 1 year of experience must include some experience with developing software in a Linux and Windows environment, C++ to develop and code 3D visualization applications, developing signal processing algorithms and utilizing Qt library. 1 year of experience as a software engineer or senior software engineer is acceptable. Send resume to HR, Rhythmia Medical, Inc., SW_job@rhythmia.com.

UNIVERSITY OF MINNESOTA - TWIN CITIES, Department of Computer Science and Engineering, Faculty Position.

The Department of Computer Science and Engineering at the University of Minnesota-Twin Cities invites applications from candidates in all areas of Computer Science for two faculty positions. We strongly encourage applications from women and members of minority groups. Candidates should have a PhD in Computer Science or a closely related discipline. The position is open until filled, but for full consideration apply at www.cs.umn.edu/employment/faculty by December 15, 2011. The University of Minnesota is an equal opportunity employer and educator.

SENIOR SERVICES CONSULTANT, (Islandia, NY & locs throughout US). Architect & implmt service mgmt products incl CA Service Desk Manager, CMDB, IT Process Automation Manager at individual product & solution level in client envrmt. Reqs: Bach's deg or for. equiv in CS, CIS, Math, Eng (any) or rel. tech + 5 yr prog. exp in job offd &/or rldt pos. Must have exp w/: Architecting & implmtg products at individual product & solution level w/in client envrmt; dsgng, scoping, assessing & delivering solutions; troubleshooting tech integrated solutions implmtn issues incl app performance; providing project planning support; workflow dvlpmt. Service Management processes; & IT implmtn consulting. Travel throughout US reqd. Work from home benefit available. Send

University of Illinois at Urbana-Champaign

The Department of Electrical and Computer Engineering invites applications for faculty positions at all levels and in all areas of electrical and computer engineering, particularly in the areas of control and communications, circuits, energy and power systems, nanoelectronics, nanophotonics, and computing. Applications are encouraged from candidates whose research programs are in traditional as well as nontraditional and interdisciplinary areas of electrical and computer engineering. The department is engaged in exciting new and expanding programs for research, education, and professional development, with strong ties to industry.

Applicants for positions at the assistant professor level must have an earned Ph.D. or equivalent degree, excellent academic credentials, and an outstanding ability to teach effectively at both the graduate and undergraduate levels. Successful candidates will be expected to initiate and carry out independent research and to perform academic duties associated with our B.S., M.S., and Ph.D. programs. Senior level appointments with tenure are available for persons of international stature.

Faculty in the department carry out research in a broad spectrum of areas and are supported by world-class facilities and programs for international work, including the Coordinated Science Laboratory, the Information Trust Institute, the Micro and Nanotechnology Laboratory, the Beckman Institute for Advanced Science and Technology, and several industrial centers. The department has one of the leading programs in the United States, granting approximately 350 B.S. degrees, 100 M.S. degrees, and 60 Ph.D. degrees annually.

In order to ensure full consideration by the Search Committee, applications must be received by December 15, 2011. Salary will be commensurate with qualifications. Preferred starting date is August 16, 2012, but is negotiable. Applications can be submitted by going to <http://jobs.illinois.edu> and uploading a cover letter, CV, research statement, and teaching statement, along with names of three references. For inquiry, please call 217-333-2301 or email ece-recruiting@illinois.edu.

Illinois is an Affirmative Action /Equal Opportunity Employer and welcomes individuals with diverse backgrounds, experiences, and ideas who embrace and value diversity and inclusivity (www.inclusiveillinois.illinois.edu).



UNIVERSITY OF MARYLAND

A. JAMES CLARK SCHOOL OF ENGINEERING DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING TENURE-TRACK AND TENURED FACULTY POSITIONS

The Department of Electrical and Computer Engineering at the University of Maryland seeks exceptionally qualified candidates for tenure-track and tenured faculty positions to begin in August 2012 in the area of: **Cyber Security**. This includes all aspects of security-oriented research in computer engineering, communications, signal processing, and networking, at the hardware, software, protocol, algorithm, system, and physical layer levels.

In addition, exceptional candidates will be considered in all research areas of interest to the department. Joint appointments are possible with other departments and with affiliated University of Maryland institutes.

Appointments at all ranks will be considered. Applicants should have received or expect to receive their PhD in Electrical Engineering or a related discipline prior to August 2012. Candidates for the rank of Assistant Professor should be creative and adaptable, and should have a high potential for both teaching and research. Candidates for the ranks of Associate and Full Professor should have distinguished records in research and a strong interest in educational programs.

For best consideration, applications should be submitted by **January 6, 2012** to <https://jobs.umd.edu> (position number 105704). Applications should include a cover letter, curriculum vitae with list of publications, research and teaching statements, and the names and contact information of at least four references.

The University of Maryland is an equal opportunity, affirmative action employer with a strong commitment to the principle of diversity. Applications from minority groups and women are especially invited.



i n v e n t

Hewlett-Packard Company is accepting resumes for the following positions:

IT Developer/Engineer

Houston, TX. (Ref. #HOUITDE31) and Austin, TX (Ref. # AUSITDE61)

Research, design, develop, configure, integrate, test and maintain existing and new business applications and/or information systems solutions, including databases through the integration of technical and business requirements.

Solutions Architect

Austin, TX. (Ref. #AUSARO1)

Work with peers to define, design, develop, and maintain a comprehensive end-to-end compliance architecture and roadmap in the Quote-to-Cash area.

Technology Consultant

Palo Alto, CA. (Ref. #RPALTC11)

Responsible for delivery of assigned tasks within the delivery cycle of a project. Maintain knowledge of a broad spectrum of HP technology in order to deliver part of a detailed technical design to meet customer requirements. Extensive travel required to various unanticipated locations throughout the U.S.

Quality Engineer

San Mateo, CA. (Ref. #SMQE11)

Design and develop an automation framework for a multi-tiered application. Develop Test Cases (both automated and manual) for the different components.

Electrical/Hardware Engineer

Corvallis, OR (Ref. #COREHW11)

Design, develop, modify and evaluate electronic parts, components or integrated circuitry for electronic equipment.

Mail resume to Hewlett-Packard Company, 5400 Legacy Drive, MS H1-6F-61, Plano, TX 75024. Resume must include Ref. #, full name, email address & mailing address. No phone calls please. Must be legally authorized to work in the U.S. without sponsorship. EOE.



is seeking

Applications Programmer, Principal

San Jose, CA • Reference Job Code: ENG8-SVCASG

Reqs MS in CS or related & 8 yrs exp. Reqs Microsoft .Net; ASP.Net, C#, VB.Net; Web Services and Windows Communication Foundation (WCF); Microsoft SQL Server programming; Microsoft Visual Studio; and object oriented design patterns.

Manager – Applications Engineering

San Jose, CA • Reference Job Code: ENG8-SVCAGA

Reqs BS in EE or related & 6 yrs exp. Reqs Traffic management and fabric architectures; Schematics and layout reviews; Software programming in reviewing and debugging software; Networking protocols.

Mail resumes to:

HR Operations Coordinator
5300 California Ave.
Bldg 2, #22108B
Irvine, CA 92617



is seeking a

Engineer, Sr Staff-IC Design

Sunnyvale, CA

Req. MS (or foreign equiv.) in Electrical Engg, CS, or rel and 2 yrs exp. Responsible for defining and implementing security architecture & roadmap for mobile devices, incl broadband chipsets & applications processors. Up to 5% domestic travel req. F/T. Must have unrestricted U.S. work authorization.

Mail resumes to:

HR Operations Coordinator
5300 California Ave.
Bldg 2, #22108B
Irvine, CA 92617
Must reference
job code ENG7-SVCAEM.

resume to: Althea Wilson, CA Technologies, One CA Plaza, Islandia, NY 11749, Requisition # 24866.

UNIVERSITY OF PENNSYLVANIA, Research Assistant Professor Position.

The University of Pennsylvania's Department of Computer and Information Science invites applicants for Research Assistant Professor. The department seeks an individual with exceptional promise for, or a proven record of, excellence in research in compositional theory for real-time systems, multimode systems, and real-time cloud computing systems. Applicants should hold a Ph.D. degree in Computer Science or Computer Engineering, and have a strong interest in applying research results to automotive systems and medical systems. The position is to be filled as soon as an appropriate candidate is identified. Research assistant professor position is for four years, renewable for additional years, contingent upon availability of research funding. Successful applicant will find Penn to be a stimulating environment conducive to professional growth in research. Please go to www.cis.upenn.edu/facultypositions to apply. The University of Pennsylvania is an Equal Opportunity/Affirmative Action Employer. The Penn CIS Faculty is sensitive to "dual career situations" and would be pleased to assist with opportunities in the Philadelphia region.

NOKIA SIEMENS NETWORKS US LLC

has the following position in Herndon, VA. Solution Consultant: Provide in-depth knowledge of UMTS/GSM/GPRS/EDGE wireless technologies including indoor/outdoor DAS systems, & their application in Service Optimization & Assurance; build & develop trust-based relationships with customers' mgmt. in relation to consulting engagements; & other duties/skills required. [Job ID: NSN2-VA11-SC]. Mail resume to Attn: NSN Recruiter, MS: 4C-1-1580, 6000 Connection Dr., Irving, TX 75039 & note Job ID#.

SOFTWARE ENGINEER (several pos.)

Dsgn & dvlp secure, reliable, scalable & performance-driven integration solutions utilizing knowl of & exp w/BizTalk Server 2009, ESQ, SOA Dsgn, WCT & Web Services, .Net Framework 3.5, WCF LOB Adapters, Web Service S/ware Factory, SQL Server 2008, Oracle 10g, MOSS 2007, MSMQ, WPF, WF, C#.Net, ASP .Net, LINQ, XML, XPATH, Code Generation. Freq travel reqd; Req MS in Comp Sci or Eng dis or rel or BS in Comp Sci or Eng dis or rel (or Foreign 4 yr Bach deg equated by USA evaluation service to be equiv to USA

Bach deg) + 5 yrs progressive wrk exp. Resumes to Infologitech Inc., 50 Cragwood Rd., Ste. 209, South Plainfield, NJ 07080.

UNIVERSITY OF WASHINGTON, Computer Science & Engineering Tenure-Track, Research, and Teaching Faculty.

The University of Washington's Department of Computer Science & Engineering has one or more open positions in a wide variety of technical areas in both Computer Science and Computer Engineering, and at all professional levels. A moderate teaching load allows time for quality research and close involvement with students. Our space in the Paul G. Allen Center for Computer Science & Engineering provides opportunities for new projects and initiatives. The Seattle area is particularly attractive given the presence of significant industrial research laboratories as well as a vibrant technology-driven entrepreneurial community that further enhances the intellectual atmosphere. Information about the department can be found on the web at <http://www.cs.washington.edu>. We welcome applicants in all research areas in Computer Science and Computer Engineering including both core and inter-disciplinary areas. We expect candidates to have a strong commitment both to research and to teaching. The department is primarily seeking individuals at the tenure-track Assistant Professor rank; however, under unusual circumstances and commensurate with the qualifications of the individual, appointments may be made at the rank of Associate Professor or Professor. We may also be seeking non-tenured research faculty at Assistant, Associate and Professor levels, postdoctoral researchers (Research Associates) and part-time and full-time annual lecturers and Sr. Lecturers. Applicants for both tenure-track and research positions must have earned a doctorate by the date of appointment; those applying for lecturer positions must have earned at least a Master's degree or have relevant teaching experience in the course area. Research Associates, Lecturers and Sr. Lecturers will be hired on an annual or multi-annual appointment. All University of Washington faculty engage in teaching, research and service. Please apply online at <https://norfolk.cs.washington.edu/apply> with a letter of application, a complete curriculum vitae, statement of research and teaching interests, and the names of four references. Applications received by December 1, 2011 will be given priority consideration. Open positions are contingent on funding. The University of Washington was awarded an Alfred P. Sloan Award for Faculty Career Flexibil-

ity in 2006. In addition, the University of Washington is a recipient of a National Science Foundation ADVANCE Institutional Transformation Award to increase the participation of women in academic science and engineering careers. We are building a culturally diverse faculty and encourage applications from women and minority candidates. The University of Washington is an affirmative action, equal opportunity employer.

IOWA STATE UNIVERSITY.

The Software Engineering Program at Iowa State University has an immediate opening for a tenure-track or tenured faculty position that will commence in August 2012. Appointments will be considered at all experience levels. Duties for the position will include undergraduate and graduate education; mentoring and engaging undergraduate as well as prospective students; developing and sustaining externally-funded research; graduate student supervision and mentoring; and professional and institutional service. An earned Ph.D. or equivalent in software engineering, computer science, computer engineering or a closely related field is required. For appointment at the level of assistant professor, the successful candidate must have demonstrated potential to establish and maintain a productive externally funded research program and potential to excel in the classroom. Commensurate experience and a proven track record will be expected for appointment at a more senior level. The tenure home in either the Department of Computer Science or the Department of Electrical and Computer Engineering will be decided in consultation with the successful candidate, with joint appointment in both departments. For more information and to apply: <http://www.se.iastate.edu/careers>. Candidates are subject to a background check. ISU is an EO/AA employer.

SOFTWARE ENGINEER (Several positions)

- Dsgn & dvlp secure, reliable, scalable & performance-driven integration solutions utilizing knowl of & exp w/BizTalk Server 2009, SOA Design, WCF & Web Services, .Net Framework 3.5, WCF LOB Adapters, SQL Server 2008, C#.Net, ASP .Net, XML, XPATH, Code Generation. Freq travel reqd; Req MS in Sci, Eng or rel. Mail resumes to Infologitech Inc., 50 Cragwood Rd., Ste. 209, South Plainfield, NJ 07080.

SENIOR TECHNICAL CONSULTANTS

sought by GSPANN Technologies, Inc. in Milpitas, CA. Responsible for providing prof'l comp consulting services in the



i n v e n t

HP Enterprise Services, LLC is accepting resumes for the following positions:

Services Information Developer

Frankfort, KY (Ref. #ESFRASID11), Monona, WI (Ref. #ESMOSID2), Pontiac, MI (Ref. #ESPONSID11),
Springfield, OR (Ref. #ESSPRSID21), and Topeka, KS (Ref. #ESTOSID11)

Conceptualize, design, develop, unit-test, configure, or implement portions of new or enhanced (upgrades or conversions) business and technical software solutions through application of appropriate standard software development life cycle methodologies and processes.

Information Testing

Springfield, OR (Ref. #ESSPRIT21) and Plano, TX. (Ref. #ESPLAYDO1)

Design, develop and execute all testing-related activities on applications, infrastructure or hardware components of IT solutions, which include both third party software and internally developed applications and infrastructure.

IT Business Consultant

Herndon, VA (Ref. #ESHEAGO1)

Provide consulting to businesses, functions and geographies that utilize IT services and drive effective business engagements for IT.

Technology Consultant

Washington, DC (Ref. #ESDCTC21) and Arlington, TX (Ref. #ESARLSPA1)

Provide technology consulting to customers and internal project teams. Provide technical support and/or leadership in creation and delivery of technology solutions designed to meet customers' business needs and, consequently, for understanding customers' businesses.

Mail resume to HP Enterprise Services, LLC, 5400 Legacy Drive, MS H1-6F-61, Plano, TX 75024. Resume must include Ref. #, full name, email address & mailing address. No phone calls please. Must be legally authorized to work in the U.S. without sponsorship. EOE.



i n v e n t

Hewlett-Packard State & Local Enterprise Services, Inc. is accepting resumes for the following positions:

IT Developer/Engineer • Sacramento, CA (Ref. #SLSACITDE21)

Research, design, develop, configure, integrate, test, and maintain existing and new business applications and/or information systems solutions including databases through the integration of technical and business requirements.

ITO Service Delivery Consultant • Trenton, NJ (Ref. #SLTREITO1)

Provide expertise for IT infrastructure, application infrastructure, and related services throughout the lifecycle of a deal in accordance with contractually established terms and conditions and established technical standards.

Services Information Developer

Vancouver, WA (Ref. #SLVANSID31), Olympia, WA (Ref. #SLOLYSID21) and Tallahassee, FL (Ref. #SLTALASH1)

Conceptualize, design, develop, unit-test, configure, or implement portions of new or enhanced (upgrades or conversions) business and technical software solutions through application of appropriate standard software development life cycle methodologies and processes.

Consulting Manager • Columbia, SC (Ref. #SLCSCSRA1)

Manage the delivery of high-quality innovative systems integration and consulting services.

Information Testing • Dublin, OH (Ref. #SLDUBIT21)

Design, develop and execute all testing-related activities on applications, infrastructure or hardware components of IT solutions, which include both third party software and internally developed applications and infrastructure.

Mail resume to Hewlett-Packard State & Local Enterprise Services, Inc., 5400 Legacy Drive, MS H1-6F-61, Plano, TX 75024. Resume must include Ref. #, full name, email address & mailing address. No phone calls please. Must be legally authorized to work in the U.S. without sponsorship. EOE.

form of systems analysis, dsgn & dvlpmt, systems integration &/or testing consulting. Min. Req. MS Mgmt Info Systems. Attn: HR, 362 Fairview Way, Milpitas, CA 95035.

UNIVERSITY OF CALGARY, Department of Computer Science, Assistant Professor Positions.

The Department of Computer Science at the University of Calgary seeks outstanding candidates for two tenure-track positions at the Assistant Professor level. Applicants from the areas of Database Management and Scientific Visualization are of particular interest. Details for each position appear at: <http://www.cpsc.ucalgary.ca/>. Applicants must possess a doctorate in Computer Science at the time of appointment, and have a strong potential to develop an excellent research record. The Department is one of Canada's leaders as evidenced by our commitment to excellence in research and teaching. It has large undergraduate and graduate programs and extensive state-of-the-art computing facilities. Calgary is a multicultural city that is the fastest growing city in Canada. Calgary enjoys a moderate climate located beside the natural beauty of the Rocky Mountains. Further information about the Department is available at <http://www.cpsc.ucalgary.ca/>. Interested applicants should send a CV, a concise description of their research area and program, a statement of teaching philosophy, and arrange to have at least three reference letters sent to: Dr. Carey Williamson, Head, Department of Computer Science, University of Calgary, Calgary, Alberta, Canada, T2N 1N4 or via email to: search@cpsc.ucalgary.ca. Completed applications received by December 15, 2011 will receive full consideration, though the review process will continue until the positions are filled. Hiring decisions will be finalized in Spring 2012, with the successful candidates joining the U of C on July 1, 2012. All qualified candidates are encouraged to apply; however, Canadians and permanent residents will be given priority. The University of Calgary respects, appreciates, and encourages diversity.

EXCHANGE ADMINISTRATOR (Unisys; Harrisburg): Manage large consolidated messaging systems, incl Exchange 2007 & Windows 2003/2008 supporting large number of mailboxes. Tier 2 support for multiple devices & domain controllers. High level troubleshooting related to Active & Sysvol Directories. Req. BS; 2 yrs. of IT exp. At least 2 yrs. exp in administering large Windows 2008/2003 Active Directory environs w/at least 50,000

users & 10 AD sites. 2 yrs. exp w/DNS & WINS in large environs w/at least 50,000 users. Extensive knowledge & exp in Exchange Administration, Windows Administration & DNS & WINS. Must live w/ in 60 minutes driving time of job site. Resumes to IEEE Computer Society, 10662 Los Vaqueros Circle, Box # COM54, Los Alamitos, CA 90720.

NOKIA INC. has a position in San Diego, CA: Senior Anite Automation Engineer: Exp. to involve testing; working w/ GSM/UMTS technologies; mobile phone test & development exp.; GSM/CDMA/WCDMA device interoperability conformance testing & debugging; test scripting & execution with base station emulator systems; & other duties/skills required. [Job ID: NOK-SD11-SAAE]. Mail resume to Nokia Recruiter, 3575 Lone Star Cir, Ste 434, Ft Worth, TX 76177 & note Job ID.

CAPTURE ADMINISTRATOR, Jersey City, NJ: Adjility Consulting seeks Capture Administrator for support & maintenance of imaging software solutions including EMC Documentum Captiva, Windows Server fileshares & IBML scanners; Req Master's in Comp Sci, Engg, Info Systems or related tech field & 2 yrs exp or Bachelor's & 5 yrs exp; Req 2 yrs exp in EMC Documentum Captiva; 1 or more of the following: -2 yrs Windows Operating System Platform administration exp, -1 yr Network Administration exp, -1 yr Package Software Installation exp, -1 yr Custom Software Installation exp, -1 yr Software administration exp; Req working knowledge of XML & XSL; Working knowledge of SQL; Working knowledge of 1 or more Systems Development Life Cycle & Methodology; Conversational English Verbal Skills; Professional English Written Skills; Experience with geographically dispersed teams; Demonstrated communication & collaboration skills; Up to 100% travel to client locations within US; Email resume to 85.adjility@helpdesk.net.

SENIOR DEVELOPER sought by Gemini Systems, LLC to dsgn & dvlp small to medium business applics using .Net, C#, VB.Net, front-end & back-end web based technologies. Master's deg in Electrical & Comp Engg or foreign educational equiv & 2 yrs rel. work exp (before or after Master's) as s/ware dvlpr reqd. Resume to: Gemini Systems, LLC, HR/K11, 61 Broadway, Ste. 925, NY, NY 10006.

FUJITSU NETWORK COMMUNICATIONS INC. has a Sr. Software Engineer (Req # FNC01736) job opportunity

available in Pearl River, NY. Design, develop and test control software for carrier grade switching platform. Submit resume to Fujitsu Network Communications, Staffing Department, 2801 Telecom Pkwy, Richardson, TX 75081. Req # must be noted or referenced when submitting resume.

SOFTWARE DEVELOPMENT ENGINEER - SERVER & TOOLS ONLINE OR OTHER needed by MICROSOFT CORPORATION in Redmond, WA. Design, implement and/or test computer software applications, systems or services, working with other engineers, working on standard or complex problems. Apply principles and techniques of computer science, engineering and/or mathematical analysis. Formulate and analyze software requirements. May be assigned to various projects that utilize the required technical skills to deploy successful product releases, from early product definition and scoping to detailed specification, implementation and roll-out phases. May work on the framework/cms and rendering framework for multiple sites utilizing ASP.NET; Silverlight; SOAP; web apps; web UI; scripting; and C#. 8AM - 5PM, Mon-Fri; \$122,605/yr, standard company benefits. Requires Bachelor's or foreign equivalent degree in Comp. Sci., Engineering, Math., Info. Sys., Physics or a related field & 1 year of software development experience formulating and analyzing software requirements utilizing C#. In lieu of a Bachelor's degree, employer will accept two additional years of front-end website development experience formulating and analyzing software requirements utilizing C#. Educ or exp must include: ASP.NET; Silverlight; SOAP; web apps; web UI; scripting; and C#. To apply, submit resumes to: Recruitment & Employment Office, MICROSOFT CORPORATION, Attn: Job Ref #: MIC47935, P.O. Box 56625, Atlanta, GA 30303.

SOFTWARE DEVELOPER (Indianapolis, IN) Dvlpmt of Linux chip drivers for Trident multichannel audio decoder chip & familiarity w/ Linux kernel i2c module incl use of audio precision eqpmt & certification tests for multichannel DTS/Dolby products. Dsgn & dvlpmt of h/ware boards for embedded products utilizing systems such as Linux, pSOS & VxWorks. BS or its equiv in Comp Sci & 5 yrs of exp reqd. Exp must incl building, testing & implmntn of Linux kernel & root configuration while adding new packages & modules as necessary to the build process Defining & tracking changes to source code as well as debugging of embedded products utilizing Dmalloc

Apple is looking for qualified individuals for following 40/hr/wk positions. To apply, mail your resume to 1 Infinite Loop 84-REL, Attn: LJ, Cupertino, CA 95014 with Req # and copy of ad. Job site & interview, Cupertino, CA. Principals only. EOE.

Supply Base Engineer/Manufacturing Engineer [Req. #9941713]

Work closely with the Industrial and Product Design teams in the early stages of a program to determine what is possible from certain processes. Requires Master's degree, or foreign equivalent, in Mechanical Engineering, or related degree, including manufacturing process development; cross functional communication with design and engineering; CNC machining; assembly fixtures; welding; painting processes; gluing processes; basic GD&T principles; and with CAD. Travel in Asia required 35% of time.

Project Lead [Req #9944109]

Design and configure SAP MM and Logistics modules and drive projects in the related areas. Requires Bachelor's degree, or foreign equivalent, in Engineering, or related degree, and 5 years professional experience in job offered or in a related occupation. Professional experience must be post-baccalaureate and progressive in nature. Also experience with managing SAP projects; SAP MM/Logistics in the areas of Supply Chain Management and Logistics/Transportation within a High Tech environment; direct procurement and master data management; integration with financials; SAP's Master Data Governance Application suites; SAP integration technologies; and, cross-functional projects with multiple systems integration. May have direct reports.

Sr. Process Engineer [Req. #10026571]

Responsible for design and process development of advanced flat panels from concept to product ramp. May require 5-10% international travel to organize technical discussion with panel suppliers in Japan and other Asian areas. Req.'s Master's degree, or foreign equivalent, in Physics, Industrial Engineering, Chemistry, Electronic Engineering, Engineering, or related plus 8 (Eight) years professional experience in job offered or in a related occupation. Must also have professional experience with TFT/CF/Cell process; TFT/CF/Cell process DOE, SPC and failure analysis; TFT device engineering, TFT LCD circuit; TFT-LCD structures and display optics; panel design and array process; negotiation with panel suppliers.

Hardware/Baseband Engineer [Req#10024960]

Design, implement, and integrate digital and analog circuit design. Req.'s Bachelor's degree, or foreign equivalent, in Electrical Engineering, or related plus Five (5) years professional experience in job offered or in a related occupation Professional experience must be post-baccalaureate and progressive in nature. Must have academic background or professional experience with: mobile processors, memory systems; analog audio, power supplies, RF components and circuits; wireless communication protocols such as GSM, WLAN. May require 10% of international travel.

Localization Specialist [Req #10050025]

Conducting an analysis of localization quality from multiple directions to avoid a slip-up in quality assurance. Req.'s Seven (7) years professional experience in job offered or in a related occupation. Must have academic background or professional experience with: localization tools and processes, software engineering, Mac OS, Windows, and UNIX. Fluency in Finnish language is also required, with working knowledge of Swedish, Italian and German.

Senior Supply Solutions Analyst [Req#10027844]

Work with key global and regional business users and extended IS&T teams to drive and implement strategic SCM solutions at Apple. Requires 10 years experience in job offered or in a related occupation, including supply chain management; project lifecycle management; SAP SCM/APO/GATP solutions; coordinating all activities between ABAP, Basis, technical infrastructure, support, application, and business teams; business disciplines including supply chain management, ATP, Allocation, Supply Network Planning, Demand Planning, and Logistics.

Software Engineer [Req#10028259]

Create, maintain, and enhance software components for internal software projects. Req.'s Bachelor's degree, or foreign equivalent, in Computer Engineering, or related plus Six (6) years professional experience in job offered or in a related occupation Professional experience must be post-baccalaureate and progressive in nature. Must have professional experience with: Java and Java Enterprise technologies (EJBs, Servlets, JDBC); server-side web technologies (Struts and JSO); relational databases (Oracle 10G); Unix and/or Linux; and client-side web technologies (HTML, CSS, AJAX, and JQuery).

& JTAG. Send resume to Ms. Landsman, D&M Holdings US, Inc. 100 Corporate Dr, Mahwah, NJ 07430.

SENIOR SERVICES CONSULTANT (Islandia, NY & locs throughout US). Efficiently dsgn., dev. & implmnt. sol., implmnt. & del. sol. using serv. implmnt life cycle. Facilitate prod. & sol. expertise, work w/ Clarity Database Schema/Functionality & Studio & interact w/ supp. team to track tech. issues. Act as Tech. Clarity Admin./Dev. & gather reqs. for Clarity proj. incl., CA Clarity, Java, Oracle/SQL Server, PL/SQL. Install & config solutions. Req's.: Bach's deg. or for. equiv. in CS, CIS, Math, Eng (any), or rel tech. fld. + 5 yrs. prog. exp. in job offd &/or rel pos. Must have exp. w/ Clarity Functionality, Clarity Studio, Oracle/SQL Server, Java Prgmg & Clarity D/base Schema; acting as a Tech. Clarity Admin./Dev.; serv. implmnt life cycle; gathering reqs. for Clarity proj. Freq. travel req. Work from home benefit avail. Please send resume to: Althea Wilson, CA Technologies, One CA Plaza, Islandia, NY 11749, Req #24992.

ENGINEERED LUMBER SOFTWARE DEVELOPER. BlueLinx Holdings Inc. (Atlanta, GA) hiring for Engineered Lumber Software Developer position. May telecommute. Send resumes to Leslie.Lovelace@BlueLinxCo.com.

CALIFORNIA HEART CENTERS GROUP in Industry, CA seeking software developer (Java) to design and implement cardiac MRI analysis system. Exp. & MS deg. Req. Pls. e-mail CV to GM ramon@chcgp.com.

DATABASE/SYSTEM ADMINISTRATOR in Chicago, IL to administer co.'s d/base systems (Therapists/Patients) using Microsoft SQL Server, Microsoft Access, & other D/base Mgmt Systems (DBMS); to maintain & administer Web server & middleware tools used to dvlp Web d/base systems & framework. Req'd: Bachelor's deg in Comp Sci or foreign deg equiv., or related fields, 2 yrs exp. in job offd. Mail resumes to Don Jorge Pagunlatan, Integrated Therapy Specialists LLC, 5946 N. Milwaukee Ave., Chicago, IL 60646. Ref #111303291. No calls. No recruiters. Job applicants only.

Optimization-based software solutions provider in Gainesville, FL, has the following openings: ***Sr. Systems Eng.** - Conduct research on network and heuristic optimizations; develop prototypes for data structures and algorithms.

Ph.D. in Operations Research, Industrial Eng., or related field req'd. ***Sr. Software Eng.** - Design, implement and test software solutions, algorithms, and data structures to solve transportation and logistics problems. Bachelor's degree, or for. equiv., in Comp. Sci., Eng., or a related field and 5 years progressive experience in the field req'd. Salary commensurate with exp. Full-time. Mail resume to: HR, Innovative Scheduling, Inc., 2153 SE Hawthorne Road, Suite 128, Gainesville, FL 32641.

LEAD SOFTWARE ENGINEER: SW design, build, test for all Core Apps; app support. Lead anal., design, build, test phases. Code reviews, automated unit tests. Contrib. to devel. standards & best prac. Mentor junior engineers. Devel/implm. SOA & core bus. object apps. J2EE app. design incl. web-based thin client arch., HTML, XML, CSS, Javascript, OOP, SQL, HQL. MS CS, Engg, related + 5 yrs relevant exp. or BS + 7 yrs. Catalyst Rx, 1650 Arch St, Phila PA 19103. Apply to hr@catalystrx.com w/ job# LSE in subj. line. EOE.

NOKIA, Inc. has a position in Sunnyvale, CA: Senior Engineer: Exp. to involve working with virtual machines on mobile & embedded platforms; code optimization for ARM processor for mobile apps; work with support & maintenance of mobile browser, flash player & Qt framework; Maemo/MeeGo development exp. & JIT virtual machines; & other duties/skills required. [Job ID: NOK-SV11-SENE]. Mail resume to Nokia Recruiter, 3575 Lone Star Cir, Ste 434, Ft Worth, TX 76177 & note Job ID#.

JIVE SOFTWARE, INC. has the following job opportunities available: Software Engineer (Analytics) in Palo Alto, CA: Act as a primary contributor to the design of Jive's next generation analytics solution. Evaluate and select the back-end datastore to store customer analytical data. Job code SEA-CA. Sr. Professional Services Engineer in Portland, OR: Develop new features, customizations and enhancements to Jive Software's core product as part of our client engagements. Job code PSE-OR. Mail resumes to: Jive Software, Inc., Attn: Matt Gradin, 915 SW Stark Street, Suite 400, Portland, OR 97205. Must reference job code.

TECHNICAL LEAD F/T (Poughkeepsie, NY) Must have Master deg or the foreign equiv in Electrical Engg, Engg, Comp Engg, Electrical & Comp Engg, or related with one (1) yr of exp leading

and managing a 3 member team & be proficient in the application analysis, design, development, implementation and testing of software applications through full product development life cycle and release process. Responsible for timely delivery of code in accordance with coding standards and best practices. Provide project plan estimation and rollout strategy in collaboration with the project manager. Mentor the junior team members by providing technology knowledge transition. Create design documents, data conversion documents, technical specifications, class diagram flowchart based on requirements session. Take ownership of the system and programming documentation. Manage production tickets workload among developers and resolve them in timely fashion. Provide subject matter expertise and implement the code using following tools/technologies: Java/J2EE, Forte, Portlets, JSF, Spring, Hibernate, EJB, AJAX, UML, Rational Rose, JDBC, Oracle, DB2, SQLDeveloper, Websphere MQ, JMS, TIBCO, LDAP, SOAP,JAX-WS, AOP, OOAD, Perl, RAD, Eclipse. Send resume: Indotronics Int.l Corp., Recruiting (VA), 331 Main St, Poughkeepsie, NY 12601.



is seeking a

Engineer, Sr Staff-SW Development

Matawan, NJ

Req. MS (or foreign equiv.) in Electrical/Computer Engg or rel & 3 yrs exp. Perform analysis of problem scenarios & generate efficient solutions. Req 10-15% domestic & intl travel. F/T. Must have unrestricted U.S. work authorization.

Mail resumes to:
HR Operations Coordinator
5300 California Ave.
Bldg 2, #22108B
Irvine, CA 92617
Must reference
job code ENG8-NJNP.



Applied Materials, Inc. is accepting resumes for the following position in **Hopewell Junction, NY**:

Process Support Engineer

Ref# NYVKA

Executes Process Engineering projects to qualify or improve the process performance of company's products. Troubleshoots system performance issues and implements appropriate resolution.

Please mail resumes with reference number to Applied Materials, Inc., 3225 Oakmead Village Drive, M/S 1217, Santa Clara, CA 95054. No phone calls please. Must be legally authorized to work in the U.S. without sponsorship. EOE.

www.appliedmaterials.com



is seeking a

RF/Wireless Engineering Manager

Sunnyvale, CA

Reqs MS in EE or related and 5 yrs exp. Reqs exp in Wireless communication protocols, including WLAN and GPS; Computer programming using the languages Python and PERL; Hardware design; RF (Radio Frequency) measurement; Windows and Linux operating systems.

Mail resumes to:

HR Operations Coordinator
5300 California Ave.

Bldg. 2, #22108B

Irvine, CA 92617

Must reference

job code ENG8-SVCABD.

ARM, Inc,

the world's leading semiconductor intellectual property (IP) supplier, has openings in **Austin TX**:

Senior Design Engineer

Job Code: DE2SR

Staff Design Engineer

Job Code: DE2ST

Principle Design Engineer

Job Code: DE2PR

Design and develop leading edge high-performance, high-density, low-power and ultra low-power memory generators which are optimized for each silicon technology.

If interested, ref job code and send resume to: ARM, Inc. Attn: Sr. Recruiter, 150 Rose Orchard Way, San Jose, CA 95134. EOE.



is seeking

Senior Staff Unix Systems Engineer

Santa Clara, CA • Reference Job Code: ENG8-SVCASV

Associate's degree (or foreign equiv.) in EE or rel + 9 yrs of professional exp. Determine method and procedures on new assignments, & provide guidance to other personnel. Broadcom Corporation. Santa Clara, CA. F/T. Must have unrestricted U.S. work authorization. Mail resumes to HR Operations Coordinator, 5300 California Avenue, Bldg. 2, #22108B, Irvine, CA 92617. Must reference job code ENG8-SVCASV.

Scientist, Sr. Staff – Electronic Design

Austin, TX • Reference Job Code: ENG8-AUTXCZ

Req. PhD (or foreign equiv.) in EE or rel + 1 yr of exp. Responsible for Block level circuit design. Broadcom Corporation. Austin, TX. F/T. Must have unrestricted U.S. work authorization. Mail resumes to HR Operations Coordinator, 5300 California Avenue, Bldg. 2, #22108B, Irvine, CA 92617. Must reference job code ENG8-AUTXCZ.

Mail resumes to: HR Operations Coordinator, 5300 California Ave., Bldg. 2, #22108B, Irvine, CA 92617

Apple is looking for qualified individuals for following 40/hr/wk positions. To apply, mail your resume to 1 Infinite Loop 84-REL, Attn: LJ, Cupertino, CA 95014 with Req # and copy of ad. Job site & interview, Cupertino, CA. Principals only. EOE.

Audit Engineer [Req #9597999]

Ensures the correct marriage between audio software and acoustic hardware in Apple's iOS devices. Req.'s Master's degree, or foreign equivalent, in Signal Processing, Electrical Engineering, Computer Engineering, or related. Five (5) years professional experience in job offered or in a related occupation. May require up to 5-10% of domestic and international travel. Must have academic background or professional experience with: Speech signal processing algorithms, including echo cancellation, noise suppression, active noise cancellation, and system identification and adaptive filtering techniques; Standard audio measurement tools, e.g. Sound Check, Audio Precision, HEAD Acoustics ACQUA, Rohde & Schwartz CMU, Rohde & Schwartz UPV/UPL; Acoustic measurements (HATS, handset positioners, artificial ear couplers, reference acoustic sources & microphones); speech intelligibility, psychoacoustics, speech quality, and perception.

Information Systems Engineer [Req #9728502]

Design, analyze, evaluate, test, debug and implement applications programs in support of various functional areas such as materials, marketing, accounting, or human resources. Req.'s Bachelor's degree, or foreign equivalent, in Computer Science or related degree plus Six (6) years professional experience in job offered or in a related occupation. Professional experience must be post-baccalaureate progressive in nature. Must also have professional experience with: Java Programming Language; Relational Database Solutions; J2EE, Application Server Frameworks; design of Enterprise Applications; and Unix/Related operating systems.

Database Administrator [Req #9745038]

Responsible for Oracle production and development administration, installation and maintenance on UNIX in a 24/7 environment. Req.'s Bachelor's degree, or foreign equivalent, in Computer Science, Computer Engineering or related plus Seven (7) years professional experience in job offered or in a related occupation Professional experience must be post-baccalaureate and progressive in nature. Must have professional experience with: SQL, OEM, TKPROF, ADDM, AWR, VMSTAT, PL/SQL.

Hardware Development Engineer (Senior Antenna Design Engineer) [Req#9746899]

Design antennas for mobile communication products. Requires Master's degree, or foreign equivalent, in Electrical or Computer Engineering, or related, including antenna design; antenna and RF test and measurement; and electrical/electronic engineering.

Apple is looking for qualified individuals for following 40/hr/wk positions. To apply, mail your resume to 1 Infinite Loop 84-REL, Attn: LJ, Cupertino, CA 95014 with Req # and copy of ad. Job site & interview, Cupertino, CA. Principals only. EOE.

Senior Software Engineer [Req#9864389]

Develop C++ software for embedded platforms. Req.'s Eight (8) years professional experience in job offered or in a related occupation Professional experience must be post-baccalaureate and progressive in nature. Also experience with: C&C ++ programming language; design patterns; embedded debugging; implementing application frameworks; driver development; multi-threaded design, implementation and debugging.

QA Engineering Lead [Req #9887027]

Direct team members to help achieve high quality software releases on iOS devices. Requires Master's degree, or foreign equivalent, in Electronics and Communication Engineering, or related degree. Must have academic background or professional experience with: UMTS/WCDMA and GSM protocols; Perl; system level design for a mobile wireless; physical layer basic communication, network simulators including Anristu, Spirent, & Racial; wireless video transmission; scripting/automation/OS; and C. May include managing direct reports.

Cellular Systems Architect [Req#9884840]

Lead creation of future cellular chipsets and technologies for iPhone & iPad products. Req.'s PhD degree, or foreign equivalent, in electrical engineering, or related plus ten (10) years professional experience in job offered or in a related occupation. Must have professional experience with: architecture of cellular chipsets & platforms for 3GPP HSPA; CPU internal architecture, SW development from firmware through upper layers, and security OR cellular RF transceiver & front-end architecture and PCB-level implementation; PCB-level application engineering of digital baseband & power management incl. host interfaces, digital audio routing, memory interfaces, & DC/DC converter issues; 3GPP standards, testing, and approval processes; identify newer wireless subsystems such as HSDPA, HSUPA, LTE, Wimax, as well as application issues of both existing and new wireless technologies as they relate to our products.

SAP BW Development Lead [Req #9917967]

Lead the design, development and deployment of reporting and analytical solutions for various business functions at Apple using SAP Business Warehouse (BW) and other reporting tools and technologies. Requires Bachelor's degree, or foreign equivalent, in Engineering, Electronic Engineering, or related degree and 5 years professional experience in job offered or in a related occupation. Professional experience must be post-baccalaureate and progressive in nature. Must have professional experience with: data modeling in BW, developing Info providers, Cubes and Operational Data Stores; deploying end-to-end solutions in BW platform for reporting and analytic needs of various business functions; extracting data from SAP and non-SAP environments and loading the data set into BW Info Providers; report design using BEx Query designer and Web Application designer on BW data models; optimization and performance management of complex BW reports using BW Accelerator and the system as a whole; ABAP development to analyze, optimize and troubleshoot the transformation in BW.

Verification Engineer [Req. #9917921]

Responsibilities include the formal verification of multiple intellectual property blocks on Apple's System-on-a-Chip (SoC) hardware designs. Requires Master's degree, or foreign equivalent, in Electrical Engineering or related field and 4 years professional experience in job offered or in a related occupation. Professional experience must be post-baccalaureate and progressive in nature. Must have professional experience with: System Verilog, Verilog, Verification Methodology Manual (VMM) and System Verilog Assertions; digital logic design and computer architecture; computer-aided design tools including Jaspergold and IFV; Tcl, PERL, C++, and Python.



i n v e n t

HP Enterprise Services, LLC
is accepting resumes for:

Presales EDS Consultant

Plano, TX

Reference: ESPLANNG21

Create, qualify, develop, and sell company
solutions to customers.

Mail resume to HP Enterprise Services,
LLC, 5400 Legacy Drive, MS H1-6F-61,
Plano, TX 75024. Resume must include
Ref. #ESPLANNG21, full name, email
address & mailing address. No phone calls
please. Must be legally authorized to work
in the U.S. without sponsorship. EOE.



Applied Materials, Inc. is accepting
resumes for the following position
in **Boise, ID**:

Process Engineer

Ref# IDSLE

Develops new or modified process
formulations, defines process or
handling equipment requirements
and specifications, reviews process
techniques and methods applied in
the fabrication of integrated circuits.

Please mail resumes with reference
number to Applied Materials, Inc.,
3225 Oakmead Village Drive, M/S
1217, Santa Clara, CA 95054. No
phone calls please. Must be legally
authorized to work in the U.S.
without sponsorship. EOE.

www.appliedmaterials.com


Applied Materials, Inc. is accepting
resumes for the following position
in **Austin, TX**:

Engineering Project Manager

Ref# TXSNA

Responsible for development of new
engineering projects, sustaining current
and legacy systems from an engineering
and cost standpoint, product change
and revision control and strategic
roadmap planning and implementation.

Please mail resumes with reference
number to Applied Materials, Inc.,
3225 Oakmead Village Drive, M/S
1217, Santa Clara, CA 95054. No
phone calls please. Must be legally
authorized to work in the U.S. without
sponsorship. EOE.

www.appliedmaterials.com


is seeking a

Business Development Manager

San Jose, CA

Req. MS (or foreign equiv.) in Electrical
Engg, CS, or rlted and 5 yrs exp. Enable
higher-level software application on the
chip & resolve internal & external
customer problems. F/T. Must have
unrestricted U.S. work authorization.

Mail resumes to:
HR Operations Coordinator
5300 California Ave.
Bldg 2, #22108B
Irvine, CA 92617
Must reference
job code BDM-SVCADN.



is seeking a

Engineer, Principal Firmware

Andover, MA

Req. MS (or foreign equiv.) in Electrical
Engg, CS, or rlted and 5 yrs exp. Enable
higher-level software application on the
chip & resolve internal & external
customer problems. F/T. Must have
unrestricted U.S. work authorization.

Mail resumes to:
HR Operations Coordinator
5300 California Ave.
Bldg 2, #22108B
Irvine, CA 92617
Must reference
job code ENG7-MAYM.



i n v e n t

HP Enterprise Services, LLC
is accepting resumes for:

Business Consultant

Schaumburg, IL
Reference: ESSCAA11

Provide business domain solution, process,
strategy, business case and change consult-
ing to client.

Mail resume to HP Enterprise Services,
LLC, 5400 Legacy Drive, MS H1-6F-61,
Plano, TX 75024. Resume must include
Ref. #ESSCAA11, full name, email address
& mailing address. No phone calls please.
Must be legally authorized to work in the
U.S. without sponsorship. EOE.

THE PROFESSION

Continued from page 112

However, these were only mock-ups; the real-world assistants were more primitive. A well-known example is the Office Assistant that came with Microsoft Office between 1998 and 2003, depicted in various ways, including as a paper clip and an Einstein caricature. This feature got a negative response from users; some went as far as to develop applications that allowed users to “shoot down”

Although speech generation and recognition might be useful in some cases, a smartphone keyboard and display can be used to communicate with the assistant. Thus, there's no need to give the assistant human-like capabilities or for it to “understand” the user. Today, we can implement a “knowledge navigator” without the magic parts, leaving the role of the human to the user.

nearby hotels that have a room available and are within the price range you normally select. In some circumstances you might want to translate the word “hotel” into another language. The system could offer this as a secondary option, or you might have to add more keywords.

In the early days of computing, an interface was an empty line on a teletype or a blinking “>” on a display. Nowadays, this has been replaced by apps and their form-based input. The drawback is that the user must choose the right app for the function, fill out forms, provide codes, and so on. Perhaps we should reintroduce the command line interface, but this time let an assistant parse the data?

The ideal interface doesn't require any input but takes its cues from the available data.

their assistant, which had the bad habit of popping up in unexpected situations, even when the function was turned off. *Smithsonian Magazine* called it “one of the worst software design blunders in the annals of computing” (R. Conniff, “What's Behind a Smile?,” Aug. 2007, pp. 46-53).

The advice the Office Assistant offered was often irrelevant, downright silly, or too trivial to be of any use. The problem was that the feature took its clue from a skerrick of context, for example, just a keyword. For example, if you started by typing “Dear,” a message would appear saying, “It looks like you're writing a letter. Would you like help?”

Today, we have more possibilities. Since we use our computers for most administrative tasks, from accessing our bank accounts to writing a shopping list, the assistant could have access to all relevant data. When it's stored in the cloud, it becomes irrelevant if a smartphone or a PC was used to input the data. That is, the assistant could have access to phone data, text messages, e-mail, contact lists, social networks, and background information from the Web. By combining this data with the user's location, an assistant running on a smartphone should be able to perform “intelligent” deductions in many situations.

CONTEXTUAL DATA

What if we offer the assistant the word “hotel?” Is this a meaningful command? Of course, there are many different interpretations, including

- I want to book a hotel.
- I need to view a booking, change it, or delete it.
- I need directions to a hotel.

However, with contextual data, things become clearer. In most cases, a simple word such as “hotel” can be given a meaningful interpretation when such data is available.

Let's assume that the system has an overview of all your bookings, your current location, and your home address. If you've booked a hotel in Rome starting today and your current location is at the airport or any place in Rome, the system could retrieve the booking; offer the hotel's name, address, and phone number; and give directions from your current location. If you have a rental car booking, the directions should be for travel by car. If you're near the hotel, the assistant could provide directions for walking to it. If you're farther away, it could offer suggestions for finding public transportation.

If you're away from your hometown and don't have a hotel reservation, the system should offer a selection of

AUTOMATIC ASSISTANCE

Some time ago, we stayed at a large airport hotel. There were six floors and as many elevators, all in the same area. When we went from the reception area to the elevators, the door of one elevator was open so that we could enter and press the button for our floor without having to wait. When we went down for dinner, we also found an elevator waiting. With a simple addition to the control program that made an elevator available at each floor, this hotel was able to offer guests a convenient service.

With some extension, this idea also can be used when elevators are busy or when there are fewer elevators than floors. A “here” command for bringing an elevator to the reception floor could be executed automatically when the receptionist hands a keycard to a guest. The same action could be performed when guests retrieve the keycard as they leave their room. Thus, a better service can be offered just by using contextual data.

This is the ideal interface—one that doesn't require any input but takes its cues from the available data.

As another example, assume that you're on your way to the bus stop. With data including the time of day,

your location, the bus stop location, and the bus route, the assistant could automatically present the bus schedule on the phone display, or better still, it could count down the minutes until your bus arrives. You could explicitly enter the route data, or the assistant might be able to deduce it from your commuting history. The data could be presented on an “I feel lucky” display on the phone. That is, you could just take a look at this display to get all the information you need. But if you pass the bus stop on your way to the grocery, the phone would clear the data. When you enter the grocery, the system would, of course, present your shopping list.

In yet another example, imagine yourself driving to the airport. The weather is bad, traffic is dense, and the cars ahead are moving along slowly. You wonder if you’ll catch the flight. Hopefully, there might be other delays such as a change in the flight schedule because of the bad weather. With a context-sensitive interface, you would only need to look at the smartphone’s “I feel lucky” screen. The system should already have deduced what information you might need, such as an updated departure time for your flight or directions that offer a way around the traffic congestion.

DATA STANDARDS

Storing data in the cloud offers the opportunity to have access to all relevant information. This is necessary to understand the user’s context. The assistant will interpret a “later flight” or “hotel” command incorrectly if it doesn’t have access to all bookings. Today, however, when we use computers for nearly everything, this data is available. But access to data isn’t enough. It must also be presented in a useful form.

This is the great challenge. It’s possible to extract the necessary information from an e-mail confirmation of a booking, such as the

hotel name, arrival date, departure date, reference number, and so on using a Web service such as TripIT (www.tripit.com). However, storing the booking information in a standardized format in the cloud would be more convenient.

Thus, to get the full advantage of having an assistant, we need a more formalized infrastructure than we have today. This will require agreements on standards (S. Ortiz Jr., “The Problem with Cloud-Computing Standardization,” *Computer*, July 2011, pp. 13-16), but standardization isn’t easy to achieve.

Initially, like its predecessors, the smartphone evolution focused on new functions. Now, there’s a need to consider the user interface. This time, the focus isn’t on organizing the input or developing convenient forms or menu systems, but on finding ways to avoid input. A personal assistant can do this for us. As we’ve seen, a human-like assistant isn’t necessary. In many cases, having background data from the cloud, as well as the time of day and the user’s location, should be sufficient.

As computing professionals, perhaps our job in the next decade

To get the full advantage of having an assistant, we need a more formalized infrastructure than we have today.

Although technical, economic, and political issues can hamper the work, there are several success stories. In Norway, for example, banks have developed a common ID system for online banking. Combining this with a national interbank system has kept transaction costs very low, and it’s also convenient for users with accounts in several banks.

The assistant’s interpretation of the context might not always be correct because it doesn’t have the most recent data or it makes the wrong deductions. However, since the system is serving as the user’s personal assistant, even if these errors are annoying, they’ll usually be recognized. The fact that most advice is on current events, especially what’s happening right now, will also be of use in detecting erroneous deductions.

As computers evolved, they left the centralized data-center, moved into local centers, then to the desktop, next to the laptop, and now into our pockets as smartphones.

isn’t just to add functions and new gadgets to current devices, but to ensure that users get the full benefits of the new technology. Ideally, we should seek “free” solutions so that the user can get valuable information without having to pay based on button clicks. ■

Kai A. Olsen is a professor at the University of Bergen and Molde University College in Norway as well as an adjunct professor at the School of Information Services, University of Pittsburgh. Contact him at kai.olsen@himolde.no.

Alessio Malizia is an associate professor in the Computer Science Department at Universidad Carlos III de Madrid, Spain. Contact him at alessio.malizia@uc3m.es.

Editor: Neville Holmes, School of Computing and Information Systems, University of Tasmania; neville.holmes@utas.edu.au

cn Selected CS articles and columns are available for free at <http://ComputingNow.computer.org>.

THE PROFESSION

Automated Personal Assistants

Kai A. Olsen

University of Bergen and Molde University College, Norway

Alessio Malizia

Universidad Carlos III de Madrid, Spain



Instead of just adding functions and new gadgets to current devices, we should ensure that users get the full benefits of the new technology.

You're in the middle of a strange city. Your hotel should be nearby, but you can't find it. There are two options. One, open an Internet connection on your smartphone, find the map service, input the city's name, download a city map, change to a convenient map scale, type the hotel's address, and let the GPS system lead you to your destination. Two, ask a passerby.

The easy choice is option two. That is, our smart devices can do the job, but in most cases they take too much effort. Keying comes at a cost. This is especially the case when we have to use an onscreen keyboard. But even when input is as simple as a button click, using small displays is time-consuming and irksome.

USING CONTEXT INFORMATION

Often, systems ask for input they could either find from available data or infer from contextual information. As an example, consider a situation that many of us have experienced.

You're in a meeting that drags on

and on. At some point, it's clear that there's no chance of catching the five o'clock plane back home. You'll have to leave the meeting, get an Internet connection, log in to your airline's website, give information such as your name and booking reference, and change the booking to a later flight. Some airlines even require you to make a phone call in this circumstance.

However, if the booking system could use contextual information, you'd be able to perform the whole operation without even interrupting the meeting. A text message to the airline saying "later flight" should suffice. The airline's system should be able to identify you by phone number, retrieve the booking for this evening's flight, and return a set of options for later flights, asking you to choose one. It could even book the next flight automatically, letting you change to another if this option isn't suitable.

But this requires the airline to offer such an option. Instead of waiting for this to happen, we could use an agent running on our smartphone to change the flight based on data from the

initial booking. This would be easy to do if the smartphone agent made the booking in the first place, but it should also be possible if our booking data is readily available, as in the cloud. The smartphone could implement this agent as an app, but embedding the agent in the operating system would perhaps be simpler.

A PERSONAL ASSISTANT

The idea of having a computer system that could act as a personal assistant has been around for many years. As early as 1987, Apple CEO John Scully described the "knowledge navigator," a device that used software agents to assist the user (*Odyssey: Pepsi to Apple ... A Journey of Adventure, Ideas and the Future*, Harper & Row).

Several videos from Apple depicted this concept, envisioning the assistant as a bow-tied butler having human properties. The assistant had natural-sounding speech, speech recognition, and the ability to grasp the underlying semantics and actually understand what the user was saying.

Continued on page 110

Organizing Committee**Chair:**

H. Kobayashi Tohoku Univ.

Vice Chairs:J. Torrellas Univ. of Illinois,
Urbana-Champaign

K. Uchiyama Hitachi

C. -M. Kyung KAIST

H. Amano Keio Univ.

Secretaries:

K. Suzuki Renesas

H. Igura NEC

Treasurers:

R. Egawa Tohoku Univ.

K. Nitta NTT

Program Chairs:

M. Ikeda Univ. of Tokyo

F. Arakawa Renesas

Publicity Chair:

M. Suzuki Panasonic

Publication Chairs:

Y. Unekawa Toshiba

Y. Hirose Fujitsu Labs.

Registration Chair:

K. Takano IBM

Local Arrangement Chairs:

Y. Nitta Renesas

A. Hashiguchi Sony

Web Manager:

Y. Sato JAIST

Advisory Committee**Chair:**

T. Nakamura Keio Univ.

Chair Emeritus:

M. J. Flynn Stanford Univ.

Advisory Emeritus:

T. L. Kunii Univ. of Tokyo

Members:

D. Allison Stanford Univ.

D. B. Alpert Camelback
Computer Architecture

A. J. Baum Intel

D. A. Draper True Circuits
(TCMCOMP Chair)

M. A. Franklin Washington Univ.

G. Goto Yamagata Univ.

Y. Hagiwara Sojo Univ./AIPS

S. Hijiya Fujitsu Labs.

S. Iwade Osaka Inst. of Tech.

L. Jow Hewlett-Packard

R. Kasai NTT Electronics

S. Kohyama Covalent Materials

T. Kunio NEC

K. Kushima NTT

T. Makimoto TechnoVision
Consulting

O. Mencer Imperial College

H. Mochida Rohm

Y. Mori CM Engineering

A. Morino SIRIJ

J. Naganuma Shikoku Univ.

M. Nishihara AIPS

T. Nukii Sharp

T. Ogura Ritsumeikan Univ.

A. Omondi Yonsei Univ.

T. Tabata Sanyo Electric

T. Watanabe Riken

N. Woo Samsung

S. Yamaguchi Panasonic

H. -J. Yoo KAIST

(in alphabetical order)

IEEE Symposium on Low-Power and High-Speed Chips**COOL Chips XV**

Yokohama Joho Bunka Center, Yokohama, Japan

(Yokohama Media & Communications Center, Yokohama, Japan)

April 18 - 20, 2012

CALL FOR CONTRIBUTIONS

COOL Chips is an International Symposium initiated in 1998 to present advancement of low-power and high-speed chips. The symposium covers leading-edge technologies in all areas of microprocessors and their applications. The COOL Chips XV is to be held in Yokohama on April 18-20, 2012, and is targeted at the architecture, design and implementation of chips with special emphasis on the areas listed below. The COOL Chips Organizing Committee will ask the IEEE MICRO to publish selected papers in a special issue on COOL Chips XV.

Contributions are solicited in the following areas:

- **Low Power-High Performance Processors for -
Multimedia, Digital Consumer Electronics, Mobile, Graphics, Encryption, Robotics,
Automotive, Networking, Medical, Healthcare, and Biometrics.**
- **Novel Architectures and Schemes for -
Single Core, Multi/Many-Core, Embedded Systems, Reconfigurable Computing, Grid,
Ubiquitous, Dependable Computing, GALS and Wireless.**
- **Cool Software including - Parallel Schedulers, Embedded Real-time Operating System,
Binary Translations, Compiler Issues and Low Power Techniques.**

Proposals should consist of a title, an extended abstract (up to 3 pages) describing the product or topic to be presented and the name, job title, address, phone number, FAX number, and e-mail address of the presenter. The status of the product or topic should precisely be described. If this is a not-yet-announced product, and you would like to keep the submission confidential, please indicate it. We will do our best to maintain confidentiality. Proposals will be selected by the program committee's evaluation of interest to the audience. Submission should be made by e-mail, (Author's kit can be obtained from <http://www.coolchips.org/>)

to: Makoto Ikeda, Program Chair e-mail: submit_xv@coolchips.org**Author Schedule: January 27, 2012 Extended Abstract Submission (by e-mail)****March 12, 2012 Acceptance Notified (by e-mail)****March 27, 2012 Final Manuscript Submission**

You are also invited to submit proposals for poster sessions by e-mail,

to: M. Muroyama, Poster Chair e-mail: poster_xv@coolchips.org**Author Schedule: March 19, 2012 Poster Abstract Submission (by e-mail)****March 27, 2012 Poster Acceptance Notified (by e-mail)**For more information, please visit <<http://www.coolchips.org/>>.For any questions, please contact the Secretariat <cool_xv@coolchips.org>.

Sponsored by the Technical Committees on Microprocessors and Microcomputers and Computer Architecture of the IEEE Computer Society. In cooperation with the IEICE Electronics Society, ACM SIGARCH and IPSJ.

**IEEE**IEEE
computer
society**Program Committee****Chairs:** M. Ikeda (Univ. of Tokyo), F. Arakawa (Renesas)**Vice Chair:** H. Shimada (NAIST)**Poster Chair:** M. Muroyama (Tohoku Univ.)**Special Session Chair:** H. Tomiyama (Ritsumeikan Univ.)**Members :**

A. Ben-Abdallah (Aizu Univ.)

T. Harada (Yamagata Univ.)

Y. Inoguchi (JAIST)

T. Kobori (NEC)

M. Kuga (Kumamoto Univ.)

K. Morioka (Fujitsu Labs.)

H. Takizawa (Tohoku Univ.)

T. Yamada (Hitachi)

J. Yao (NAIST)

K. -R. Cho (Chungbuk National Univ.)

K. Hashimoto (Fukuoka Univ.)

H. Kawaguchi (Kobe Univ.)

T. Kodaka (Toshiba)

G. Lee (Korea Univ.)

B. -G. Nam (Chungnam National Univ.)

N. Togawa (Waseda Univ.)

T. Tsutsumi (Meiji Univ.)

K. S. Yeo (NTU Singapore)

A. Gupta (Freescale)

T. Hashimoto (Panasonic)

K. Kimura (Waseda Univ.)

Y. Kodama (Univ. of Tsukuba)

S. -J. Lee (TI)

Y. Shibata (Nagasaki Univ.)

T. -H. Tsai (NCU Taiwan)

H. Yamauchi (Samsung)

(in alphabetical order)

(As of October 12, 2011)



We Live Quality

Software quality is in our DNA. For over 15 years, we've lived and breathed it. The reason is simple: Your software affects our friends, our families, and ourselves.

Whether it's the latest video game or a secure banking web site or the software that analyzes medical test results, we want it to work right because we rely on it.

From our expert Consulting and Agile Services teams, to our award-winning application lifecycle management (ALM) solutions, to our world-class customer support, Seapine has helped thousands of companies worldwide build, test, and deploy quality software.

Go with Seapine, and get serious about software quality.

www.seapine.com



© 2011 Seapine Software, Inc. All rights reserved.