I-LIGHT

I-Light Symposium
March 2004
Proceedings

Program Committee

Gary Bertoline, Associate Vice President
Envision Center, Information Technology at Purdue

Steve Dunlop, Project Director
Information Technology at Purdue

Greg Flint (photos)

Diane Jung, Special Events Director
Office of the Vice President for Information Technology and CIO, Indiana University

Wendy Lin, Systems Programmer
Information Technology at Purdue

Greg Moore, Communications Specialist
Office of the Vice President for Information Technology and CIO, Indiana University

Mary Moyars-Johnson, Associate Vice President of Communications
Information Technology at Purdue

Bryan Putnam, User Services Programmer
Information Technology at Purdue

John Samuel, Director, Center for Statistical and Mathematical Computing
Office of the Vice President for Information Technology and CIO, Indiana University

Nancy Smigiel, Secretary for Envision
Information Technology at Purdue

Craig A. Stewart, Director
Research and Academic Computing and Director, Indiana Genomics Initiative
Information Technology Core, Indiana University

Julie Wernert, Communications Specialist
Office of the Vice President for Information Technology and CIO, Indiana University

Karen Whitney, Manager, Research Interactions and Communications
Information Technology at Purdue

Acknowledgments

The 2004 I-Light Symposium was sponsored by the Office of the Vice President for Research and Information Technology, Indiana University, and by the Office of the Vice President for Information Technology and CIO, Purdue University.

The I-Light network was made possible by a special State appropriation approved by the Indiana Legislature and the late Govenor Frank O'Bannon in 2001.

Research projects described in these Proceedings were supported by various grants. Key grants include:

*Analysis and Visualization of Instrument-Driven Data (AVIDD) - National Science Foundation Grant No.EIA- 0116050.

*Scientific Applications on Arrays of Multiprocessors (SCAAMP) - National Science Foundation under Grant No. CDA-9601632."

*Indiana Genomics Initiative. The Indiana Genomics Initiative of Indiana University is supported in part by Lilly Endowment, Inc."

*IBM Shared University Research grants. Support from IBM, Inc. in the form of Shared University Research grants from IBM, Inc. to Indiana University is gratefully acknowledged, as is IU's relationship with IBM as an IBM Life Sciences Institute of Innovation.

*IBM Shared University Research grants. Support from IBM, Inc. in the form of Shared University Research grants from IBM, Inc. to Purdue University is gratefully acknowledged, as is Purdue's relationship with IBM for the Network for Computational Nanotechnology.

*Sun Center of Excellence. Support from Sun Microsystems, Inc. via its Sun Center of Excellence is gratefully acknowledged.

Nota Bene: Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

About the Symposium

I-Light Symposium March 9, 2004

In 2003 legislators approved a plan for I-Light2, an expanded fiber optic network for the State of Indiana. On March 9, 2004 researchers gathered at the IUPUI Conference Center in Indianapolis to discuss how the original I-Light network, launched in December 2001, had impacted their work and to explore with their colleagues from institutions throughout the state the opportunities offered by the expanded network.

This was the second time that groups from throughout Indiana came together to discuss the uses and impact of the unique I-Light high-speed fiber optic network. The original network connects Indiana University Bloomington, Indiana University-Purdue University Indianapolis, and Purdue University West Lafayette with each other and with Abilene, the national high-speed Internet2 research and education network. I-Light2 will make this connectivity available to other universities, colleges, libraries, and related institutions around the state.

The Symposium brought together researchers from across the State of Indiana. Presentations and poster sessions filled the day, and special insights into the future were provided by the keynote and capstone speakers: Miron Livny, University of Wisconsin, and Donna Cox, National Center for Supercomputing Applications. Livny's discussion of high–throughput computing and Cox's examples of her internationally–acclaimed work in scientific visualization contributed significantly to the creative energy of the Symposium.

Building on the foundation provided by I-Light and the increased linkages available through I-Light2, Indiana can rightly claim to be at the crossroads of the Internet.

James Bottum
Vice President for Information Technology and CIO
Purdue University

Michael A. McRobbie
Vice President for Research and Information Technology
Indiana University

Description of I-Light

What is I-Light?

I-Light is a very high speed optical fiber network that links together the flagship campuses of Indiana University (in Bloomington) and Purdue University (in West Lafayette) by way of Indiana University-Purdue University Indianapolis. I-Light also hooks all three campuses into the national Internet infrastructure, including Internet2.

This vital artery running between the three main research campuses of the state's public universities connects many of Indiana's finest scholars, scientists, and researchers to one another and to the world via data, voice, and video.

Project Background

I-Light's roots trace back to discussions held in 1998; three years and a $5.3 million state appropriation later, Governor Frank O'Bannon symbolically launched I-Light in December 2001, making Indiana the first state in the nation to have such a network fully operational.

The two universities own and manage the optical fiber network in collaboration, with each responsible for its respective connection to IUPUI. A steering committee with representatives from Indiana's Intelenet Commission, IU, and Purdue pioneered I-Light's implementation.

What is the capacity of I-Light?

Before I-Light, Purdue and IU were limited in network capacity. The previous data access speed between Bloomington and West Lafayette was 30 million bits per second. I-Light increases access speed initially to 1 billion bits per second and is expandable hundreds of times over.

While the I-Light system is orders of magnitude faster than the standard copper-wire-based networking technologies it replaces, a bigger issue is the increased volume of information scientists and researchers will now be able to exchange. I-Light is capable of moving data equivalent to the entire written contents of either university's library from one campus to the other in seconds or to other universities nationwide through Internet2.

The I-Light Advantage

University ownership of the optical fiber infrastructure is a key advantage of I-Light. It should easily provide enough networking capacity for the next 10 to 20 years between the constituent campuses and the national optical fiber infrastructure. This long-term infrastructure investment—made by the state in good economic times—will help retain and strengthen Indiana's advantages in information technology in the future.

By significantly reducing digital barriers, I-Light has ushered in a new age of collaboration. Via I-Light, IU and Purdue have pooled their high-end computational resources in initiatives such as the Indiana Virtual Machine Room, the first university supercomputing grid to surpass the teraflop level of computation.

I-Light presents countless possibilities for collaborative research and an unparalleled platform for distance education. Moreover, with I-Light, IU and Purdue will have greater leverage and potential for federal grants and can help Indiana become a more substantial player in the information economy.

What does the future hold for I-Light?

Future expansion of the network could see it linked to important regional fiber structures such as the Illinois I-Wire initiative, Michigan's Merit Network, and Ohio's OARNet. Few other states have Indiana's geographical advantage when it comes to tapping into existing fiber pathways/crossroads. The result will be an optical fiber network fabric that will allow the institutions to engage in computing grids, share resources, and position IU and Purdue faculty more competitively for federal research grants and other opportunities.

Meeting Summary and Description

James Bottum, Vice President for Information Technology and CIO at Purdue University, welcomed attendees to the I-Light Symposium, the second of its kind, and thanked the State of Indiana, and the two universities that made the I-Light network possible. Vice President Bottum described the importance of this I-Light network and pointed out that this event provides an occasion for Indiana University Bloomington, Indiana University–Purdue University Indianapolis, and Purdue University, to explore opportunities for collaboration, research, and scholarship using the L-Light Optical Fiber Infrastructure. This network also allows these institutions to connect to Abilene, the national high-speed Internet2 research and education network. Vice President Bottum also pointed out that soon I-Light2 will link more educational institutions to this important network and extensively broaden collaborative opportunities.

Dr. Miron Livny was the keynote speaker of this event. Dr. Livny is a professor in the Computer Sciences department at the University of Wisconsin-Madison. Dr. Livny's research focuses on distributed processing and data management systems and data visualization environments. At the University of Wisconsin Dr. Livny is currently leading the Condor project. His recent work includes the Condor high throughput computing system, the DEVise data visualization and exploration environment and the ZOO scientific database management framework. The topic of Dr. Livny's keynote address was Data placement in widely distributed systems.

The largest portion of the day was devoted to an impressive array of presentations by researchers and information technologists from Indiana University Bloomington, Indiana University–Purdue University Indianapolis, and Purdue University West Lafayette. A great diversity of disciplines was represented in the presentations. The strength of the research done at all three research campuses was evident from the presentations. Several mid-morning and afternoon break demonstrations and poster sessions by researchers from all three campuses indicated how I-Light has facilitated their research.

Dr. Donna Cox delivered the Lunch and Capstone presentation. Dr. Cox is an international pioneer in scientific visualization and computer art. She is associate director for Experimental Technologies at the National Center for Supercomputing Applications (NCSA) and professor in the School of Arts and Design at University of Illinois, Urbana-Champaign (UIUC). Dr. Cox's presentation topic was Beyond Computing: The Search for Creativity.

Dr. Gary R. Bertoline, Associate Vice President for Visualization Computing at Purdue University closed the 2004 I-Light Symposium by thanking all of the participants for their contributions to make the Symposium a success. He emphasized the importance of such gatherings to share information by researchers from all disciplines, and concluded that I-Light has changed the way research is done at these institutions.

Sessions Conveners for the 2004 I-Light Symposium

Karen K. Whitney
Research Coordinator, Information Technology at Purdue
Purdue University

John V. Samuel
Director, Center for Statistical and Mathematical Computing
Indiana University

Program

8:00 a.m. Registration and Check-in
IUPUI University Place Ballroom

8:30 a.m. Welcome
University Place Conference Center
Auditorium

James Bottum
Vice President for Information Technology
and CIO, Purdue University, West Lafayette

8:45 a.m. Keynote Address
Data Placement in Widely Distributed Systems
Miron Livny
University of Wisconsin

9:15 a.m. *I-Light and I-Light 2 Updates*
Brian Voss
Associate Vice President for
Telecommunications, Indiana University

9:30 a.m. Demonstrations and Refreshments
University Place Conference Center
Breezeway

Advances in Parallel Metacomputing of Solid-Fluid Interaction (SFI) Problems
Hasan Akay
Amit S. Baddi
Resat U. Payli
Indiana University–Purdue University,
Indianapolis

Introduction to Virtual Environments – A Collaborative Classroom Offering between PU and IU
Laura Arns
Purdue University, West Lafayette

Live and Computational Experimentation in Bio-Terrorism Response
Tejas Bhatt
Purdue University, West Lafayette

Ingestion, Analysis and Distribution of Real-time and Archival Satellite Data for Indiana: The Purdue Terrestrial Observatory, the Laboratory for Analysis of Remote Sensing and IndianaView
Larry Biehl
Purdue University, West Lafayette

Variations2 Digital Music Library
Jon Dunn
Mark Notess
Indiana University, Bloomington

Building Community Grids: Collaboration and Distributed Computing
Geoffrey C. Fox
Indiana University, Bloomington

Using I-Light and the Purdue Nanohub Computing Resources to Run Computationally Intensive Codes in Nanotechnologies
Sebastien Goasguen
Purdue University, West Lafayette

Computer Simulation of Neutron Rich Matter Using the MDGRAPE-2
Chuck Horowitz
Donald K. Berry
Indiana University, Bloomington

Conducting Remote Sensing Seminars across the United States
Chris J. Johannsen
Purdue University, West Lafayette

Large-Scale Distributed Rendering of 3D Animation and Imagery
W. Scott Meador
Purdue University, West Lafayette

iVDGL Grid Operations Center and I-Light
Rob Quick
Indiana University–Purdue University
Indianapolis

Digital Library Resources Over I-Light
John Walsh
Indiana University, Bloomington

10:00 a.m. Presentations
University Place Conference Center
Auditorium

The MxN Problem: Coupling Parallel Components
Randall Bramley
Felipe Bertrand
Indiana University, Bloomington

10:20 a.m. *Introduction to Virtual Environments – A Collaborative Classroom Offering between PU and IU*
Laura Arns
Purdue University, West Lafayette

10:40 a.m. *Science Education Using Networked IT*
Randy Heiland
Indiana University–Purdue University Indianapolis

11:00 a.m. *The Biocomplexity Institute at IU–Combining Experimental and Computational Systems Biology*
James Glazier
Indiana University, Bloomington

11:20 a.m. *Conducting Remote Sensing Seminars across the United States*
Chris J. Johannsen
Purdue University, West Lafayette

11:40 a.m. *A Global Grid Analysis of Invertebrate Evolution*
Craig Stewart
Indiana University, Bloomington

12 noon - 1:30 p.m.
Lunch and Capstone Presentation
University Place Conference Center Hotel - Bistro (2nd floor, just above Hotel Lobby)

Beyond Computing: the Search for Creativity
Donna Cox
National Center for Supercomputing Applications (NCSA)

1:30 p.m. Presentations continued
University Place Conference Center Auditorium

An Integrated Environmental Monitoring Network
Lenore Tedesco
Pauline Baker
Indiana University–Purdue University, Indianapolis

1:50 p.m. *Using I-Light and the Purdue Nanohub Computing Resources to Run Computationally IntensiCodes in Nanotechnologies*
Sebastien Goasguen
Purdue University, West Lafayette

2:10 p.m. *Live and Computational Experimentation in Bioterrorism Response*
Tejas Bhatt
Purdue University, West Lafayette

2:30 p.m. *Demos and Refreshment Break*
University Place Conference Center
Auditorium Lobby and Breezeway

3:00 p.m. Presentations continued
University Place Conference Center Auditorium

Security Challenges for the Future
Brian King
Indiana University–Purdue University, Indianapolis

3:20 p.m. *Ingestion, Analysis and Distribution of Real-time and Archival Satellite Data for Indiana: The Purdue Terrestrial Observatory, the Laboratory for Analysis of Remote Sensing and IndianaView*
Gilbert L. Rochon
Purdue University, West Lafayette

3:40 p.m. *Large-Scale Distributed Rendering of 3D Animation and Imagery*
W. Scott Meador
Purdue University, West Lafayette

4:00 p.m. *Methods for Network-based Graphics and Visualization*
Eric Wernert
Indiana University

4:20 p.m. Closing Remarks
Gary Bertoline
Associate Vice President of Information Technology, Purdue University, West Lafayette

Keynote Speakers

Miron Livny

University of Wisconsin–Madison

Miron Livny received a B.Sc. degree in Physics and Mathematics in 1975 from the Hebrew University and M.Sc. and Ph.D. degrees in Computer Science from the Weizmann Institute of Science in 1978 and 1984, respectively. Since 1983 he has been on the Computer Sciences Department faculty at the University of Wisconsin-Madison, where he is currently a Professor of Computer Sciences and is leading the Condor project.

Dr. Livny's research focuses on distributed processing and data management systems and data visualization environments. His recent work includes the Condor high throughput computing system, the DEVise data visualization and exploration environment and the ZOO scientific database management framework.

Donna Cox

National Center for Supercomputing Applications (NCSA)

Donna Cox is Associate Director for Experimental Technologies at the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Urbana-Champaign (UIUC). A renowned expert on computer visualization, Cox also is a professor in the School of Art and Design at the University of Illinois. She has authored many papers and monographs on computer graphics, information design, education, and scientific visualization and has exhibited computer images and animations in more than 100 invitational and juried exhibits. Her animations of scientific data have appeared on international television, including episodes of *NOVA*, Discovery Channel documentaries, CNN, and *NBC* Nightly News.

In 1997, Cox and a coworker earned an Oscar nomination for their work on the IMAX film *Cosmic Voyage.* Cox's team developed data-driven scientific visualizations for the HDTV Nova/WGBH show "Runaway Universe," which received the 2002 Golden Camera, International Film and Video Festival award.

The MxN Problem: Coupling Parallel Components

Randall Bramley and Felipe Bertrand

The original vision of the Grid was that of a platform for running a scientific simulation consisting of large numbers (on the order of millions) of parallel communicating processes. That vision has since evolved to one of the Grid as a fabric for tying together heterogeneous resources: data servers (sensors, instruments, databases, and large archival storage), computation resources, and visualization systems. Increasingly important is the problem of connecting parallel programs, possibly running on geographically distributed machines, each using a different number of processes. This problem occurs in climate modeling, space weather systems, and magnetically–confined fusion energy simulations. In each case, large legacy parallel codes from different scientific communities are being connected to create new multidisciplinary simulations. This presentation describes the problems involved and introduces some middleware solutions for combining parallel programs even when they use incommensurate numbers of processors.

Introduction to Virtual Environments — A Collaborative Classroom Offering between Purdue and IU

Laura Arns

Purdue University, West Lafayette

The Access Grid (AG) is an emerging technology that allows for unique learning experiences. Currently, we are conducting a class titled “Introduction to Virtual Environments” (Purdue TECH 519V) via the AG in collaboration with Indiana University (IU Informatics I590). The class meets twice a week for lectures over the Access Grid.

Conducting a class via the Access Grid offers several advantages over traditional classroom instruction. Because the topic of the class (virtual reality) is relatively new to Purdue and is quite specialized, class size is rather small (approximately 10 students at Purdue and 6 at IU). Utilizing AG allows us to hold lectures that combine the students from both the Purdue and IU classes, creating a larger virtual class size. This allows both students and instructors to interact with a larger, more diverse, student population.

Conducting joint lectures also gives students the benefit of learning from two instructors with different backgrounds and experiences, providing more information and viewpoints than a single instructor alone could present. Additionally, the instructors can share the lecture load, allowing each to create better lectures and lab materials, and to spend more time interacting with students. Because AG technology is cheaply and easily accessible at a large number of locations, expert guest lecturers can also join the lecture sessions remotely, eliminating the time and expense that would normally be incurred, and allowing students to hear from leaders in the field that might not otherwise be accessible to them.

The topic of virtual reality presents unique challenges for traditional distance education methods such as videotaping and Web streaming of lectures. In particular, it is important that students not only learn about the technology of virtual reality, but also experience immersive and interactive virtual environments as part of their class activity.

Science Education using Networked IT

Randy Heiland
Indiana University–Purdue University Indianapolis

This presentation considers some projects and ideas for projects that benefit from networked information technology. As a relative newcomer to this endeavor (and to Indiana), the speaker welcomes and hopes to hear follow-on ideas from the audience.

The Biocomplexity Institute at IU — Combining Experimental and Computational Systems Biology

James Glazier
Indiana University, Bloomington

The Biocomplexity Institute at IU, in collaboration with Purdue and Notre Dame, addresses issues in post-human-genome-project interdisciplinary biosciences, particularly the integration of quantitative experiments and computation. One of our chief goals is to provide predictive multiscale models of the development, which the National Institutes of Health have identified as a priority for the next 10 years (see *http://nihroadmap.nih.gov*). This goal requires both large-scale simulations (currently based around the Cellular Potts Model (CPM)–*www.nd.edu/~lcls/compucell*) and the development of new experimental techniques. The College of Arts and Sciences and the School of Informatics at IU are greatly expanding in these areas, with five hires completed, five planned for this year, and similar numbers planned for the next two years. Collaborations include projects on vascular, heart and limb development and regeneration (at the IU School of Medicine and the School of Science at IUPUI), on neurobiology, microfluidics and biofilms (at IU), and on algorithm development (at Notre Dame).

Conducting Remote Sensing Seminars across the United States

Chris J. Johannsen
Purdue University, West Lafayette

In 1999, I developed a remote sensing seminar course to bring a focus on the latest information about remote sensing topics. It was very popular and, during the next year, we decided to offer the course to Indiana State and Mississippi State University at the same time via IHETS. This meant that I only needed to arrange for about one-third of the presentations and that we could learn about remote sensing work at other universities.

The purpose of the seminars is to provide a forum for graduate students to: 1) interact with academic, aerospace, and government personnel working with remote sensing, global positioning systems, and geographic information systems called "spatial technologies," 2) hear the latest information about spatial technologies and how they are being used, and 3) provide opportunities to explore future career paths involving remote sensing, GIS, and GPS technologies.

During 2001, 2002, and 2003, we added the University of Nebraska, the Delta Research and Education Center (Stoneville, MS), the NASA Stennis Space Center, the EROS Data Center, and the South Dakota School of Mines as participants. We also changed the approach and are using the Internet and Information Technology at Purdue (ITaP) to provide higher–quality seminars. Now we are able to ask remote sensing scientists from across the United States to present a seminar to our students by transmitting the seminar from their locations. ITaP then transmits the seminar to the other locations.

The non-university locations are participating to provide their professional staffs with the latest information on remote sensing. Each of the participating universities sign up students using their course registration, and students receive one hour credit for participation. The students earn their credit by attending the seminars and providing a five-minute presentation on their research with a limit of five PowerPoint slides. The quality of the seminars has improved every year with changes in the technology of presenting and transmitting the seminars.

During this current semester, Purdue is going international with this approach and we are offering a special remote sensing seminar jointly with the University of Leuven, Belgium. This is part of a joint Master of Science degree program on Earth Observation with Belgium.

A Global Grid Analysis of Invertebrate Evolution

Craig Stewart
Indiana University, Bloomington

In recent years the evolutionary relationships of invertebrates with six legs (insects and their relatives) have been a topic of considerable debate in the biological community. There are two steep challenges in pursuing this question: assembling enough genetic sequence data to answer the question effectively, and assembling the computer resources required to analyze the data. The Center for Genomics and Bioinformatics and IU, along with experts from University Information Technology Services (UITS), assembled a large data set of genetic data.

A team led by University Information Technology Services and the High Performance Computing Center of Stuttgart lead an international effort that created a global grid to analyze this data. More than 600 computer processors were applied to the task of analyzing this data within a global grid spanning every continent on Earth except Antarctica. Indiana University supercomputers and Teragrid resources were key components of this grid.

This talk describes this study, our conclusions to date, and also discusses the future of biological computing and computing grids. This project was awarded a High Performance Computing Challenge award at the international ACM/IEEE SuperComputing 2003 conference held in November 2003.

Beyond Computing: The Search for Creativity

Donna Cox
National Center for Supercomputing Applications (NCSA)

Modern computing arose from the interplay of science, engineering and defense needs, and hardware and software technology advances. Computational science began, like most science, as a small and localized group activity. These islands of research are increasingly connected — the truly grand challenges require the skills of multidisciplinary groups, often internationally distributed, working collaboratively as global renaissance teams. Transformative collaborative, networked technologies, coupled with high-end computing and large-scale data archives, enable success through enhanced human interaction, community building, collective problem solving and innovation.

Recent studies on human innovation span the sciences and humanities, and they point to a paradigm shift in the way that we think about creativity, they demonstrate the power of understanding one knowledge domain in terms of another, and they reveal the contingencies of our perspectives. Fresh perspectives result from cross-domain interactions through common themes.

These common technological themes across the sciences and humanities include high-performance data collection, retrieval and integration, text and image data mining, software fusion, visualization, collaborative tools and human-computer interfaces. Enlarging the scope of high-performance computing to include a broad range of disciplines will enrich discovery and expand the applications of computing technology.

Both creativity and courage will be necessary to address the really big questions: understanding the matter and structure of our universe; modeling life and its complex processes; and enriching the human condition. From mapping the cosmos to mapping the brain; from interactive sensor databases to real-time severe weather prediction; from bioinformatics to situational awareness, “Thinking out of the Box” requires interdisciplinary, diverse and global interactions. How we enable these creative practices will shape the future of our HPC community.

An Integrated Environmental Monitoring Network

Lenore Tedesco and Pauline Baker

Ecological studies depend on the ability to monitor an environment, collect data at appropriate spatial and temporal scales, and analyze that data from the diverse viewpoints of many relevant disciplines. Historically, environmental studies have been conducted by small teams of researchers, usually hand-collecting data at set but low frequency, and organizing it according to ad hoc, project-specific goals. Recent years have seen dramatic advancement in the ability to gather environmental data remotely and therefore at much higher frequency.

We are working to create a dynamic and integrated network of environmental sensors in natural environments to acquire real–time data and to create tools for visualization appropriate for different audiences to advance both ecological research and educational exploration. Visualization of real–time data from remote sensors distributed throughout Central Indiana provides numerous challenges. The benefits of successfully integrating remotely–deployed environmental sensors in a post–9/11 world are obvious.

Our work will bridge both the extremes associated with the frequency of data collection and the lack of data coordination by creating techniques for data networking and retrieval, as well as data management and analysis. We are working to integrate multiple data streams into a coherent database and to create applications that allow users to view data from multiple instruments at different sites.
A key outcome of this program is to create visualizations of real–time, dynamic data from the everyday world, and deliver it via Web applications and innovative display spaces. These visualization capabilities need to operate across a range of computing platforms to make this data immediately accessible and useful to a range of interested parties across multiple disciplines.

Our goal is to use the instrumented sites to create analysis and presentation applications to foster a community of learners interested in understanding these complex ecosystems, and the larger environmental issues that they represent. This broad-based community will include environmental researchers, university faculty in lecture halls, K-12 math and science teachers, university and K-12 students, civic leaders, and educators at informal learning centers.

Using I-Light and the Purdue Nanohub Computing Resources to Run Computationally-IntensiveCodes in Nanotechnologies

Sebastien Goasguen
Purdue University, West Lafayette

The NSF-funded Network for Computational Nanotechnology (NCN) has a dual mission of research and providing computing services for the nanotechnology community. The research performed at the NCN is focusing on three areas: Nanoelectronics, nanoelectromechanical systems (NEMS), and nano-bio. In each of the research themes, projects are defined that are ready for a coordinated approach from a team of experts that can tackle the atomistic level, the device level, and up to the system level.

These computational challenges create useful computational tools that are made available to the community through the NCN computing portal, the Nanohub. The Nanohub employs a middleware infrastructure that allows users to run simulations seamlessly without knowing where and how the program is running. The front end is a standard Web site, but the back end is made of workstations, the SMP machine, and Linux clusters. Linking the Nanohub to I-Light and the Teragrid is of prime importance to the NCN in order to make available simulators that require a considerable amount of computing power to run in parallel.

Live and Computational Experimentation in Bio-Terrorism Response

Tejas Bhatt
Purdue University, West Lafayette

The study of bio-terror threats requires a significant improvement in our capabilities to analyze human responses to an attack, both at the level of citizens (individual freedom of mobility, infection rate, and the collective feeling of well-being) and at the level of responders and policy makers (coordination, control, planning, and policy formulation).

We address these issues using a geography-based synthetic environment with artificial and human agents, developing computational models of artificial agents' positions, mobility, infection-susceptibility and the state of well-being. Intuitive interfaces are provided for human agents to experiment with complex coordination roles. Together, their behaviors are used to analyze how a bio-terror attack may spread through the population and how its impact may be mitigated by different intervention strategies. The behavior of the artificial agents is calibrated in accordance with standard models from the fields of epidemiology, psychology, and economics, and the agents' movements reflect the actual behavior patterns in the cities concerned. The representation of human behavior in artificial agents reflects human capabilities, cognitive processes, limitations, and conditions that influence behavior (e.g., morale, stress, panic). We use individual-based epidemiological models for person-to-person contamination scenarios. An important aspect of this work is the development of a scalable architecture for distributed tera-scale grid computing.

Our results indicate that if the intervention is done early, city-block vaccination is most effective; in the intermediate term-trace vaccination will be most effective; and if the response is delayed, then mass vaccination is the best strategy. In terms of quarantine strategies, early extreme quarantine is more effective than the city block quarantine. However, for delayed intervention, there is no significant difference between the extreme and city block quarantine strategies.

Security Challenges for the Future

Brian King
Indiana University–Purdue University, Indianapolis

Network technology advancements such as I-Light have provided greater network connectivity and increased bandwidth. This greater connectivity, together with advancements in hardware technology, will bring us to a point where we will be able to compute on demand, utilizing whatever computing devices are available, no matter how limited. Computing will be everywhere. In many cases, this computing will be unseen and invisible.

As these advancements bring small lightweight devices to the computing mainstream, new applications that were once foreign will be unleashed on these devices. Consequently security must be addressed.

In this presentation we will survey some of the more challenging problems that are being examined in our Information Security Lab.

Ingestion, Analysis, and Distribution of Real-time and Archival Satellite Data for Indiana: The Purdue Terrestrial Observatory, the Laboratory for Analysis of Remote Sensing, and IndianaView

Gilbert L. Rochon
Purdue University, West Lafayette

This presentation will discuss the imminent acquisition in realtime of panchromatic, multi-spectral, radar and, potentially, hyper-spectral data by the Purdue Terrestrial Observatory (PTO). Plans are detailed for near-real-time analysis of both archival and newly-acquired remotely-sensed data through Purdue's Laboratory for Applications of Remote Sensing (LARS), soon celebrating its 40th anniversary, and the distribution of such data, facilitated by I-Light, I-Light2, the National Science Foundation (NSF) Extensible Teragrid Facility (ETF), and by the Web-enabled IndianaView affiliate of the U.S. Geological Survey (USGS)-supported AmericaView Program. Opportunities for collaborative multidisciplinary research, distributed instruction, economic development, and community outreach utilizing these technological resources also will be delineated.

Large-Scale Distributed Rendering of 3D Animation and Imagery

W. Scott Meador
Purdue University, West Lafayette

The rendering of 3D animation and imagery is often a computationally intensive task. Rendering an entire animation may take many days to finish and with the standard refinement process to correct mistakes or make changes, many more days are needed to get the final results.

Distributing the rendering process to all available computers can dramatically decrease the time it takes to produce the finished animation.

Purdue University's Distributed Rendering Environment (DRE) initiative uses as many computers as are available to process 3D animation. In order to grow the number of available computers and extend the functionality to other campuses in Indiana, IU and IUPUI have begun to use and help develop the system over I-Light.

Methods for Network-Based Graphics and Visualization

Eric Wernert
Indiana University, Bloomington

Recent increases in network bandwidths and graphics capabilities promise significant improvements in the way that researchers, artists, educators, and students are able to visually interact with their data, concepts, presentations, and colleagues. We will provide a brief survey of methodologies for harnessing graphics resources over a network, focusing on visual telecollaboration and remote visualization and rendering. We will illustrate these methods with a number of projects carried out in conjunction with IU's Advanced Visualization Lab, and we will describe potential enhancements to the research, education, and creative activities at Indiana University.

Network-based Visualization and Visual Collaboration

iLight 2004 Workshop

Eric A. Wernert
UITS Advanced Visualization Lab
Indiana University

UITS Advanced Visualization Lab

- Unit of University Information Technology Services at Indiana University
  - Part of Research & Academic Computing division
- Labs in Bloomington and Indianapolis
  - 3 staff members at IUB
  - 4 staff members at IUPUI
  - Operate as single cross-campus unit
- Actively expanding to select regional campuses

AVL Mission

… to provide…
consulting • development • hardware & software resources
…in the areas of…
visualization • virtual reality • high-end graphics • visual collaboration
… to support the…
research • education • creative activities
…missions of IU across…
all departments • all campuses

AVL Resources

John-e-Box Portable Stereo Display

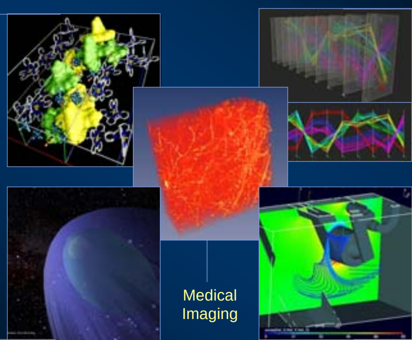
Tiled Displays & Clusters

CAVE Immersive VR

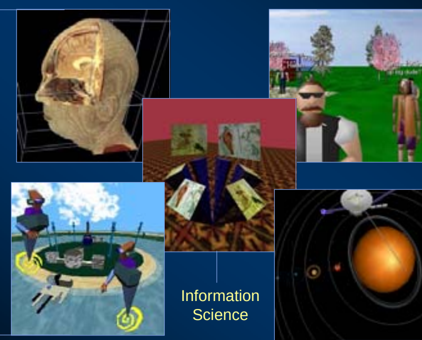
Haptics & Multi-modal Interaction

Tele-Collaboration

AVL – Support for Research



AVL – Support for Education



Network-based Visualization and Visual Collaboration • iLight Symposium 2004
Eric A. Wernert • UITS Advanced Visualization Lab • Indiana University • 3/9/04

AVL – Support for Creative Activities

Network-based Visualization and Visual Collaboration • iLight Symposium 2004
Eric A. Wernert • UITS Advanced Visualization Lab • Indiana University • 3/9/04

Presentation Objectives

- To describe the range of methods for interactive, networked visualization and collaboration
- To explain some specific, ongoing projects and investigations at AVL, IU, and elsewhere
- To discuss our vision for an ideal visualization infrastructure

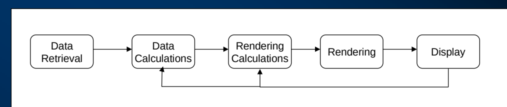
Ultimate Goal: To help stimulate your thinking on how you can apply networked visualization and visual collaboration techniques to benefit your work

Network-based Visualization and Visual Collaboration • iLight Symposium 2004
Eric A. Wernert • UITS Advanced Visualization Lab • Indiana University • 3/9/04

Challenges of Networked Visualization & Visual Collaboration

- diversity of users and applications
- diversity of advanced displays
- developing a compelling remote visualization resource
- interfacing with other high-performance services: computing, storage, networking
- demands for "interactivity"
- location of user - last mile problem

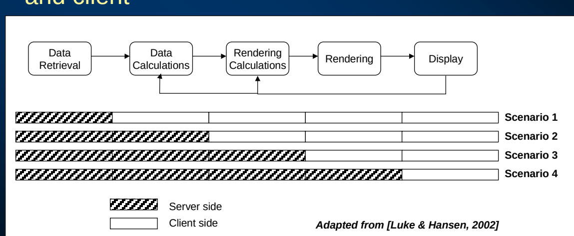
The Visualization Pipeline



1. Data access / retrieval
2. Data calculations (filtering and conversion)
3. Rendering calculations (data mapping)
4. Rendering
5. Display

Network-based Visualization

- 4 logical places to divide the work between server and client



Network-based Visualization Scenarios

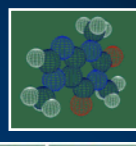
- Scenario 1/2 - Data transfer
  - requires significant local storage, computation & rendering
- Scenario 2/3 - Graphics command transfer
  - computation & rendering load shared between systems
- Scenario 4 - Image transfer (includes video streaming)
  - storage, computation, and rendering load remain on server; good for lightweight clients

- *Must consider data complexity, visual quality, display requirements, and bandwidth availability to determine best method for any given application*

Network-based Visualization - Example

- Data Transfer
  - `atom 1 (type, location), atom 2,…`

- Graphics Commands Transfer
  - `color(red), sphere(radius 0.57), …`



- Image Transfer
  - `image 1, image 2, image 3, …,`

Interactivity Requirements

- What is "interactive"? - It depends on the task:

30 FPS (frames per second)	for simulation & training
10 FPS	for manipulation
1-3 FPS	for passive observing
30-60 seconds per frame	for complex data analysis (within the span of human attention)
5-30 minutes	for complex image or simple movie preview
2 hours – 1 day	for complex movie preview or batch rendering

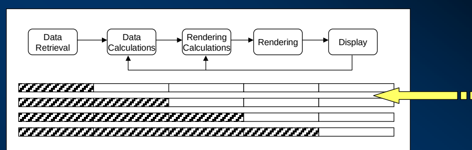
Display Requirements

	Monoscopic	Stereoscopic
High Resolution	• Tiled display wall • IBM T221 display	• Tiled stereo wall • CAVE
Standard Resolution	• Desktop displays • Web, PDAs	• John-e-Box • Desktop stereo

Bandwidth Requirements (for Image Transfer)

- **John-e-Box**
  - 1024x768 pixels x 24bits x 2 (stereo) x 10 FPS = 188 Mbps
- **IBM T221 (9.2 megapixel display)**
  - 3840x2400 pixels x 24bits x 3 FPS = 664 Mbps
- **6-Tile SXGA Display Wall**
  - 1280x1024 x 24bits x 6 tiles x 10 FPS = 1.88 Gbps
- **CAVE**
  - 1280x1024 x 24bits x 4 walls x 2 (stereo) x 30 FPS = 7.55 Gbps

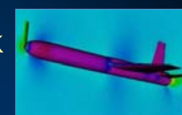
Data Transfer – Techniques



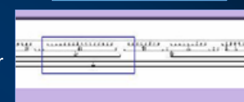
- Good: download and visualize
- Better: network-based querying/ data mining
- Better Still: querying and visualization over real-time data streams

Data Transfer – Examples

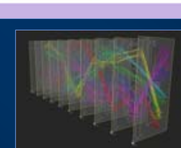
- Good: FEA visualization w/VTK
  - data download to workstation



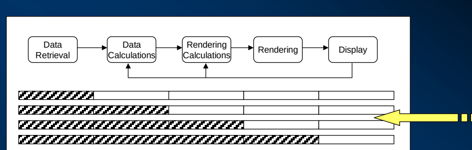
- Better: Pedigree Tree Vis
  - data queries to Access server



- Better Still: Phylogenetic Vis
  - Real-time data from running computation



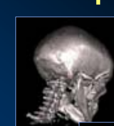
Graphics Command Transfer – Techniques



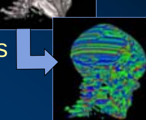
- Good: remote graphics display (e.g., Xwindows)
- Better: geometry optimization methods
- Better still: application-transparent streaming and optimization

Graphics Command Transfer – Examples

- Good: remote OpenGL via GLX
  - works with most applications

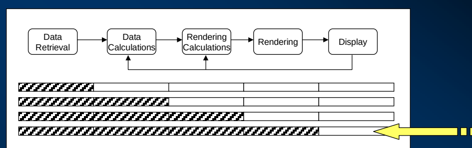


- Better: Marching Cubes optimizations
  - Minimize number of vertices transferred (Lakshmipathy, et. al.)



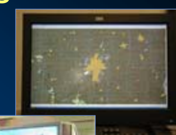
- Better still: Chromium tiled rendering
  - Application-transparent sorting, caching, and state optimizations

Image Transfer – Techniques



- Good: individual (monoscopic) images
- Better: stereo images
- Better still: composited or tiled images

Image Transfer – Examples

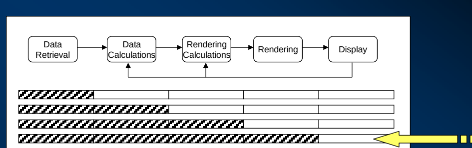


- Good: VNC w/ GIS applications
  – point-to-point and multipoint collaboration

- Better: VNC for stereo
  – shared side-by-side stereo applications on John-e-Box

- Better Still: Compositing with VTK or Chromium
  – Parallel rendering provides scalability of data, rendering quality, and speed

Video Streaming – Techniques



- Good: point-to-point streaming
- Better: multicast streaming
- Better still: multiple, synchronized visualization streams

Video Streaming – Examples

- Good: Polycom
  – Point-to-point or 4-way MCU

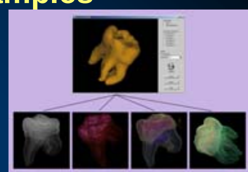
- Better: Multicast Streaming
  – Access Grid
  – AVL “virtual distributed lab”

- Better still: synchronized streams
  – Chromium H262 SPU (Argonne)
  – TeraVision (EVL @ UIC)

Image from EVL TeraVision web site

Hybrid Methods – Examples

- AVL's "ToothPics" Framework
  - lightweight client to establish viewpoint
  - High-quality, server-based rendering
  - Image transfer to client

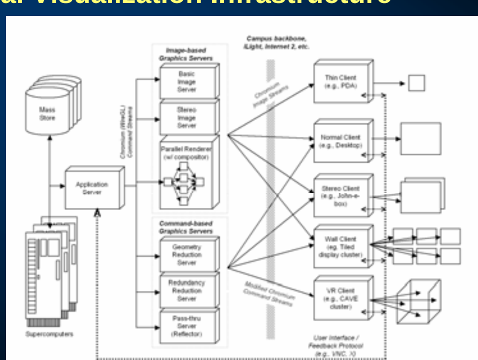


- OpenDX, Paraview, Visit
  - explicit distribution of the visualization pipeline
  - Potential for parallelization at most steps
  - Support for local or remote rendering

Desirable Traits of Networked Visualization

- *Scalable* – number of users, size of data, size/quality of visualization
- *Powerful* – provide a resource an order of magnitude better than desktop systems
- *Flexible* – one central visualization computing resource utilized in many different ways
- *Efficient* – makes best use of network and local compute resources
- *Simple* – for users and developers
- *Application-Transparent (Translucent)* – requires no (minimal) modifications to applications

Ideal Visualization Infrastructure



Conclusions

Network-based visualization and visual collaboration:

- Many techniques are currently possible
- Newer, better techniques are being developed
- The best technique depends on a number of factors
- Consult with the AVL at IU (or Envision Center at Purdue) if you have a visualization or visual collaboration problem or opportunity

More Information & Acknowledgements

For more information, please visit

http://www.avl.iu.edu

Acknowledgements:

- AVL staff & clients
- Office for the VP for Information Technology at IU
- UITS colleagues
- Indiana Genomics Initiative (INGEN)
- NSF/MRI grant (AVIDD)
- Researchers at other institutions

I-Light & I-Light2 Updates

Brian D. Voss

Associate Vice President for Telecommunications
Office of the Vice President for Information Technology & CIO
Indiana University
&
Chief Operating Officer
Pervasive Technology Labs @ Indiana University

I - L I G H T

I-Light Success – Fostering Research

Providing the foundation for innovation

I-Light makes possible extensive growth in research across a wide variety of applications and areas.

This increases Indiana's involvement in national and international research activities

Educational Advances

Visualizing Complex Data

Homeland Security

Virtual Reality

I-Light's Continuing Impact

I - L I G H T

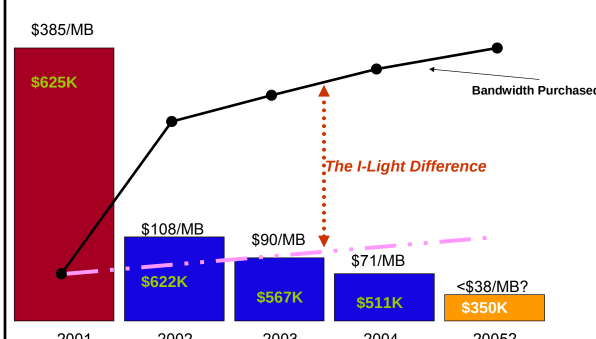
- **Intercampus bandwidth remains completely unconstrained**
  - The Doritos Principle of bandwidth scarcity – *Go ahead and crunch it … we'll just light more!*
- **Positioning with co-location and our Quilt membership (as a GigaPoP) has allowed us to negotiate even better commodity internet rates (dropped 10%! … may drop another 10% next year)**
- **Eliminates barriers to collaboration**
  - Witness the many examples in this year's program

✓ **Educause 2003 Award for Best Practices in Networking**

2003 EDUCAUSE Award for Excellence in Networking WINNER

Costs for Internet Usage

(IU as an example)



- **What's Next?**

  - Connecting to Chicago … ***the Holy Grail***
  - Connecting to National Cyberinfrastructure
    - TeraGrid
    - National LambdaRail

NSF's Extensible Terascale Facility

a.k.a. The TeraGrid

TERAGRID

I-LIGHT IP•grid TERAGRID

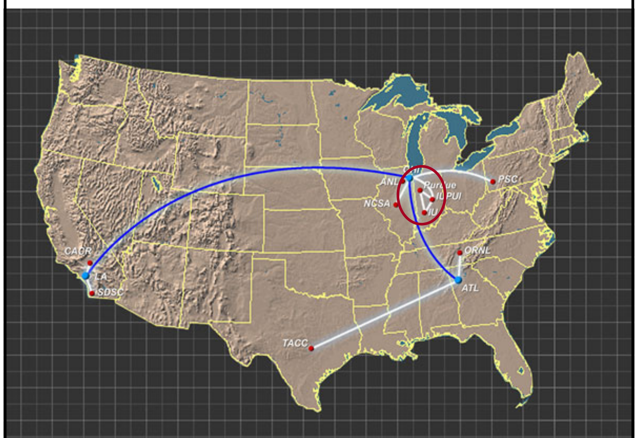
IU and Purdue wrote a proposal in June 2003 to join the TeraGrid Community in response to an NSF call for proposals

- Current members included NCSA, Argonne National Labs, SanDiego Supercomputing Center, CalTech, and Pittsburgh Supercomputing Center
- Quid-pro-quo – we contribute computation, storage, and data resources and NSF funds our connectivity to the nearest 'hub' …

IP•grid IP•grid IP•grid IP•grid IP•grid IP•grid

IU and Purdue formed IP-grid in our proposal to the NSF

- **We were chosen (with Oak Ridge National Labs and Texas Advanced Computing Center) to join this elite high-performance computational/ research community!**
- **NSF formal announcement 9/29 - $3-million in funding**
- **Expect our connection to be completed *THIS MONTH*!**
  - **We'll be the first of the new sites connected – by >5months!**



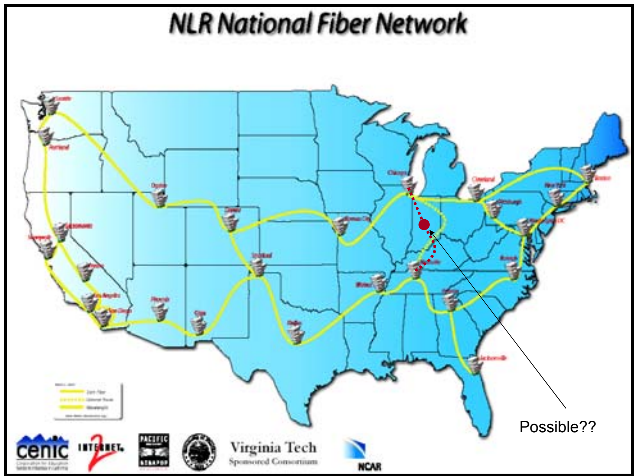
What does it mean?

- We have acquired a **pair of fibers** and run advanced dense-wave-division-multiplexing (DWDM) connectivity across it
  - Were the first customer of FiberCo – the new Internet2 subsidiary providing fiber across the nation for regional optical network development
- IU and Purdue will each have a **10-Gbit Wave** connecting us to the TeraGrid backplane
- We will have **co-location facilities** in a key carrier facility in downtown Chicago, as well as connection to StarLight
- We not only got $3-million for 2003-05, but also have a path to pick up ~$1-million *per year per institution* for the next four years after that!

What does ***that*** mean???

- **We will have the ability to add more waves to get other kinds of connectivity in Chicago**
  - Likely even cheaper commodity internet
- **We will have a network 'beach-head' in one of the largest network 'hubs' in the nation/world**
  - We may be able to overcome some peering issues with national high-speed internet services to homes
  - We can connect to others as the national cyberinfrastructure grows
- **We have managed to get Federal Grant money to fund a part of I-Light2**
  - *More on that in a moment …*

National LambdaRail (NLR)



What is NLR?

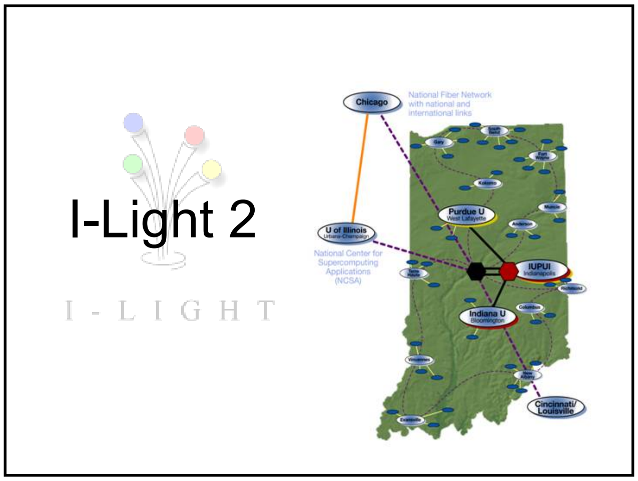
- Basically a larger, consortium-owned version of I-Light
  - Designed to facilitate point-to-point λ (10 Gbit) 'waves' (i.e., individual light-waves on a fiber)
  - Research focused
- A experimental network
  - A place to test new networking technologies

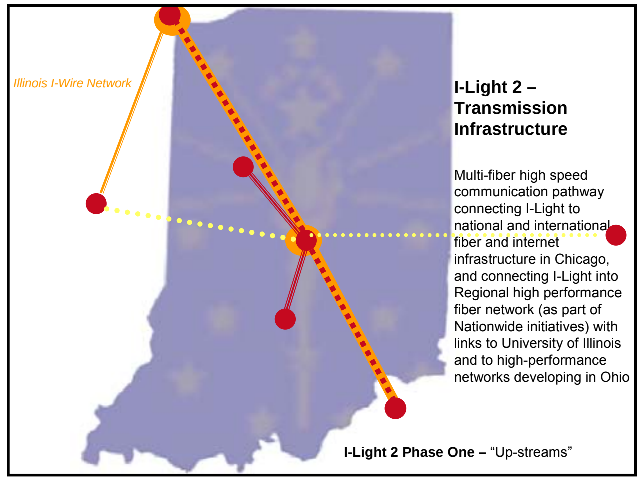
What is NLR ***not*** ?

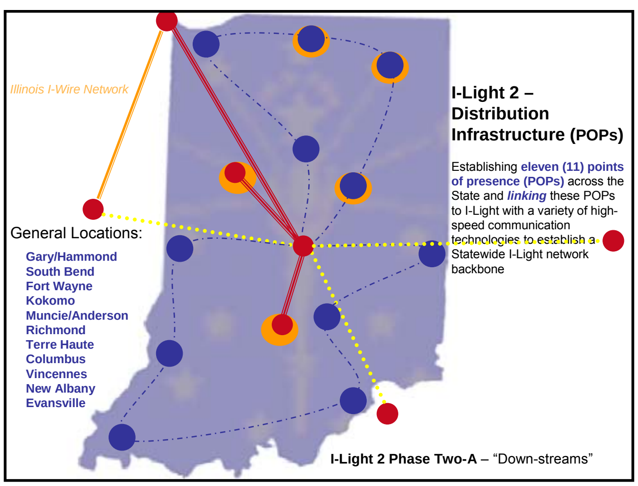
- ***NOT*** **A replacement for Internet2/Abilene**
  - UCAID is an active Partner
  - UCAID will provide it's own advanced services to I2 members using 2 λ's on the backbone
  - UCAID may eventually make use of NLR λ's for it's production Abilene traffic
- ***NOT*** **A new carrier**
  - Think of it as the Sam's Club of coast-to-coast high-speed bandwidth for research collaboration

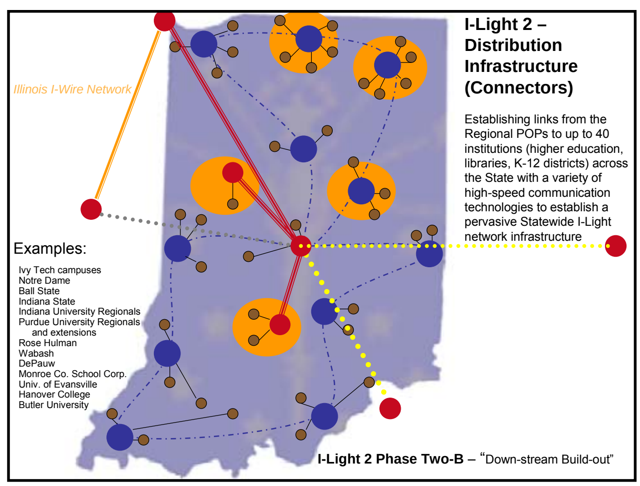
How are Purdue and IU involved?

- **Joined NLR as part of the CIC**
  - 13 institutions splitting the $5-mil investment (5-years)
- **IU led the CIC efforts to form the NLR**
- **IU was selected to provide NOC and L2/L3 Engineering Services to NLR**


I-Light 2
I - L I G H T
Chicago
National Fiber Network with national and international links
U of Illinois
National Center for Supercomputing Applications (NCSA)
Purdue U
IUPUI
Indiana U
Cincinnati/ Louisville


Illinois I-Wire Network
I-Light 2 – Transmission Infrastructure
Multi-fiber high speed communication pathway connecting I-Light to national and international fiber and internet infrastructure in Chicago, and connecting I-Light into Regional high performance fiber network (as part of Nationwide initiatives) with links to University of Illinois and to high-performance networks developing in Ohio
I-Light 2 Phase One – "Up-streams"


Illinois I-Wire Network
I-Light 2 – Distribution Infrastructure (POPs)
Establishing eleven (11) points of presence (POPs) across the State and linking these POPs to I-Light with a variety of high-speed communication technologies to establish a Statewide I-Light network backbone
General Locations:
Gary/Hammond
South Bend
Fort Wayne
Kokomo
Muncie/Anderson
Richmond
Terre Haute
Columbus
Vincennes
New Albany
Evansville
I-Light 2 Phase Two-A – "Down-streams"


Illinois I-Wire Network
I-Light 2 – Distribution Infrastructure (Connectors)
Establishing links from the Regional POPs to up to 40 institutions (higher education, libraries, K-12 districts) across the State with a variety of high-speed communication technologies to establish a pervasive Statewide I-Light network infrastructure
Examples:
Ivy Tech campuses
Notre Dame
Ball State
Indiana State
Indiana University Regionals
Purdue University Regionals and extensions
Rose Hulman
Wabash
DePauw
Monroe Co. School Corp.
Univ. of Evansville
Hanover College
Butler University
I-Light 2 Phase Two-B – "Down-stream Build-out"

What is I-Light2?

- Collaborative effort of a new Partnership:
    - State Budget Agency
    - Intelenet Commission
    - IHETS
    - Indiana Department of Commerce
    - Purdue University
    - Indiana University
- A way to "tap in" and use some of the miles of unused 'dark fiber' cables in place generating no revenue for the private sector owners and no benefit today to Indiana Communities
- A way to help connect regional networking projects and technology centers around the State

What Else Will I-Light2 Do?

- Act as an **'Anchor Tenant'** to lighting the dark fiber, lowering the cost of lighting more of it for providing broadband services in communities across the State.
- Help **eliminate the 'academic digital divide**' by providing high-speed access beyond just IU and Purdue to the rest of Indiana's higher education community, and K-12, Libraries, etc.
- **Enhance the research, education, and economic development** through the State and expand Indiana's growing reputation in advanced high-speed networking

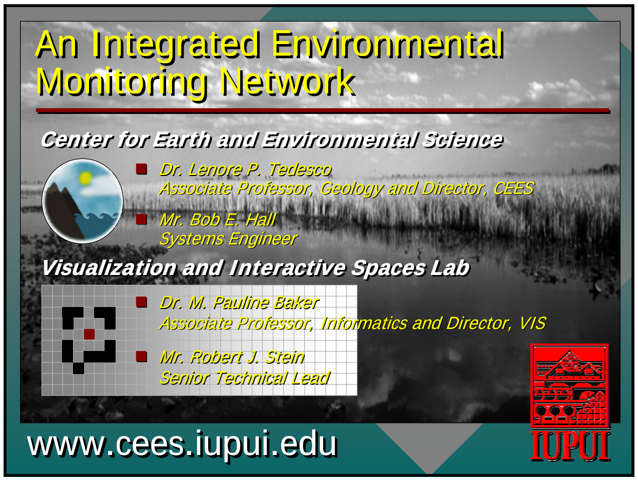
Project Status (March 2004)

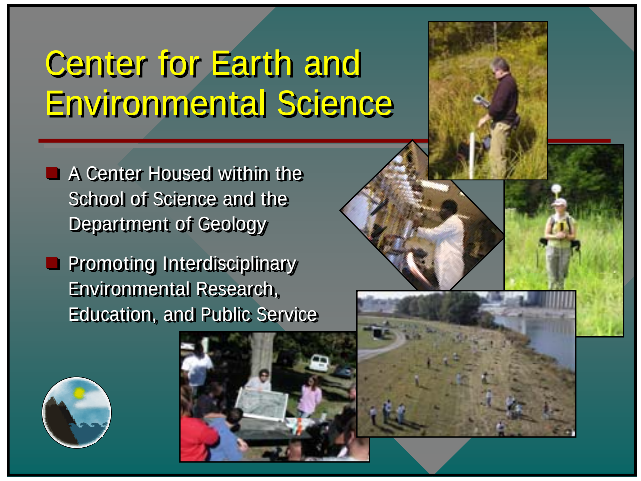
- **State appropriated $10-million in the 2003-04 biennium for the project**
    - Was estimated to cost $15-million -- $5-mil each phase
    - Here's where that TeraGrid grant really paid off!
        - Transmission Phase is essentially externally funded (though $1-mil held in reserve for extending/expanding these out-of-state connections)
- **RFP distributed to build in-state PoPs and acquire connectivity fiber paths**
    - RFP responses returned last month – in evaluation now by a select committee made up of people from the Universities, IHETS, Intelenet, Budget, and
- **Plans being discussed to "jump-start" I-Light2 by using I-Light's *State-Use Fibers* to establish PoPs in Bloomington and West Lafayette**

I-light – Watch it in 2004!

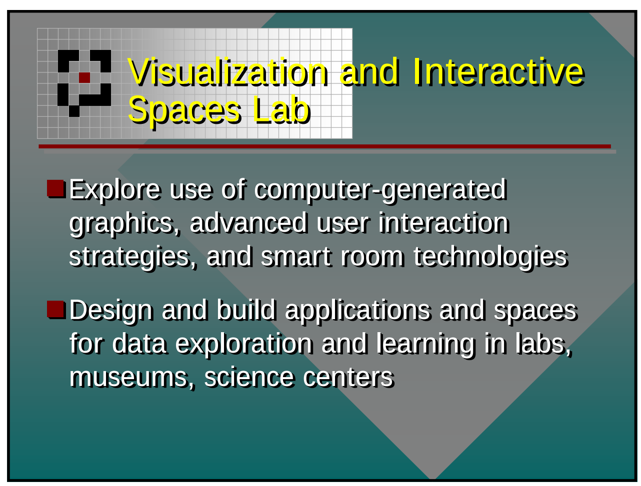
- Watch the IP-grid network come live and yield not only TeraGrid connectivity, but broader connectivity and services for I-Light
- Watch national cyberinfrastructure developments relating to Internet2 and NLR
- Watch I-Light2 Unfold across the State with the first PoPs being lit before the end of the year

I - L I G H T


An Integrated Environmental Monitoring Network
Center for Earth and Environmental Science
Dr. Lenore P. Tedesco
Associate Professor, Geology and Director, CEES
Mr. Bob E. Hall
Systems Engineer
Visualization and Interactive Spaces Lab
Dr. M. Pauline Baker
Associate Professor, Informatics and Director, VIS
Mr. Robert J. Stein
Senior Technical Lead
www.cees.iupui.edu
IUPUI


Center for Earth and Environmental Science
A Center Housed within the School of Science and the Department of Geology
Promoting Interdisciplinary Environmental Research, Education, and Public Service


CEES Interdisciplinary Environmental Research Focus Areas
Wetland Ecosystem Restoration
Water Resource Evaluation
The Fate and Transport of Environmental Contaminants
Environmental Data Management, Mapping, and Visualization


Visualization and Interactive Spaces Lab
Explore use of computer-generated graphics, advanced user interaction strategies, and smart room technologies
Design and build applications and spaces for data exploration and learning in labs, museums, science centers

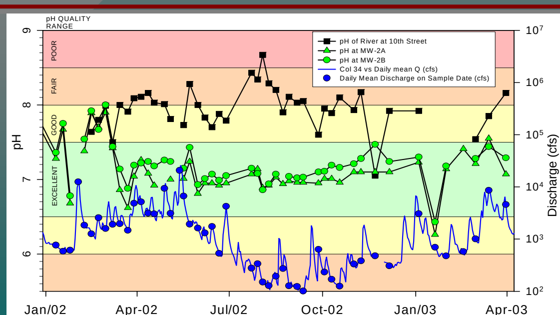
Our Project Partners

- School of Science, IUPUI
- Pervasive Technology Labs
- Eli Lilly Endowment
- Eli Lilly Company
- Efroymson Fund of the Central Indiana Community Fund
- Veolia Water Indianapolis, LLC.
- The Rotary Club of Indianapolis
- Indy Parks
- Indy Parks/Greenways
- U.S. Geological Survey
- City of Indianapolis, Department of Public Works
- Indiana State Museum
- IMAGIS
- Hamilton County Parks

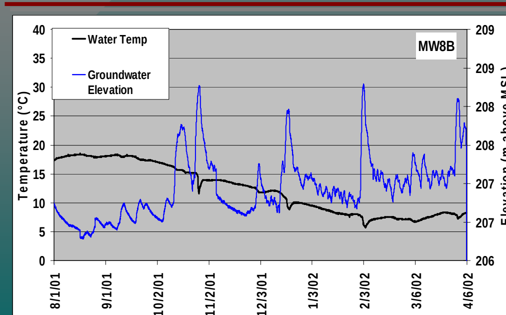
Interdisciplinary Ecosystem Research: The Challenges

- Complex environmental systems
- Multicomponent dynamic systems with each component being heterogeneous both spatially and temporally
- Scientists suffer from an inability to capture the spatial and temporal rates of change and feedbacks using manual observation and data collection methods – luck driven
- Multidisciplinary teams approach questions from different perspectives, measuring/observing different things at different resolution
- Even the best research teams have only infrequent discussions among investigators that are typically unable to evaluate the complete spectrum of data and interrelate the various systems
- Most straightforward data interpretations come from oversimplified observations that focus on limited system dynamics

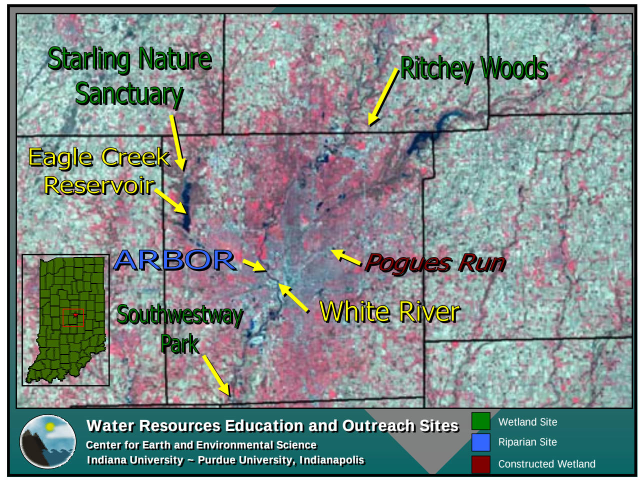
Manual Sampling of pH of River and Groundwater Relative to Automated Sampling of River Discharge



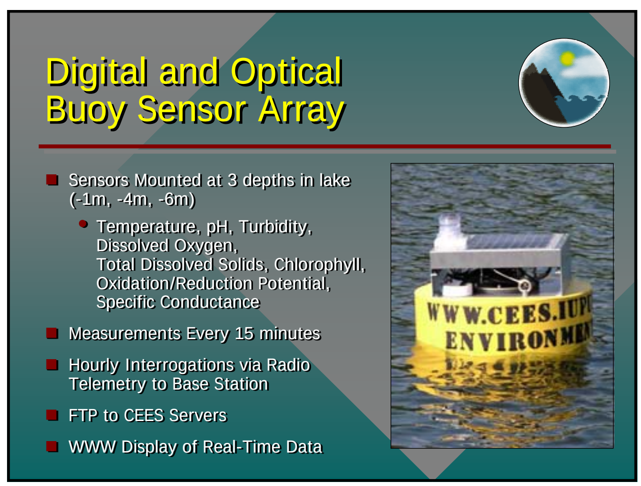
Groundwater Elevation and Temperature from Automated Monitoring Well

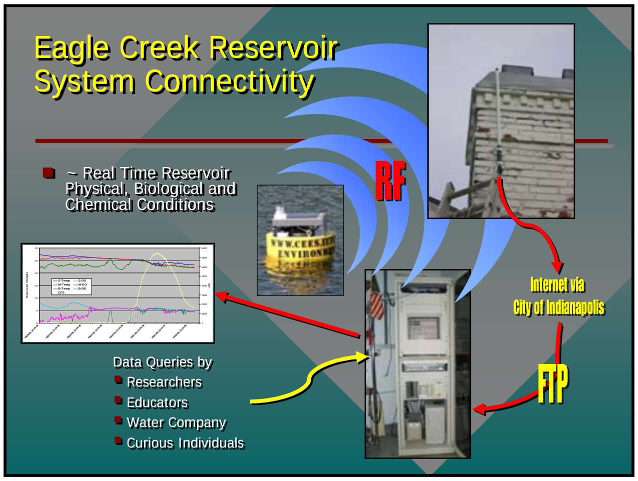


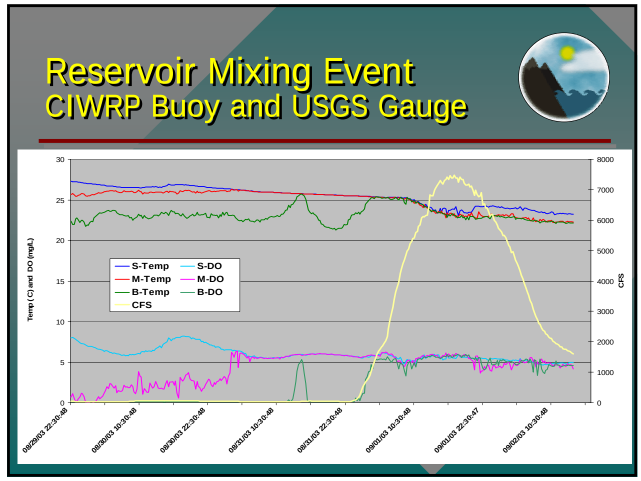

Research Site Network
A Series of Wetland, Riverine and Lake Ecology Sites
Distributed Throughout Marion and Hamilton Counties
Being Restored or Studied by CEES Facilitated Interdisciplinary Teams of Faculty, Staff, Students and Volunteers
Utilized for Ecosystem Research and Teaching for University, K-12, and Community Education

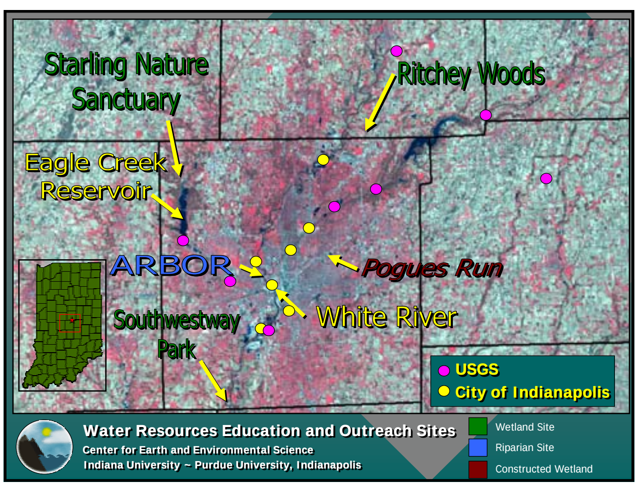

Starling Nature Sanctuary
Ritchey Woods
Eagle Creek Reservoir
ARBOR
Pogues Run
Southwestway Park
White River
Water Resources Education and Outreach Sites
Center for Earth and Environmental Science
Indiana University – Purdue University, Indianapolis
Wetland Site
Riparian Site
Constructed Wetland

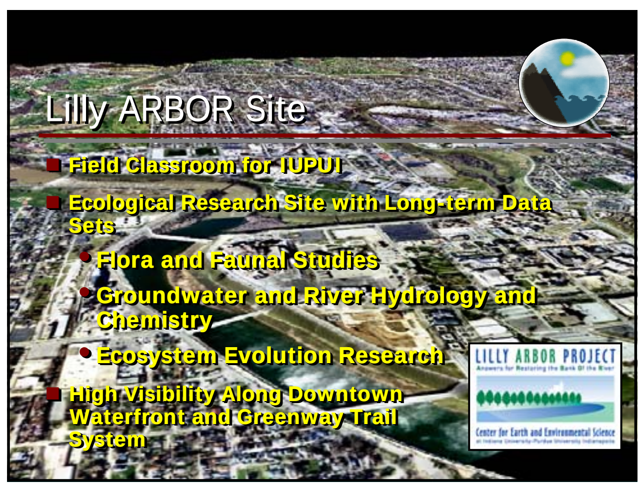

Eagle Creek Reservoir Ecology Site
Eagle Creek Reservoir Water Quality Monitoring via Real-time Water Quality Sensors
Key Component of Research for Management of the Reservoir
Involves Study of Biology, Geology, and Chemistry of Altered Lake System Impacting Drinking Water Supply


Digital and Optical Buoy Sensor Array
Sensors Mounted at 3 depths in lake (-1m, -4m, -6m)
Temperature, pH, Turbidity, Dissolved Oxygen, Total Dissolved Solids, Chlorophyll, Oxidation/Reduction Potential, Specific Conductance
Measurements Every 15 minutes
Hourly Interrogations via Radio Telemetry to Base Station
FTP to CEES Servers
WWW Display of Real-Time Data


Eagle Creek Reservoir
System Connectivity
~ Real Time Reservoir Physical, Biological and Chemical Conditions
RF
Internet via
City of Indianapolis
FTP
Data Queries by
Researchers
Educators
Water Company
Curious Individuals

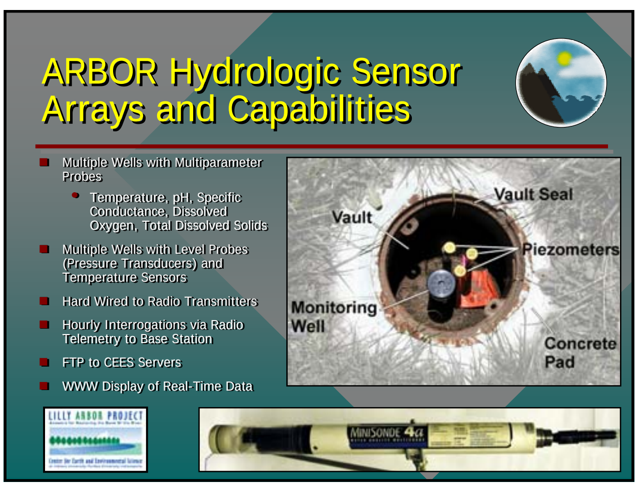

Reservoir Mixing Event
CIWRP Buoy and USGS Gauge
Temp (C) and DO (mg/L)
CFS
S-Temp
M-Temp
B-Temp
CFS
S-DO
M-DO
B-DO

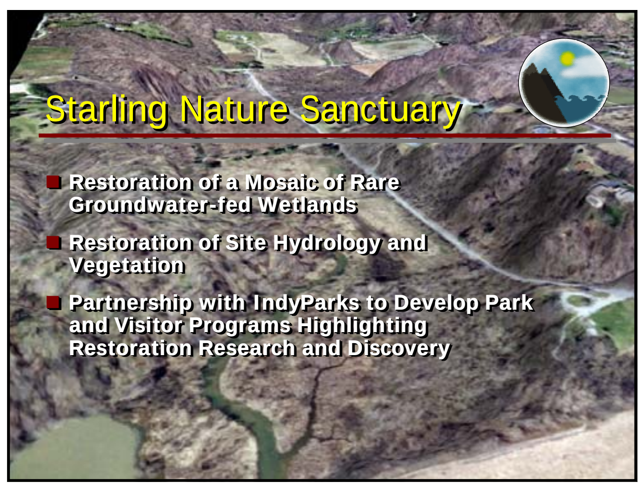

Starling Nature Sanctuary
Ritchey Woods
Eagle Creek Reservoir
ARBOR
Pogues Run
Southwestway Park
White River
USGS
City of Indianapolis
Water Resources Education and Outreach Sites
Center for Earth and Environmental Science
Indiana University ~ Purdue University, Indianapolis
Wetland Site
Riparian Site
Constructed Wetland


Lilly ARBOR Site
Field Classroom for IUPUI
Ecological Research Site with Long-term Data Sets
Flora and Faunal Studies
Groundwater and River Hydrology and Chemistry
Ecosystem Evolution Research
High Visibility Along Downtown Waterfront and Greenway Trail System
LILLY ARBOR PROJECT
Center for Earth and Environmental Science


LILLY ARBOR PROJECT
Answers for Restoring the Bank Of the River
Center for Earth and Environmental Science
Scheduled Wireless Environment
City of Indianapolis Multiparameter Water Quality Probe with Hourly Interrogations via Telephone Lines
Scheduled (PTZ) Digital Cameras Mounted Under Bridges
Scheduled USGS Stage Gauge and Sensor Array Transmitting via Satellite Station on Hourly Schedule to USGS returned to IUPUI via Internet
Weather Station
Radio Receiver
Groundwater Well
Digital Camera
USGS Gauge
Water Quality Probe


LILLY ARBOR PROJECT
Answers for Restoring the Bank Of the River
Center for Earth and Environmental Science


ARBOR Hydrologic Sensor Arrays and Capabilities
Multiple Wells with Multiparameter Probes
Temperature, pH, Specific Conductance, Dissolved Oxygen, Total Dissolved Solids
Multiple Wells with Level Probes (Pressure Transducers) and Temperature Sensors
Hard Wired to Radio Transmitters
Hourly Interrogations via Radio Telemetry to Base Station
FTP to CEES Servers
WWW Display of Real-Time Data
Vault
Vault Seal
Piezometers
Monitoring Well
Concrete Pad
LILLY ARBOR PROJECT
MINISONDE 4a


Starling Nature Sanctuary
Restoration of a Mosaic of Rare Groundwater-fed Wetlands
Restoration of Site Hydrology and Vegetation
Partnership with IndyParks to Develop Park and Visitor Programs Highlighting Restoration Research and Discovery

Starling Wetland Restoration

- Groundwater Well Clusters
- Stream Flow and Water Chemistry Monitoring
- Vegetation Monitoring
- Bird Utilization Studies

Existing Groundwater Well
Future Groundwater Well
Test Cores
Vegetation Monitoring
Stream Monitoring
Spring

Starling Hydrologic Sensor Arrays and Capabilities

- Wells Instrumented with Multiparameter Probes
- Water Level Loggers with Temperature Sensors
- Multiparameter Water Quality Probe on Fishback Creek
- Hourly Interrogations via Radio Telemetry to Base Station Shared with Eagle Creek Reservoir Buoy
- FTP to CEES Servers
- WWW Display of Real-Time Data

Meeting the Challenge

- Increased Observations and Measurement of Complex Ecosystems
  - Overcome Randomness of Sampling
  - Sampling at Frequencies Closer to Natural Variability
- Creating a Centralized Data Management System
- Goal is to Provide a Common Platform for Interdisciplinary Analysis and Visualization

Visualization and Interactive Spaces

Database/Data Management

- Database Systems Under Development
- Programming to Handle Data QA/QC via Identification of Unusual Data Ranges
  - Probe charge
  - Known Normal Ranges for Parameters (trained data set from background data and literature values)

Visualization and Interactive Spaces

Database Requirements

- Receive Incoming Data from Remote Sensors, Manual Sampling, and Continually Growing Geographic Base Data
- Sort, Categorize, Associate, Perform Standard Processing, and QA/QC
- Produce Generalized Active Web Displays
- Allow Remote Web Queries and Return Results In Both Tabular and Graphical Forms

Visualization and Interactive Spaces

Relatively Large Data Sets

- Aerial Photography
  ~7.5 GB per year, per County
- Multi Season, Multi Year Satellite Imagery
  ~1 GB per scene (~1/3 state)
- Digital Elevation Models
  ~50 MB per quad (~9 quads to each County)
- Lidar Elevation Data
  ~167 GB per County
  (per year?)
- Investigation Specific GIS Coverages
  ~21 GB (50k files) and growing
- Streaming Probe Data
  ~15 MB text file per probe per year
- Photography (pre automation)
  ~11 GB (18k images) and growing

Note: Two Common Projections

Environmental Education Program

- Create Web and Virtual Ecosystem Education Modules
- Utilize Virtual Experiments Based Around Real-time Data Queries
- PercepTable at Indiana State Museum

Visualization and Interactive Spaces

Center for Earth and Environmental Science

Fostering a New Level of Interdisciplinary Collaboration at IUPUI and Beyond

www.cees.iupui.edu

Global Analysis of Arthropod Evolution

Craig Stewart, Richard Repasky, John Colbourne, David Hart, Donald Berry, Matthias Mueller, Raymond Sheppard, Eric Wernert, Mary Papakhian, John N. Huffman

University Information Technology Services
&
Center for Genomics and Bioinformatics

In collaboration with
High Performance Computing Center Stuttgart

Outline

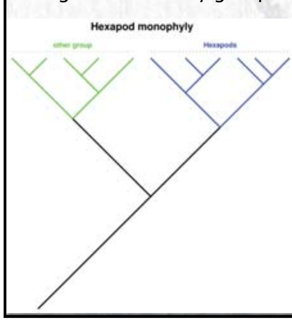
- The SCxy conference and the HPC Challenge
- The biological problem
- The software used
- The global grid
- The results!
- Acknowledgements

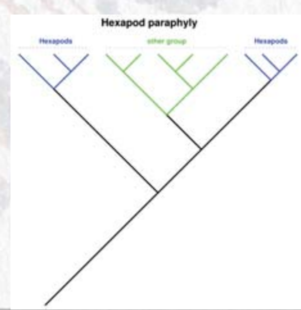
The SCxy conference and the HPC Challenge

- Supercomputing Conference (sponsored by ACM and IEEE)
- High Performance Challenge – demonstrate new capabilities in advanced computing systems

Biological problem

Are Hexapods a single evolutionary group? Are ecdysozoans a single evolutionary group?





INDIANA UNIVERSITY UITS

A partial bestiary

5

INDIANA UNIVERSITY UITS

Software and data analysis

- Non-grid preparatory work
  - Download sequences from NCBI (67 Taxa, 12,162 bp, mitochondrial genes for 12 proteins)
  - Align sequences with Multi-Clustal
  - Determine rate parameters with TreePuzzle
- Grid preparatory work
  - Analyze performance of fastDNAml with Vampir
  - Meetings via Access Grid & CoVise
- The grid software
  - PACXMPI – Grid/MPI middleware
  - Covise – Collaboration and visualization
  - fastDNAml – Maximum Likelihood phylogenetics

6

INDIANA UNIVERSITY UITS

PACX-MPI

- A project of HLRS (High Performance Computing Center Stuttgart)
- PACX-MPI (**PA**rallel **C**omputer e**X**tension) enables seamlessly execution of MPI-conforming parallel applications on a Grid.
- Application recompiled and linked w. PACX-MPI.
- Communication between MPI processes internally is done with the vendor MPI, while communication to other parts of the Metacomputer is done via the connecting network.
- Key advantages:
  - Optimized vendor MPI library is used.
  - Two daemons (MPI processes) take care of communication between systems – allows bundling of communication.

7

INDIANA UNIVERSITY UITS

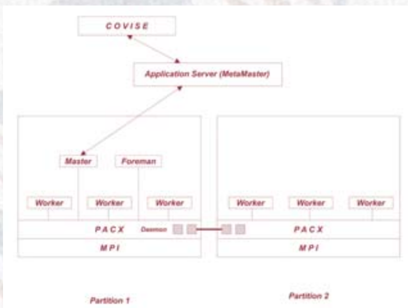
COVISE

- <u>CO</u>llaborative <u>VI</u>sualization and <u>S</u>imulation <u>E</u>nvironment
- A project of HLRS (High Performance Computing Center Stuttgart)
- Focus on collaborative and interactive use of supercomputers
- Interactive startup of calculation on a Computational Grid
- Real-Time visualization of the results and the performance of computation.

8

fastDNAml

- ML analysis of phylogenetic trees based on DNA sequences
- Foreman/worker MPI program
- Heuristic search for best trees
- For 67 taxa: $2.12 \sim 10^{109}$ trees
- Goal: 300 bootstraps, 10 jumbles per – 3000 executions (more than 3x typical!)



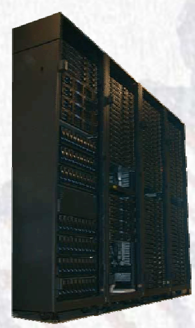
Why this project on a grid?

- Important & time-sensitive biological question requiring massive computer resources
- A biologically-oriented code that scales well
- Grid middleware environment & collaboration tool well suited to the task at hand
- Opportunity to create a grid spanning every continent on earth (except Antarctica)

IBM Research SP (Aries/Orion Complex)

- 1.005 TeraFLOPS. 1st University-owned supercomputer in US to exceed 1 TFLOPS peak theoretical processing capacity.
- Geographically distributed at IUB and IUPUI
- Initially 50th, now 302nd in Top 500 supercomputer list
- An enabler of collaborative research using very large scale computations

AVIDD

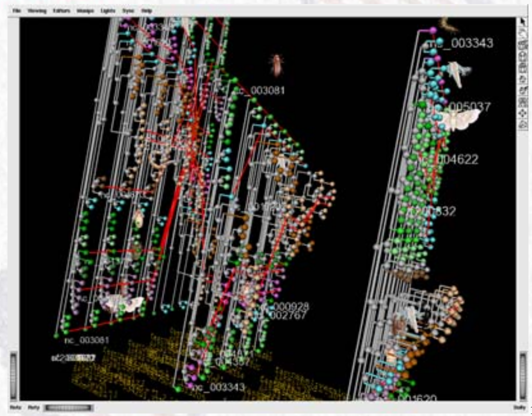


- Analysis and Visualization of Instrument-Driven Data
- Distributed Linux cluster. Three locations: IUN, IUPUI, IUB
- 2.164 TFLOPS, 0.5 TB RAM, 10 TB Disk
- First distributed Linux cluster to achieve more than 1 TFLOPS on Linpack benchmark
- Originally 50th, currently 114th on Top500 list

The metacomputers

• One	Origin 2000	32	Spain
	Linux cluster	64	Japan
	Linux cluster	12	Australia
	IBM SP	32	US
• Two	T3E	128	Germany
	IBM SP	64	US
	Dec Alpha	4	Brazil
	Sun fire 6800	16	Singapore
• Three	Hitachi SR8000	32	Germany
	Cray T3E	128	UK
	Cray T3E	32	US
	IBM SP (Blue Horiz)	32	US
• Four	Dec Alpha (Lemieux)	64	US
• Five	Linux System	1	Tunisia

Five functional units, 8 types of systems (several on Top500 list); 6+ vendors; 641 processors; 9 countries, 6 continents



The results

- Hundreds of trees were analyzed during the course of the week
- The biological results are still being analyzed
- Our HPC challenge project was awarded the prize for "Most geographically distributed application"

For further information

- fastDNAml: http://www.indiana.edu/~rac/hpc/fastDNAml/
- PACXMPI: www.hlrs.de/organization/pds/projects/pacx-mpi
- COVISE: www.hlrs.de/organization/vis/covise
- HLRS: www.hlrs.de
- UITS: uits.iu.edu
- Center for Genomics and Bioinformatics: www.cgb.indiana.edu
- SCxy: www.supercomp.org
- about.uits.iu.edu/divisions/rac/index.html
- about.uits.iu.edu/divisions/rac/pubsstaff.html
- ingen.iu.edu
- it.iu.edu

Acknowledgments

- This research was supported in part by the Indiana Genomics Initiative. The Indiana Genomics Initiative of Indiana University is supported in part by Lilly Endowment Inc.
- This work was supported in part by Shared University Research grants from IBM, Inc. to Indiana University.
- This material is based upon work supported by the National Science Foundation under Grant No. 0116050 and Grant No. CDA-9601632. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).
- Assistance with this presentation: John Herrin, Malinda Lingwall, W. Les Teach
- Thanks to the SciNet team and SC2003 organizers!

Our partners

Rainer Keller, Matthias Hess	HLRS, University of Stuttgart
Richard Repasky	UITS, Indiana University
John Colbourne	Center for Genomics and Informatics, Indiana University
Craig Stewart, David Hart	UITS, Indiana University
Jennifer Steinbachs	Center for Genomics and Bioinformatics, Indiana University
Uwe Woessner	HLRS, University of Stuttgart
Donald Berry	UITS, Indiana University
Matthias Mueller	HLRS, University of Stuttgart
Huian Li	UITS, Indiana University
Gary W. Stuart	Center for Genomics and Bioinformatics, Indiana University
Michael Resch	HLRS, University of Stuttgart
Eric Wernert	UITS, Indiana University
Martin Aumüller, Ulrich Lang	HLRS, University of Stuttgart
Markus Buchhorn	Australia National University
Hiroshi Takemiya	National Institute of Advanced Industrial Science & Technology, Japan
Rim Belhaj	ISET'Com, Tunesia
Wolfgang E. Nagel	ZHR, Technical University of Dresden
Sergui Sanielevici	Pittsburgh Supercomputing Center
Sergio takeo Kofuji	LCCA/CCE-USP
David Bannon	Victorian Partnership for Advanced Computing, Australia
Norihiro Nakajima	Japan Atomic Energy Research Institute
Rosa Badia	CEPBA-IBM Research Institute
Mark A. Miller	San Diego Supercomputer Center
Hyungwoo Park	Korea Institute of Science and Technology Information
Rick Stevens	Argonne National Laboratory
Fang-Pang Lin	National Center for High Performance Computing
John Brooke	Manchester Computing
David Moffett	Purdue University
Tan Tin Wee	National University of Singapore
Greg Newby	Arctic Region Supercomputer Center
J.C.T. Poole	CACR, Cal-Tech
Ramched Hamza	Sup'com, Tunesia
Mary Papakhian, John N. Huffman	UITS, Indiana University
Leigh Grundhoeffer	UITS, Indiana University
Ray Sheppard	UITS, Indiana University
Peter Cherbas	Center for Genomics and Bioinformatics, Indiana U.
Stephen Pickles, Neil Stringfellow	CSAR, University of Manchester

Thank you!

Questions?

Indiana's Optical Fiber Initiative

I-LIGHT

PURDUE UNIVERSITY

Ingestion, Analysis and Distribution of Real-time and Archival Satellite Data for Indiana: The Purdue Terrestrial Observatory (PTO), Laboratory for Analysis of Remote Sensing (LARS) and IndianaView

Gilbert L. Rochon, Ph.D., MPH, Associate Vice President, Collaborative Research & Engagement, Purdue University, ITaP

Larry Biehl, Systems Manager, Purdue Terrestrial Observatory

Joseph Quansah, Graduate Research Assistant, Purdue University

2004 I-Light Symposium
Indianapolis, IN March 9, 2004

Recognition:

PURDUE UNIVERSITY

2003 BEST places to work ACADEMIA The Scientist

The Scientist www.the-scientist.com

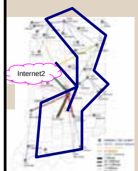
Best Places to Work - 2003

- **The top overall results for US institutions:**
  - 1- Fox Chase Cancer Center, Philadelphia, PA
  - 2- Purdue University, West Lafayette, IN
  - 3- Yale University, New Haven, CT

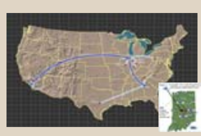
http://www.the-scientist.com/press_releases/academia_031020.html

Balancing Networking Investments

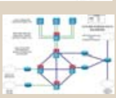
PURDUE UNIVERSITY



National
- Internet2 (10 Gbps)
- TeraGrid (40 Gbps)
- NLR – participation via CIC



State
- Invested in fiber optic for Universities
- I-Light - $5 M (2 Gbps)
- I-Light 2 - $10 M allocated in '04

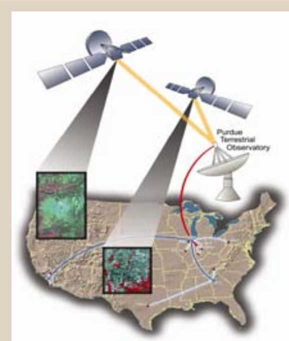


Campus
- Reengineering Purdue's backbone network ($6.9 M)
- Gigabit between buildings, 10/100 Mbps to desktops and GigE on demand
- Full campus wireless deployment ($1.3 M)
- "Shadow network" (dark fiber) provides experimental environment

NSF: Extensible Terascale Facility (ETF)

PURDUE UNIVERSITY

- **$3 million NSF grant will provide a high band width data transmission rate to acquire, exchange and disseminate geographic data to partnering facilities nationwide**
- **The ETF will better enable distributed data centers to engage in virtual data mining, data fusion, and data product development in support of decision-making, research, instruction and economic development**



Real-Time Remote Sensing (RTRS)

PURDUE UNIVERSITY

- RTRS Definition: Near Zero Latency for Ingestion, Analysis, Intelligent Archiving & Distribution of Data & Data Products from Orbiting Satellites, Reconnaissance Aircraft and *In Situ* Monitoring Devices
- RTRS Incentives:
  - Proliferation of Satellites & Access to Current & Archival Satellite, Airborne and Ground Truth Data
  - Need for anticipation and rapid response to biogenic and anthropogenic disasters

Utility of RTRS

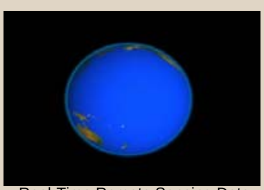
PURDUE UNIVERSITY

- **Singular:**
  - Emergency Response
  - Disaster Relief
- **Combined with Intelligent Ready Access to Archival RS Data:**
  - Environmental Impact Estimation
  - Change Detection
  - Early Warning
  - Threat Determination
  - Vulnerability Assessment
  - Damage Analysis
  - Forecasting & Hindcasting
  - Alternative Future Scenario Generation

Purdue Terrestrial Observatory (PTO)

PURDUE UNIVERSITY

<u>Participating Departments</u>
Agronomy
Ag. & Bio. Engineering
Agriculture
Agricultural Economics
Atmospheric Science
Chemistry
Civil Engineering
Computer Science
Earth & Atmospheric Sciences
Electrical & Computer Eng.
Env. Science & Engr. Inst.
Forestry & Natural Resources
ITaP
Management
Nuclear Engineering



Real-Time Remote Sensing Data

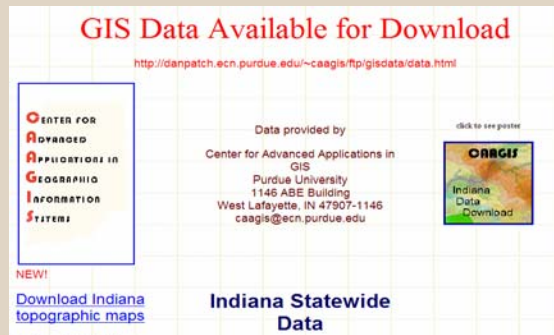
<u>Sample Applications</u>:
Precision Farming
Homeland Security
Disaster Intervention
Forestry & Ecology
Urban Planning
Epidemiology
Oceanography
Industrial Development
Geology & Mineralogy

Satellite Masks for Purdue & Univ. of Alaska ASF

PURDUE UNIVERSITY

LARS 40th Anniversary

CAAGIS:Center for Advanced Applications of GIS



AmericaView

To: IndianaView Consortium
From: AmericaView, Inc. Board
1/9/04

"Congratulations - the Board of AmericaView has approved your application for Affiliate Status with AmericaView for IndianaView."
- Tim Warner
Tim.Warner@Mail.wvu.edu West Virginia View: www.wvview.org

IndianaView will permit acquisition of archival satellite data for the entire State of Indiana, thereby providing decision-makers with the benefit of virtual hindsight, through dynamic change detection, and virtual foresight, through visualization of alternative future scenarios, incorporating real-time data from the Purdue Terrestrial Observatory (PTO).

AmericaView

Member Organizations:

Ball State University - Muncie
Indiana State University - Terre Haute
IUPUI - Indianapolis
Indiana University – Bloomington
Indiana University – South Bend
Martin University – Indianapolis
Purdue University – West Lafayette
Purdue University – Calumet
University of Notre Dame – South Bend
Indiana Geological Survey (IGS)
Indiana Geographic Information Council (IGIC)
Indiana Space Grant Consortium (ISGC)

AmericaView

PURDUE UNIVERSITY

Benefits of IndianaView's Affiliation with AmericaView:

- **Discounted pricing on acquisition of satellite data**
- **Shared access to the distributed satellite data archives for the other 14 affiliated member states and their academic partners**
- **Facilitates inter-state collaboration for major joint research initiatives, distance learning, community outreach, and regional economic development**
- **Serves a complimentary role with the real-time data stream from the Purdue Terrestrial Observatory, thereby enabling change detection between periods covered within the archive and contemporary observation**
- **Offers the potential for international replication in Asia, Africa & South America to support Third World Development plans**

Utility of RTRS

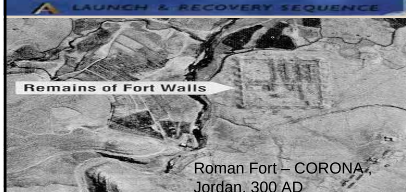
PURDUE UNIVERSITY

- **Singular:**
  - Emergency Response
  - Disaster Relief
- **Combined with Intelligent Ready Access to Archival RS Data:**
  - Environmental Impact Estimation
  - Change Detection
  - Early Warning
  - Threat Determination
  - Vulnerability Assessment
  - Damage Analysis
  - Forecasting & Hindcasting
  - Alternative Future Scenario Generation

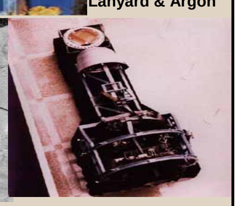
Corona, Lanyard & Argon

PURDUE UNIVERSITY

Satellite: 6-8ft.resolution approved by Eisenhower-1958 Declassified by Clinton-1995. 850,000 images: 1960-1972:Corona, Lanyard & Argon



Roman Fort – CORONA, Jordan, 300 AD



U2

PURDUE UNIVERSITY

U2 Spy Plane Versions-Lockheed

Fargo, North Dakota-Image taken by U2

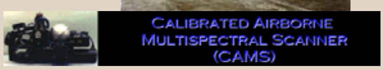
Airborne Remote Sensing:

PURDUE UNIVERSITY

NASA Wallops Island
Hot Air Balloon

NASA Stennis Lear Jet

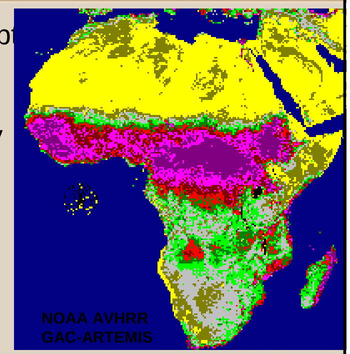
Airborne remote sensing has heretofore been the method of choice for high resolution over flights, utilizing experimental sensors & pre-launch testing of precursor instrumentation.



NOAA:

PURDUE UNIVERSITY

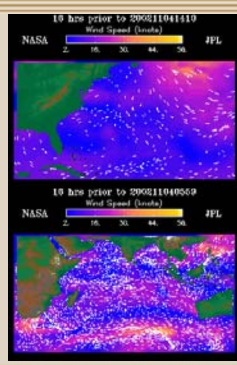
- DOC NOAA- US Dept of Commerce, National Oceanic & Atmospheric Agency
  - Landsat
  - AVHRR HRPT
    - » GAC
    - » LAC
  - GOES



NASA:

PURDUE UNIVERSITY

- NASA:
  - MODIS
    - » Terra
    - » Aqua (USA, Brazil & Japan)- atmospheric temperature, humidity, clouds, sea surface temp
  - GRACE (USA & Germany)- gravity field temporal and geographic variations
  - ICEsat- Ice sheet topography
  - ACRIMSAT
  - UARS
  - SORCE-solar irradiance
  - SeaWinds



MODIS:

PURDUE UNIVERSITY

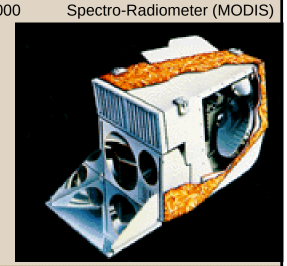
http://modland.nascom.nasa.gov/gallery/?Korea.A2000097.0235.825x1075.jpg

MODIS Image: April 6, 2000

Moderate Resolution Imaging Spectro-Radiometer (MODIS)

Korea

36 Bands



Private Sector:

Space Imaging's IKONOS: Pyramids at Gisa: 1 m. resolution- panchromatic

Private Sector:

- Pyramid at Gisa at 0.6 meter resolution- Digital Globe's Quickbird

Europe:

- **ESA- European Space Agency**:
  - Mt. Aetna-Sicily Smoke Plume-Oct. 30, 2002 (Proba CHRIS)
  - Proba-Project for on Board Autonomy CHRIS (Compact High Resolution Imaging Spectrometer)
  - MERIS-Medium Resolution Imaging Spectrometer (Envisat)
  - ATSR- Along Track Scanning Radiometer (ERS-2)

India:

- Indian Space Research Organization (ISRO)
  - IRS (1B, 1C, P3, P4)
  - INSAT (3A-3E)
  - Successor Systems:
    - Resourcesat- 5.8 m resolution - multispectral
    - Cartosat series- 2.5 m and 1 m spatial resolution sensors

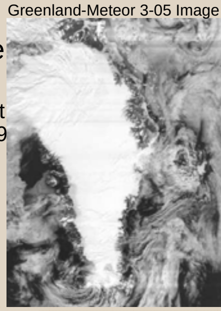
Colorado River- IRS

Russia:

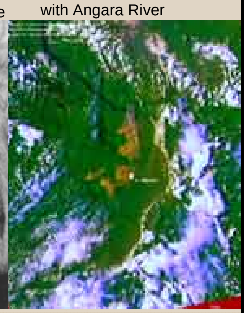
PURDUE UNIVERSITY

- Glavcosmos /Russian Aero Space Agency:
  - Meteor- First Launch 1969
  - Okean
  - Resurs
  - Electro

July 5, 2001
Greenland-Meteor 3-05 Image



Irkutsk, Eastern Siberia- with Angara River



Korea:

PURDUE UNIVERSITY

- Korea Aerospace Research Institute (KARI) (Daejun), Electronics & Telecommunications Research Institute (ETRI), KAIST and Multiple Universities – TRW contractors for KOMPSAT Series
- KITSAT
- Remote Sensing Research Lab of the Meteorological Research Inst. (METRI)
- National Fisheries Research Institute (NFRDI) Remote Sensing

KOMPSAT - Korea Multipurpose Satellite

Japan:

PURDUE UNIVERSITY

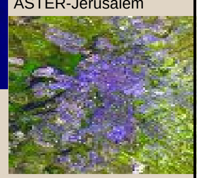
- NASDA – (National Space Development Agency of Japan): Monitoring ozone and greenhouse gases, pursuant to the Kyoto Protocol (mission-specific/small size)
  - JERS
  - ADEOS
  - ASTER (NASDA/NASA)
  - ODUS- Ozone Dynamics Ultraviolet Spectrometer
  - SOFIS-Solar Occultation
  - Fourier Transform Spectrometer

ASTER: 10/20/2002

The Isle of Jersey, Channel Island in the Bay of Mont St Michel off the north-west coast of France.

JERS-1

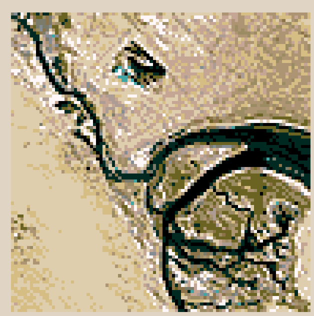
ASTER-Jerusalem



France:

PURDUE UNIVERSITY

- CNES (Centre National d'Etudes Spatiales)
  - Systeme Probatoire pour L'Observation de la Terre (SPOT)
    - » Haut Resolution Visible (HRV)
    - » Panchromatic



Remote Sensing Diversity:

PURDUE UNIVERSITY

- CONAE (Argentina) – Sac-C
- Tubitak (Turkey)- EURASIASAT-1
- INPE (Brazil) – CBERS (Chinese Brazilian Earth Resource Satellite)
- CNSA (China) -Feng Yun
- ISA (Israel) –EROS (Res. 1.8m; 480 km orbit, 3-7 day revisit, depending upon latitude)

Disaster Monitoring Constellation (DMC):

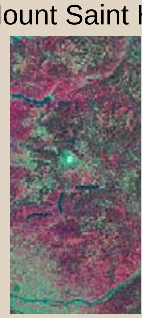
PURDUE UNIVERSITY

- **Disaster Monitoring Constellation (DMC):**
  - The DMC will comprise seven Earth observation microsatellites launched into low Earth orbit to provide daily imaging revisit anywhere in the world.
  - The DMC Consortium:
    - » Algeria
    - » China
    - » Nigeria
    - » Thailand
    - » Turkey
    - » Vietnam
    - » United Kingdom
  - Each country is building an advanced yet low-cost Earth observation microsatellite to form the first ever constellation specifically designed and dedicated to monitoring natural and man-made disasters. The first DMC microsatellite is scheduled to be launched for Algeria in 2002 and subsequent microsatellites into the same orbit in 2003 & 2004.
  - The objective of the Consortium is to derive the maximum mutual benefit from the constellation through collaboration and cooperation between the DMC Partners. The partners in the DMC Consortium agreed to exchange their DMC satellite resources and data to achieve a daily Earth observation imaging capability for disaster monitoring and other dynamic phenomena.

Landsat MSS & TM:

PURDUE UNIVERSITY

- Mount Saint Helens Eruption Sequence: MSS



Sept. 15, 1973

May 22, 1983

August 31, 1988

Natural Phenomena RTRS:

PURDUE UNIVERSITY

- Vincente (NASA Goddard) & Costa (U. of Vicosa, Brazil): flash flood monitoring in the Brazilian Amazon
- Zhang & Xie (China Seimological Bureau) Earthquake Disaster Reduction
- Pichel, Clemente-Colon, et al (NOAA NESDIS); Marschalk, et al. (German Remote Sensing Data Center) Near Real-Time Sea Ice and Ocean Surface, oil slicks w/ SAR
- Palm, et al. (SSAI) & Spinhirne, et al. (Goddard) Geoscience Laser Altimeter System (GLAS) for monitoring polar cloud climatology
- Yang, et al. (Sichuan Normal U.-China) Rapid Assessment of flood disasters, w/ SAR and GIS

Human/Computer Interface for RTRS:

PURDUE UNIVERSITY

- Young & Harrah (NASA Langley) and de Haag (Ohio University): (Terrain Awareness and Warning Systems (TAWS) and Synthetic Vision Systems (SVS) to avert aircraft accidents in real time, associated with loss of pilots' situational awareness, by deploying stored spatial data, radar altimeters, forward looking autonomous integrity monitoring (FLAIM), etc.
- Diego Loyola (German Aerospace Center) Combining Neural Networks and Near-Real-Time Processing of Satellite Data for the Global Ozone Monitoring Experiment (GOME)
- Fiebe, Kloos, et al. (University of Hannover, Germany) Compact Real-Time SAR Processing System
- Array Systems (Canada): Parallelized software for analysis of Radarsat data within a Beowulf cluster or supercomputing environment

Developing Initiatives

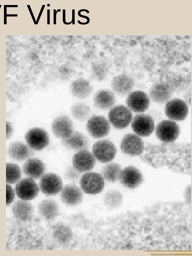
PURDUE UNIVERSITY

- NASA SEDAC UWG-Columbia University
- Disaster Informatics/ Crisis Grid
- Institute for Geo-Informatics Research
- NATO Committee on he Challenges to Modern Society (CCMS) Pilot Study (Budapest Meeting) Clean Products & Processes; Remote Sensing & Sustainable Technology
- Ecological Acoustics Research (EAR) Facility
- Center for Climate Change Research (9 new faculty hires)
- Center for Intercultural Communications
- Remote Sensing to Mitigate Epidemics & Epizootics

Rift Valley Fever

PURDUE UNIVERSITY

- Culex pipiens & Aedes
- RVF Virus

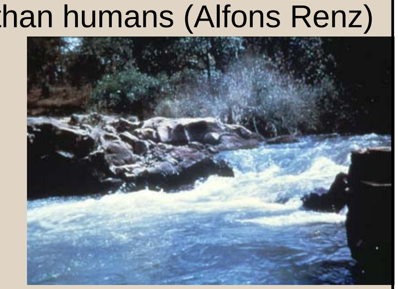


Linthicum, KJ, et al. (1999) Climate & Satellite indicators to forecast Rift Valley Fever epidemics in Kenya. *Science* 285, 397-4000.

River Blindness

PURDUE UNIVERSITY

- 27 species of onchocerca affecting animals other than humans (Alfons Renz)



Leishmaniasis

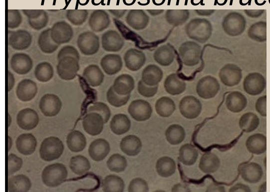
PURDUE UNIVERSITY

- Phlebotomus paptasi>>>
- Leishmania

Trypanosomiasis

PURDUE UNIVERSITY

- Tsetse fly (Bantu Name) *Glossina morsitans*
- *Trypanosoma brucei*

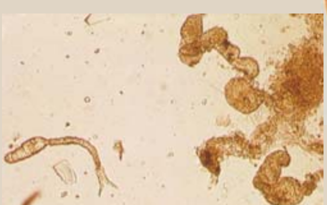


Rogers, DJ & S. Randolph. "Monitoring Trypanosomiasis in Space & Time." Parasitology. 1993; 351:739-741.

Schistosomiasis

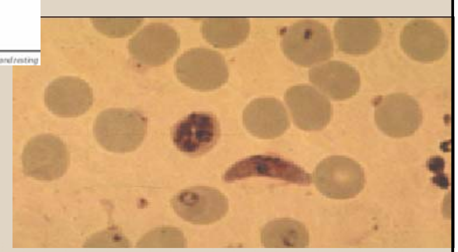
PURDUE UNIVERSITY

- *Biomphylaria >>>>>> & Bulinus truncatus*
- *Schistosoma mansoni*



Plasmodium Falciparum: Anopheles gambiae

PURDUE UNIVERSITY



Time-Sensitive Anthropogenic Disaster Assessment:

PURDUE UNIVERSITY

The World Trade Center Site Military EC-3 Recon Photo

Ground Zero:

PURDUE UNIVERSITY

Disaster Monitoring

PURDUE UNIVERSITY

Disaster Assessment – 9/11

PURDUE UNIVERSITY

Source:
Mete Sozen, CE
Sami Kilic, CE
Chris Hoffmann, CS
Ahmed Sameh, CRI
Voicu Popescu, CS
Scott Meador, CGT
William Whitson, ITaP
Hendry Lim, CS;
Amit Chourasia, CGT

Ad Hoc Purchase vs. Real-Time Acquisition:

PURDUE UNIVERSITY

- University-Based Real-Time Remote Sensing in the USA (Partial Listing)
  - Alaska SAR Facility (Univ. of Alaska-Fairbanks)
  - University of Wisconsin at Madison
  - University of Miami Rosenstiel School of Oceanography
  - Louisiana State University (LSU)
  - University of Texas at Austin
  - Rutgers University
  - Scripps Institute of Oceanography, University of California, San Diego
  - City University of New York
  - Howard University
  - University of Puerto Rico at Mayaguez
  - Elizabeth City State University
  - University of California at Berkeley (Real-Time Solar Observations)
  - Purdue University: [in development]: The Purdue Terrestrial Observatory (PTO)
- University-Based Real-Time Remote Sensing in Other Countries (Partial Listing)
  - Singapore National University
  - Istanbul Technical University
  - University of Reading – UK
  - Bradford University – UK
  - University of Dundee -UK
- Archival Data Sources (Partial Listing)
  - USGS Eros Data Center. Sioux Falls, South Dakota [http://www.usgs.gov]
  - NASA
  - NOAA
  - UN FAO ARTEMIS for AVHRR NDVI [in collaboration with NASA Goddard Space Flight Center and the University of Reading, UK]

Remote Sensing Diversity:

PURDUE UNIVERSITY

- CONAE (Argentina) – Sac-C
- Tubitak (Turkey)- EURASIASAT-1
- INPE (Brazil) – CBERS (Chinese Brazilian Earth Resource Satellite)
- CNSA (China) -Feng Yun
- ISA (Israel) –EROS (Res. 1.8m; 480 km orbit, 3-7 day revisit, depending upon latitude)

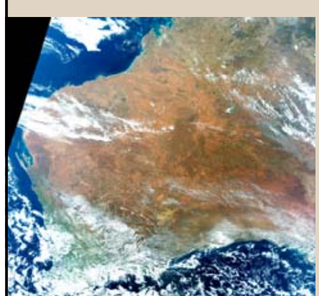
Disaster Monitoring Constellation (DMC):

PURDUE UNIVERSITY

- **Disaster Monitoring Constellation (DMC):**
  - The DMC will comprise seven Earth observation microsatellites launched into low Earth orbit to provide daily imaging revisit anywhere in the world.
  - The DMC Consortium:
    - » Algeria
    - » China
    - » Nigeria
    - » Thailand
    - » Turkey
    - » Vietnam
    - » United Kingdom
  - Each country is building an advanced yet low-cost Earth observation microsatellite to form the first ever constellation specifically designed and dedicated to monitoring natural and man-made disasters. The first DMC microsatellite is scheduled to be launched for Algeria in 2002 and subsequent microsatellites into the same orbit in 2003 & 2004.
  - The objective of the Consortium is to derive the maximum mutual benefit from the constellation through collaboration and cooperation between the DMC Partners. The partners in the DMC Consortium agreed to exchange their DMC satellite resources and data to achieve a daily Earth observation imaging capability for disaster monitoring and other dynamic phenomena.

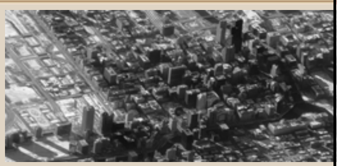
Need for Course Resolution & High Resolution Data:

PURDUE UNIVERSITY



Need for Moderate Resolution & High Resolution Data:

PURDUE UNIVERSITY



ITaP INFORMATION TECHNOLOGY AT PURDUE

COLLABORATIVE RESEARCH & ENGAGEMENT

49

PURDUE UNIVERSITY

High Performance Classroom-Distributed Rendering Environment (HPC-DRE)

W. Scott Meador, Angel Hernandez, Lee Gooding
Envision/ICS

I-Light Symposium

High Performance Classroom (HPC) Definition

PURDUE UNIVERSITY

- Classroom that has access to unrestricted computational resources (or nearly so)
- Components
  - Storage
  - CPU cycles and large memory
  - High Performance Applications
  - Queuing and scheduling
  - Data sets

Examples

PURDUE UNIVERSITY

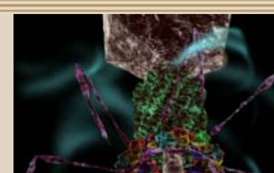
- Three dimensional rendering/animation (in progress)
- Computational chemistry (upcoming)
- Discrete Event Simulation
- Parallel Programming

Purdue DRE

PURDUE UNIVERSITY

What is the need for a DRE?

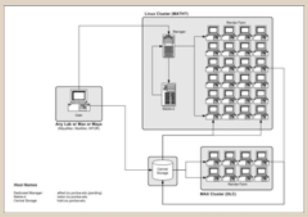
- Time
  - » Renders can sometimes take hours per frame.
- Lab availability
  - » In the past, students were forced to “baby-sit” long renders in the lab.
- Complexity



Purdue DRE

Purdue's Distributed Rendering Environment (DRE)

- What is it?
  - » Purdue DRE is a network of multi-platform machines working in conjunction to render computer generated 3D imagery

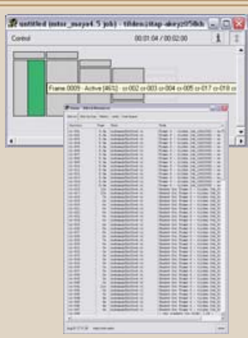


Purdue DRE

What about the DRE machines?

- Hardware:
  - » 48 Recycled Lab machines running Linux
  - » About ~170 Pentium 4s running Windows XP Professional provided by the SoT
  - » Hoping to expand to other labs on campus
  - » Hoping to expand to other campuses

Purdue DRE



What about the software?

- Software:
  - » Alias Maya 5.x
  - » Discreet 3ds max 6.x
  - » Pixar RenderMan 11.5.x
  - » Pixar RenderMan Artist Tools
  - » Animal Logic MayaMan and MaxMan

Purdue DRE

The Submission Tools

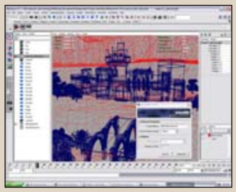
- Submission tools come in two flavors

Purdue DRE

PURDUE UNIVERSITY

Submission Tool Features:

- Supports Renderman 11.5.x
  - » Scene Translation supports both Pixar MTOR (Maya version only) and Animal Logic Software
- Supports Maya and 3ds max Native Renderers
- Email Notification
- Cross-Platform Rendering
- “Fire and Forget” render submission
- Artist Integration



Purdue DRE

PURDUE UNIVERSITY

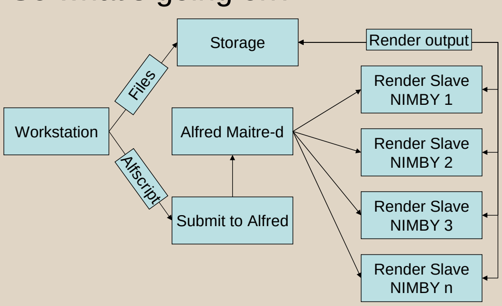
So what’s going on?

- The current scene is analyzed for required render information and data
  - » Scene files and textures are gathered to a temp. location
  - » A render to-do list (Alfscript) is generated
- The files are uploaded to central storage
- Alfscript is sent to the Manager via SSH
- Manager allocates resources and begins the job.
  - » Is NIMBY running?
  - » Does the machine pass metrics test?
- Each node renders one frame at a time, then outputs it back to the central storage.

Purdue DRE

PURDUE UNIVERSITY

So what’s going on?



Purdue DRE

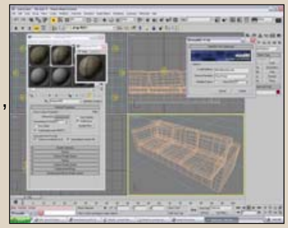
PURDUE UNIVERSITY

Demo in Maya

Purdue DRE

Where we have been:

- Fall 2003
  - » Offered Unattended Rendering via PRMan (Maya, Max), and Maya Native
- Spring 2004
  - » Render Nodes Potentially Available 24/7 –NIMBY
  - » Added 3ds max 6



Purdue DRE

Future Work:

- Fine-tune current implementation
  - » Add Mental Ray if possible
- Incorporate automatic compilation of individual frames to video
- Web-based submission system and new dispatcher
- Build into other pipelines – become rendering back-end for scientific visualization applications
- Cross campus/state/country rendering
- Grid Rendering

Purdue DRE

Any Questions?

Data Placement in Widely Distributed Environments (WiDE)

Miron Livny
Computer Sciences Department
University of Wisconsin-Madison
miron@cs.wisc.edu

" … Since the early days of mankind the primary motivation for the establishment of *communities* has been the idea that by being part of an organized group the capabilities of an individual are improved. The great progress in the area of inter-computer communication led to the development of means by which stand-alone processing sub-systems can be integrated into multi-computer *'communities'*. … "

Miron Livny, " *Study of Load Balancing Algorithms for Decentralized Distributed Processing Systems*.", Ph.D thesis, July 1983.

Claims for "benefits" provided by Distributed Processing Systems

- High Availability and Reliability
- High System Performance
- Ease of Modular and Incremental Growth
- Automatic Load and Resource Sharing
- Good Response to Temporary Overloads
- Easy Expansion in Capacity and/or Function

"What is a Distributed Data Processing System?" , P.H. Enslow, Computer, January 1978

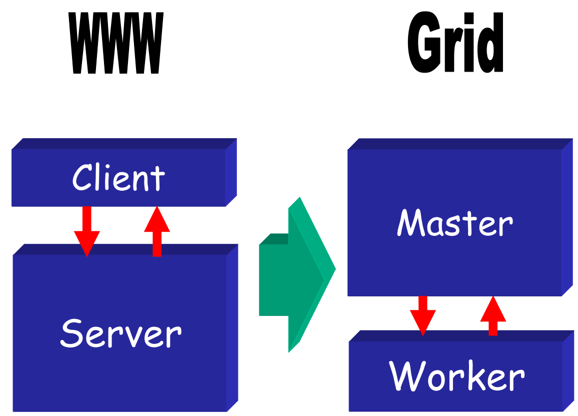
Benefits to Science

- Democratization of Computing – "you do not have to be a SUPER person to do SUPER computing." (accessibility)
- Speculative Science – "Since the resources are there, lets run it and see what we get." (unbounded computing power)
- Function shipping – "Find the image that has a red car in this 3 TB collection." (computational mobility)

The Ethernet Protocol

IEEE 802.3 CSMA/CD - A truly distributed (and very effective) access control protocol to a shared service.

- Client responsible for access control
- Client responsible for error detection
- Client responsible for fairness



Being a Master

Customer "deposits" task(s) with the master that is responsible for:

- Obtaining resources and/or workers
- Deploying and managing workers on obtained resources
- Assigning and delivering work unites to obtained/deployed workers
- Receiving and processing results
- Notify customer.

Application Responsibilities

- Use algorithms that can generate very large numbers of independent tasks – "*use pleasantly parallel algorithms*"
- Implement self-contained portable workers – "*this code can run anywhere!*"
- Detect failures and react gracefully – "*use exponential back off, please!*"
- Be well informed and opportunistic – "*get your work done and out of the way* !"

High Throughput Computing

For many experimental scientists, scientific progress and quality of research are strongly linked to computing throughput. In other words, they are less concerned about instantaneous computing power. Instead, what matters to them is the amount of computing they can harness over a month or a year --- they measure computing power in units of scenarios per day, wind patterns per week, instructions sets per month, or crystal configurations per year.

High Throughput Computing is a 24-7-365 activity

MB/y $\neq$ *(60*60*24*7*52)*MB/s*

Customer requests:

Place *y = F(x)* at L!

Master delivers.

Data Placement

Management of storage space and bulk data transfers play a key role in the end-to-end performance of an application.

- Data Placement (DaP) operations must be treated as "first class" jobs and explicitly expressed in the job flow
- Fabric must provide services to manage storage space
- Data Placement schedulers are needed.
- Data Placement and computing must be coordinated
- Smooth transition of CPU-I/O interleaving across software layers
- Error handling and garbage collection

www.cs.wisc.edu/condor

A simple plan for y=F(x) → L

1. Allocate (size(x)+size(y)+size(F)) at SE(i)
2. Move x from SE(j) to SE(i)
3. Place F on CE(k)
4. Compute F(x) at CE(k)
5. Move y to L
6. Release allocated space

Storage Element (SE); Compute Element (CE)

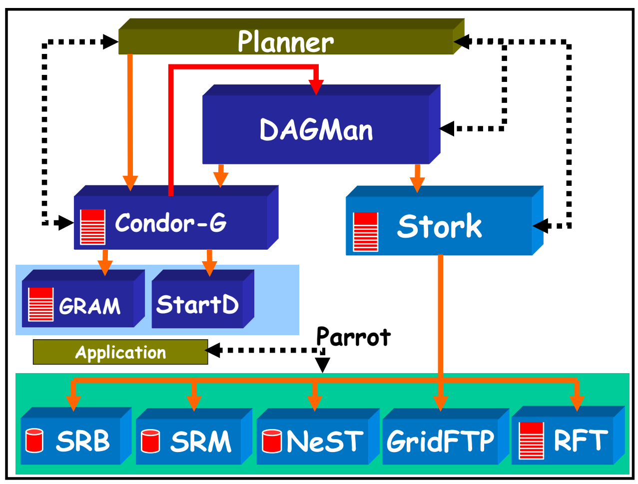
www.cs.wisc.edu/condor

Current Status

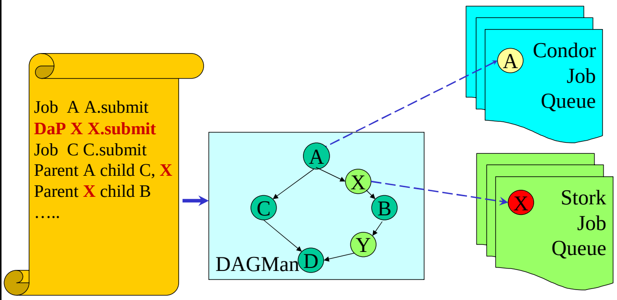
Implemented a first version of a framework that unifies the management of compute and data placement activities.

- DaP "aware" Job Flow (DAGMan).
- Stork – A DaP scheduler
- Parrot – A tool that "speaks" a variety of distributed I/O services
- NeST – A portable Grid enabled storage appliance

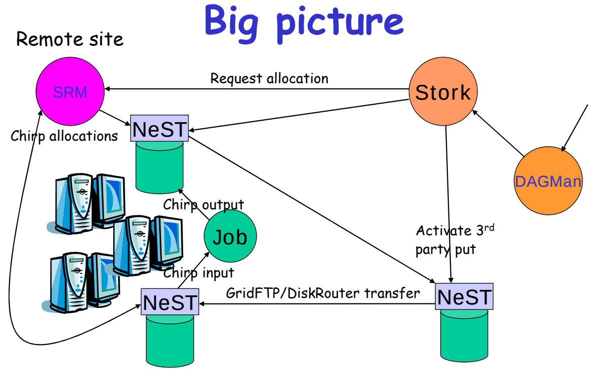
www.cs.wisc.edu/condor



DaP aware DAGs



Big picture



WCER

The Wisconsin Center for Education Research is one of the oldest, largest, and most productive education research centers in the world. A part of the University of Wisconsin-Madison's School of Education, WCER provides a productive environment where some of the country's leading scholars conduct basic and applied education research. The WCER roster includes research centers and projects that are currently investigating a variety of topics in education.

Managing Video Data

- ***Managing Video Data***

The use of video and other related data-images of student work, transcripts of dialogue, test scores, observational notes, and the like-opens up new areas of inquiry and gives us new ways of investigating traditional questions. However, the expanding role of video and other accompanying multimedia data also presents new challenges in data acquisition, management, analysis, and dissemination. The Digital Insight project www.wcer.wisc.edu/digitalinsight/ is automating elements of video data analysis that are currently "handcrafted."

Digital Insight provides a model for managing the entire research process as well as providing a set of tools designed to allow researchers to answer complex questions based on randomly-addressable large video collections

Video Encodeing

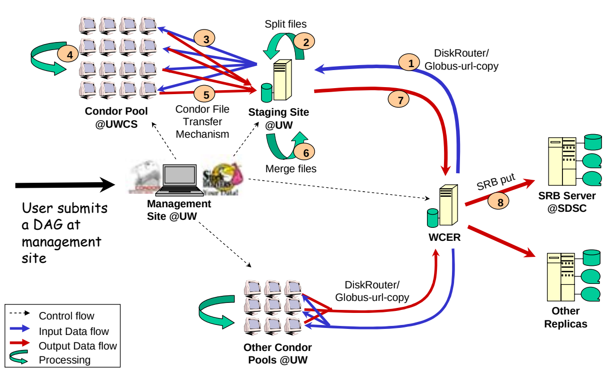
- Steps in Encoding to MPEG-1 and MPEG-2
  - Decode DV to YUV (4X increase in Size)
  - Apply de-noising
  - Apply scaling
  - Encode the video to the desired format
  - Extract audio from the DV
  - Encode audio to mp2
  - Multiplex audio and video into the final mpeg file
- Tools Used
  - MJPEG tools, SMIL utils, libdv,ffmpeg
- Time Taken For Processing
  - 2 hours for 1 hour mpeg-1 video at 320x240, 9 ½ hours for mpeg-2 at 640x480 on 2Ghz Pentium 4

www.cs.wisc.edu/condor

(more) Video Encoders

- Steps in Encoding to MPEG-4
  - Dump the audio stream to file
  - Encode audio to mp4
  - Perform two pass encoding of DV video to mpeg4 using 'mencoder' (increasing quality for a given fixed size)
  - Multiplex mp4 video and audio into a mp4 container.
- Tools Used
  - MPlayer, Mencoder, FAAC, MPEG4IP tools,libdv
- Time Taken For Processing
  - 2 ½ hours on 2 GHz Pentium 4

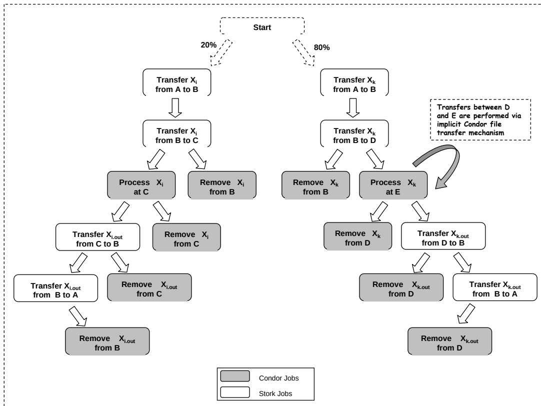
www.cs.wisc.edu/condor



The Palomar Digital Sky Survey (DPOSS)

✓ We have built a data processing pipeline in order to process ~3 TB of DPOSS data (2611 x 1.1 GB files).
✓The data is transferred from NCSA's UniTree server to Condor pools at UW and Starlight; processed there; and the output files moved back NCSA.
✓The data transfers are scheduled and executed by Stork. Computational jobs are scheduled and executed by Condor. The whole process is managed by DAGMan.

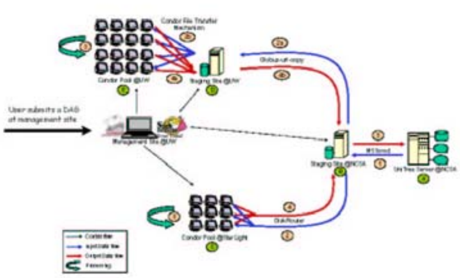
www.cs.wisc.edu/condor



DPOSS Data

- Without any parallelism, end-to-end transfer & processing of a single image file (1.1 GB) takes 14 minutes. 5.5 minutes of this is spent for stage-in, 8 minutes for processing the data, and 30 seconds for stage-out.
- With the current configuration and parallelism level, end-to-end processing of a single image file in average takes only 3 minutes.
- You click only one button, and all of your ~3TB data will be processed and the results will be sent back to you in less than 6 days.
- This will be done fully automated, reliably, and without any human interaction.

www.cs.wisc.edu/condor



www.cs.wisc.edu/condor

Problem

How big?

Solving

When?

Solution Time

determines

Problem Size

How can we accommodate

an unbounded

need for computing with

an unbounded

amount of resources?

www.cs.wisc.edu/condor

Security challenges for the future*

Brian King
Purdue School of Engineering and Tech., IUPUI
briking@iupui.edu

*Supported by the Lily Endowment and IU Pervasive Technology Labs

Security Challenges for the future

The Institute for Information Infrastructure Protection (I3P)* focusing on cybersecurity and information infrastructure research and development has identified key research areas they are:

- Enterprise Security Management
- Trust Among Distributed Autonomous Parties
- Discovery and Analysis of Security Properties and Vulnerabilities
- Secure System and Network Response And Recovery
- Traceback, Identification, and Forensics
- Wireless Security
- Metrics and Models
- Law, Policy, and Economic Issues

**http://www.thei3p.org/documents/2003_Cyber_Security_RD_Agenda.pdf*

Enterprise Security Management

- Pieces of the information infrastructure are operated by single entities yet they are interconnected
- Research needs to continue to study ways to integrate several security mechanisms into a consistent capability for managing both access of resources and the use of enterprises

Trust Among Distributed Autonomous Parties

- Individuals, organizations, and entities need to be able to establish a “relationship over this interconnected world (*cyberspace*)"
- We need to be able to do this without having a resource like a *“trusted central authority*”
- Research needs to continue to focus on methods to provide seamless ways for parties to establish “relationships” over this interconnected world

Discovery and Analysis of Security Properties and Vulnerabilities

- Information infrastructures are LARGE and COMPLEX, pieces of the infrastructures are located in:
  - Hardware Software Firmware Multimedia
- Research needs to continue to focus on understanding the different designs of these information structures—look for design flaws, incompleteness, anything that may lead to understanding their potential security vulnerabilities

Secure System and Network Response and Recovery

- Information infrastructures are complex and large — *response from* and the *ability to recover* from *attacks* are affected by the size and complexity
- Further research needs to continue on improving their survivability skills and developing proactive intrusion detection

Traceback, Identification, and Forensics

- After an attack has been initiated on a system—the organization needs to make an appropriate response and address this attack—including determining *who, what, where* and *when*
- Research needs to identify improved methods that allow these organizations to trace such attacks

Wireless Security

- Wireless networks have become essential in our ability to deliver services and resources.
- Unfortunately the devices that are interconnected in this wireless world are DIVERSE, such devices include laptops, PDAs, cellular handsets, sensors, …..
- It is difficult to initiate an *all-in-one solution* to the security concerns that can fit the potential security vulnerabilities when working within the wireless world on “lightweight devices”

Metrics and Models

- We rely on our information infrastructures. To gauge that these infrastructures are adequately protected and secured we need to develop acceptable levels of risk.
    - what are they?
    - don't risks differ for some?
- The basis of these risk levels need to be grounded on well analyzed models and metrics for security and so further research needs to be developed to provide such models

Law, Policy, and Economic Issues

- Decisions affecting the security of information infrastructures are often made in a mind-set in which one either misunderstands or lacks enough complete information concerning the different issues of economics, laws, regulations and government policies
- Research needs to be done to magnify the importance of the cybersecurity problem and the complex relationship of the issues and organizations that affect the protection of the information infrastructure

Some security challenges we are focusing on at IUPUI

- Enabling secure applications
    - Electronic voting & electronic laws—verifiable democracy
- Lightweight security/cryptography
    - Public-key cryptography
    - Key management
- Security applications/securing the delivery of information
    - Secure transmission of images
    - Secure M-commerce

Lightweight security/cryptography

- Network technology advancements like **I-light** have provided greater network connectivity and bandwidth.
- The greater connectivity together + advancements in hardware technology

compute when needed
with whatever computing devices that are available—no matter how limited

Lightweight security/cryptography

- Today more and more “lightweight devices” devices are required to participate in the computing mainstream
  - – Problems arise
    - Resource starved devices
      - – Small bandwidth, limited processing, limited storage
    - Devices that were not originally developed for the internet
- Traffic (data) that may not previously have traveled through these communication systems will now frequent them. New vulnerabilities will be introduced.

Lightweight security/cryptography

- Public-key cryptography is needed for exchanging keys, authentication, confidentiality algorithms, ….
- Characteristics of public-key cryptosystems
  - – Larger keys e.g for RSA we should be using 1024 bits
  - – Significant more processing
- We look for efficient public-key cryptosystems—as computing becomes more pervasive, more devices are required to support public-key cryptography

Elliptic Curve cryptography (ECC)

- ECC is a **much more** mathematically complex public-key cryptosystem than the integer based (like RSA)
- Why study/why use?
  - – Better utilization of resources
    - Key size
    - Bandwidth
    - Processing speed
      - – Digital signature generation is faster than RSA
      - – Key exchange faster than Diffie-Hellman

Public-key cryptography

key comparison symmetric vs. asymmetric cryptosystems

symmetric key size	ECC over Zp size of p	ECC over GF(2^n) size of n	RSA modulus size
80	192	163	1024
112	224	233	2048
128	256	283	3072

comparison between RSA and Elliptic Curve Cryptography for comparable security levels

	1024-bit RSA	163 bit ECC
certificate size key and signature	over 256 bytes	over 62 bytes
key generation (ms)	285, 630	397
signature generation (ms)	20,208	528
signature verification (ms)	900	1,142

Elliptic Curve cryptography—some recent results from our lab

– New *Galois Field arithmetic* algorithms and a new elliptic curve scalar addition technique using point halving—integral towards achieving the least amount of processing required when using ECC over *Galois Fields*

– New point compression algorithm which is optimal for certain elliptic curves—an integral algorithm to achieve the minimal amount of bandwidth required when using ECC

Security applications/securing the delivery of images

- The research is a collaboration with Paul Salama (IUPUI)
- Typically in an application which requires both compression and encryption, compression will be applied first and then we encrypt the result—in terms of processing we need to visit each member of the bitstream twice
- Our goal was to develop an **efficient secure image encryption** system that will allow one to secure content at a level for which the content demands – dynamically set the security level

Security applications/securing the delivery of images

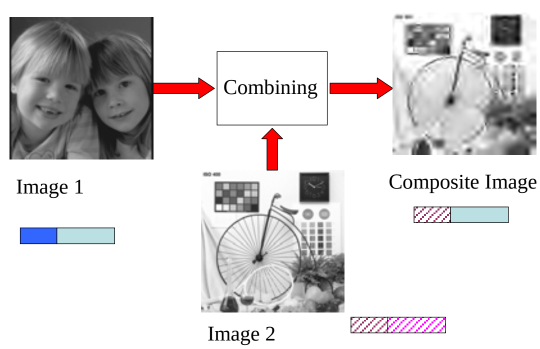
Input stream to SE BOX

Data Dependency

0110010101011010 .....

Image Quality

What is Selective Encryption and why it works?



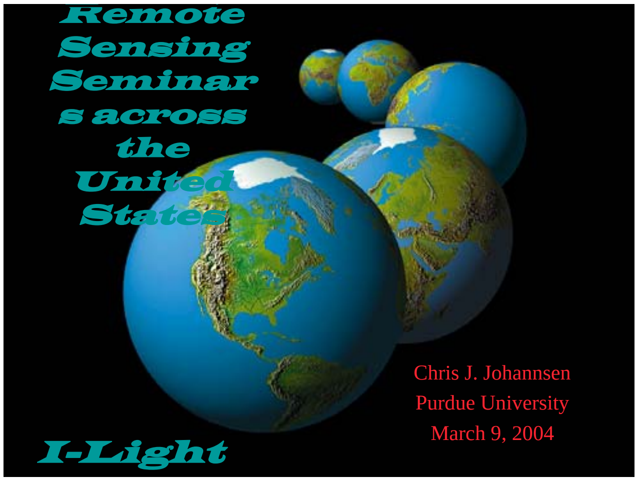
What are the benefits of Selective Encryption?

- The result is that we have
  - less bandwidth usage (since we are compressing)
  - Less processing (rather than processing the bitstream twice we are processing bitstream approximately once)
- The result is an efficient method which both encrypts and compresses
- We can increase security if we wish by increasing the amount of data that is encrypted

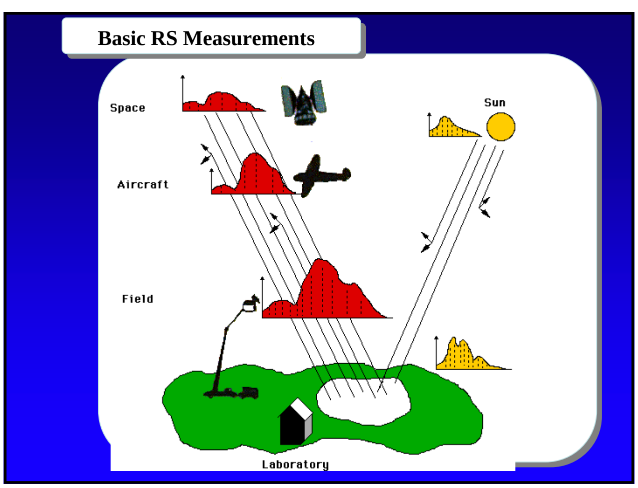
Conclusion

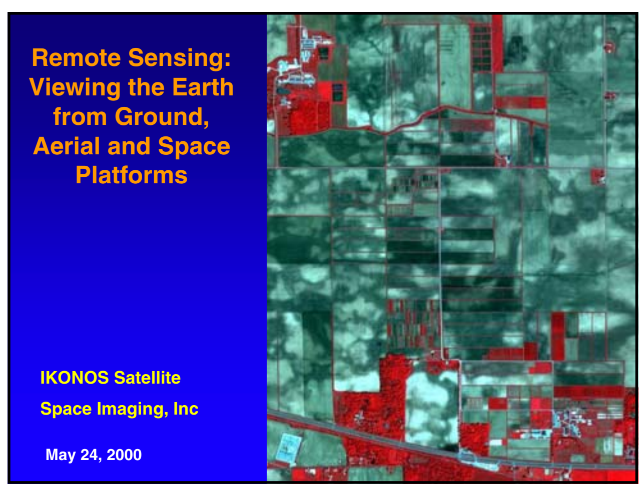
- Information infrastructures are key resources and need to be protected. Security today plays an important role in providing access to these infrastructures to those that are permitted access, as well as securing the services and enterprises that they support.

- As technology improves, smaller lightweight devices will play an increasingly more important role in the computing mainstream and security will become even more critical.


Remote Sensing Seminar s across the United States
Chris J. Johannsen
Purdue University
March 9, 2004
I-Light


In order to apply Spatial Technologies to Agriculture,
we use: "Tools"
Remote Sensing
Geographic Information Systems
Global Positioning Systems
Communications
Spatial Statistics


Basic RS Measurements
Space
Sun
Aircraft
Field
Laboratory


Remote Sensing: Viewing the Earth from Ground, Aerial and Space Platforms
IKONOS Satellite
Space Imaging, Inc
May 24, 2000

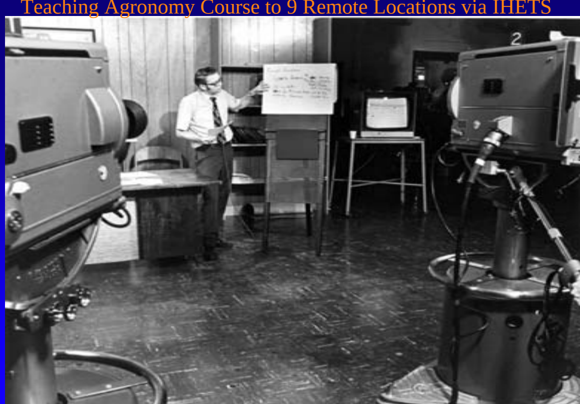
Conducting TV Interviews

Background - Extension opportunities with radio and TV

Filming the Corn Blight Watch Experiment

Background - Filming the Corn Blight Watch Experiment

Teaching Agronomy Course to 9 Remote Locations via IHETS



Background - The “Agricultural Tonight Show”

A “Remote Experience”

Background -Taking a question from a student at Ft. Wayne

Opening PowerPoint - 1998

AGRY 598G

Remote Sensing Seminar

Credit: 1 hour

Instructor: Chris J. Johannsen

3-353 LILY, 494-6248

August 31, 1998

Objectives of Remote Sensing Seminars

Provide a forum for graduate students to:

- interact with academic, aerospace and government personnel working with remote sensing technology
- hear the latest information about remote sensing technology
- provide opportunities to explore future career paths involving remote sensing technologies.

Sequence of the Growth of the RS Seminars

1998 and 1999 Agronomy Department Seminar

2000 Indiana State University - IHETS Connection

2001 Added Mississippi State University and Stennis Space Center

2002 Added the University of Nebraska

2003 Added EROS Data Center and SD School of Mines

2004 Conducting special Earth Observation Seminars with the Katholieke Universiteit Leuven (KUL) Belgium

2000 Remote Sensing Seminar Students and Department

Name	Department
Marshall Beatty	Agronomy
*Greg Blumhoff	Agronomy
*Richard J. Caldanaro	Civil Engineering
*Paul Carter	Agronomy
*Matt Coon	Soc/Anthropology
Krysten DeBroke	EAS
*Peng Du	Agronomy
*Bruce Erickson	Agronomy
*Dan Getman	EAS
*Brenda Hofmann	Agronomy
William King	Forestry & NR
Kyoung Jae Lim	Ag & BioEngineering
*Huagen Liu	EAS
*Pablo Mercuri	Agronomy
*Keith Morris	Agronomy
*Suresh Muthukrishnan	EAS
*Kent Ross	Agronomy
*Maggie Sullivan	Ag & BioEngineering
Van Etten, Jeffrey	Horticulture/LA
Wallace, Robert M.	Civil Engineering
*Wenchun Wu	Forestry & NR

Note: students registered from 7 departments within the Schools of Agriculture, Engineering and Science.

*Registered for credit

Information Extraction Principles for Hyperspectral Data

David Landgrebe
Professor of Electrical & Computer Engineering
Purdue University
landgreb@ecn.purdue.edu

Outline

- A Historical Perspective
- Data and Analysis Factors
- Hyperspectral Data Characteristics
- Examples
- Summary of Key Factors

More then 15 Purdue faculty provided seminars during the series

Speakers Participating in RS Seminars

George Buttner, Head, Environmental Applications, Remote Sensing Centre, Budapest Hungary "CORINE Land Cover Program in Europe"

David Shaw, Director, Remote Sensing Technologies Center, Mississippi State University, "Remote Sensing: From Pandora to the Oracle"

Ryan Jensen, Geology, Geography and Anthropology, Indiana State University, "Integrating artificial neural networks and TM data to estimate forest LAI"

Donald Runquist, University of Nebraska, "University of Nebraska Airborne Remote Sensing Program"

Speakers (Continued)

+Kass Green, President, Pacific Meridian

+Patrick Willis, National Ag Statistics Service, USDA

+Bruce Davis, Chief Scientist, Commercial Remote Sensing Program, Stennis Space Center

+Roger King, Chief Technologist, Applications Division, NASA Headquarters

+Michael F. Goodchild, Uni. of California, Santa Barbara, CA

+Kevin Gallo, NOAA/NESDIS, USGS EROS Data Center

+Mark Vanacht, President, AG Business Consultants, St. Louis

+Tina Cary, Cary and Associates, Longmont, CO

+Joanne Gabrynowicz, Director, Nat. RS and Space Law Center, University of Mississippi

Speakers (Continued)

Tim Warner, Geography and Geology Department, West Virginia University, Morgantown, WV

Todd Helt, President, Telemorphic Inc. San Francisco, CA

Michael Renslow, Manager, Spencer B. Gross, Portland, OR

Bryan Pijanowski, Michigan State University, Lansing, MI

Susan Carson Lambert, Governor's Office for Technology, Commonwealth of Kentucky

Harold Reetz, Jr., Midwest Director, Potash Phosphate Institute

The Future

- Invited IndianaView and AmericanView Universities to participate in seminars for Fall 2004
- Expect to provide seminars during Fall and Spring Semesters
- Discussing the possibilities of teaching AGRY 545 Remote Sensing of Land Resources to the United Arab Emirates
- Will conduct Graduate Student Thesis Defenses with KULeuven using the Purdue Communications Services

& Math

Science Education using Networked IT

I-Light 2004

Randy Heiland, Assoc Director
Scientific Data Analysis Lab (IUPUI)
heiland@indiana.edu

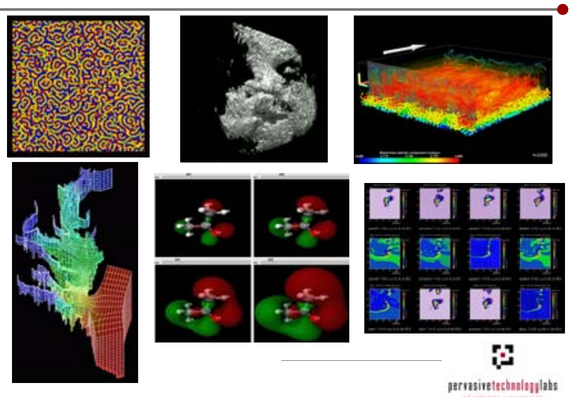
pervasivetechnologylabs
AT INDIANA UNIVERSITY

Speaker's background

- **B.S. Comp Math, Eastern Illinois U.**
- **M.S. Comp Science, U. of Utah**
- **M.A. Mathematics, Arizona State U.**

- **14 yrs scientific computing; data visualization/analysis for HPC (NCSA/UIUC, Gov't Labs)**

Snapshots of Past Projects



NSF TeraGrid ⇦ IP-grid

- Large simulations
- Distributed clusters
- High bandwidth
- Data repositories
- Visualization ⇦

Scientific Data Analysis Lab

- Self-explanatory? ☺
- Scientific visualization/analysis for HPC; data mining
- Collaborations with Indiana+ researchers/businesses
- Education/Outreach

Motivation for this talk

- HPC scientists ⇔ K-12
- Opportunity
  - Came to IUPUI summer `03
  - Internet/Web/Grids
- Middle age
  - learning ⇨ teaching/sharing
- Self-interest (kids)

Broad Categorization

Networked computer-based tools for science & math education

Bandwidth

- Passive
- Active ~ Exploratory
- Interactive ~ Participatory

Passive (dist'd computing)

SETI@home The Search for Extraterrestrial Intelligence

Folding@home

GIMPS The Great Internet Mersenne Prime Search

ENTROPIA PC GRID COMPUTING

PARABON COMPUTATION

Active ~ Exploratory

Online encyclopedias, applets, games

mathworld.wolfram.com

science.howstuffworks.com

sdss.org

Interactive ~ Participatory

Online gaming?

-Multiplayer/broadband
-chessclub.com
-DARPA FutureMAP

Local/regional science projects?
-eco monitoring
-americanbirding.org

“clickworkers” – NASA, public participation for classifying Mars craters

Targeting K-6: squeakland.org

- Free/open graphics tool
- Fun, interactive
- Math & science concepts through simulations
- Alan Kay – “father of the personal computer”

Advisory Board

JP Barlow, G. Bell,
R. Dawkins, M. Minsky,
S. Papert, M. Resnick,
D. Hillis, L. Smarr,
Quincy Jones …

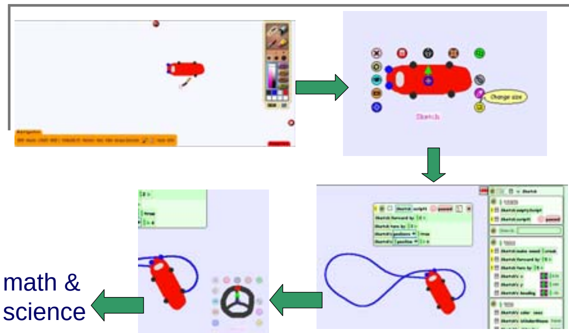
Squeak Media

Book/Tutorial

DVD (Ball State/Lilly)

-Articles & references
-Online community

Squeak Project #1: drive-a-car



math & science

“Computers, Networks and Education”

- Alan Kay, SciAm Sept ’91
- Computers as supporting media
- Value of computer simulation for education
  - Visual explanations
  - “what if” scenarios
- Value of networked computers
  - Benefits of peer interaction

Summary

- How can we use I-Light/Grids for K-16 science/math education?
- Need student-friendly tools, data, projects
- Squeak: an interesting, open tool; collaborative/networked squeak
- Like to hear your ideas

NCN NETWORK for COMPUTATIONAL NANOTECHNOLOGY

Using I-light and the Purdue Nanohub Computing Resources to Run Computationaly Intensive Codes in Nanotechnology

Sebastien Goasguen
IT manager

Mark Lundstrom
Director

Gerhard Klimeck
Technical Director

1

NCN NETWORK for COMPUTATIONAL NANOTECHNOLOGY

Vision

To help move nanoscience to nanotechnology by connecting experiment, theory, simulation, and computation in a new way ***and*** to become a resource to the NNI.

NNIN > NASA URETI > NCN > Ilight

Mission

- Research
- Education
- Infrastructure

2

NCN NETWORK for COMPUTATIONAL NANOTECHNOLOGY

Outline

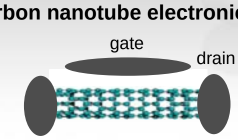
- The Network for Computational Nanotechnology…A research center.
- Building a community and providing services…A user facility.
- Using Ilight and Teragrid to enhance research and education…An infrastructure

3

NCN NETWORK for COMPUTATIONAL NANOTECHNOLOGY

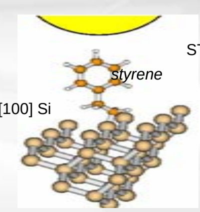
Nanoelectronics

carbon nanotube electronics



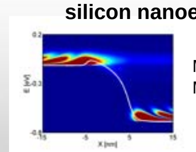
Theory: Cho, Leburton, Lundstrom, Roy
Experiment: Dai

molecules on silicon



Theory: Datta, Ratner, Roy
Experiment: Hersam (NASA)

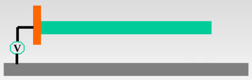
silicon nanoelectronics



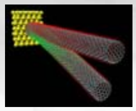
MARCO
MSD Center

4

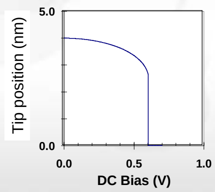
NEMS



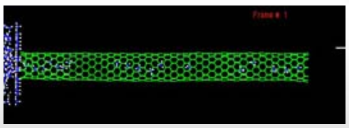
QM Models for NEM switches
Theory (Aluru, Hess, Dutton)
Experiments (Lyding, Liu)



Compact models and circuits for arrays of NEMS



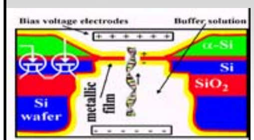
Theory (Dutton, Aluru)



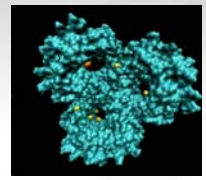
Fluids through nanotubes and nanopores
Theory (Sinnott, Murthy, Aluru)
Experiments (Shannon, Bohn)

5

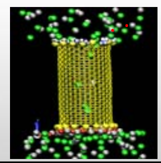
nano-bio



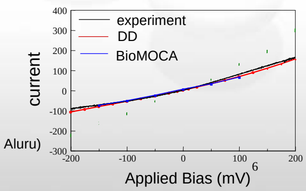
DNA sequencing using artificial nanopores
Theory (Hess, Leburton, Schulten)
Experiments (Timp)



Ion Channels in Biological Membranes
Theory (Hess, Ravaioli, Schatz)



Ion transport in artificial channels
Theory (Hess, Schatz, Aluru)
Experiments (Lyding)



6

The NCN research is about

-**research** that makes a difference

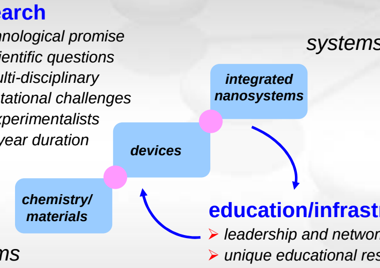
nanoelectronics, NEMS, nano-bio
tight links to experiment
advancing science, exploring technologies
new theory
new computational algorithms and approaches
open source software

7

More than Research...

research

- significant technological promise
- challenging scientific questions
- multi-scale, multi-disciplinary
- serious computational challenges
- tight links to experimentalists
- clear focus, 3-year duration



education/infrastructure

- *leadership and networking*
- *unique educational resources*
- *new theory and approaches*
- *new software*
- *collaborative services*
- *web-based computing*

nanoelectronics
.... ***nano-bio***
NEMS

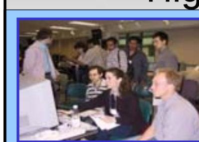
8

Community Building

- Bringing people together
- Offering valuable content
- Fostering debates
- Sharing software and education material
- Use of new technology (Access Grid, Ilight…)
- On-line simulation

9

High Value Content



"Introduction to Computational Nanotechnology"
June 7-18, 2004 Beckman Institute, UIUC

"From Atoms to Transistors"
Supriyo Datta

- On-line courses, streaming video and lecture notes
- On-line seminars, streaming video, slides and forums
- On-line simulation for summer school, labs and lectures on-line.

Fundamental Limits of Digital Computation

1) Victor Zhirnov, SRC
2) Craig Lent, Notre Dame
3) Kostya Likharev, SUNY-SB
4) Vwani Roychowdhury, UCLA
5) …

10

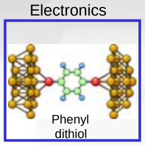
Sharing research codes

http://nanohub.purdue.edu

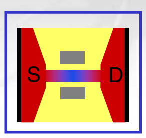
In 2003:

20 tools / 987 users / 86,152 jobs

Molecular Electronics



Si CMOS



MolcToy Hückel-IV 2.0 *ab initio* DFT	FETToy nanoMOS 2.6

2003: 261 on-line users (9,334 jobs)
Total downloads: 772

11

More than great content…

- A real user facility, nanohub.purdue.edu
- Providing on-line simulation to the nanotechnolgy community
- Using tomorrow's cyberinfrastructure…
  - … Ilight and Teragrid

12

Online simulation

- Need for applications
- Need for middleware
- Need for hardware
- Need users …..

PUNCH 5.0
network computing
middleware

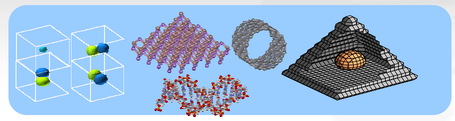
1) software applications → 2) computing resources

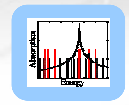
3) Results

13

Example of Application

NEMO 3D (Klimeck, et al., JPL)





Atomic Orbitals size: 0.2nm + Structure → Nanoscale Quantum States (Artificial Atoms, size 20nm) → Designed Optical Transitions **Sensors**

Computational Cost: 1-10 hours on 20-40 CPUs running in parallel

- ***3D quantum simulation is limited by computation***
- ***We want to allow users to run these type of applications transparently***

14

Infrastructure Middleware

next-generation network computing

J.A.B. Fortès
Univ. of Florida

http://invigo.acis.ufl.edu

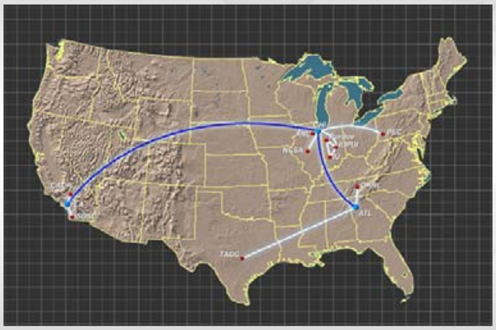
-Access to virtual machine on demand
- Shared workshpace for collaboration
- Enhanced security and use of open source packages
- Grid aware, backends protocols for PBS, Condor, Globus

15

Some Hardware...

16

Some more...



17

The NCN goals

- Every three years state of the art production level applications are made available on the nanohub
- Some of these are and will be computationally intensive but this should not be an issue for great research and great teaching/learning
- Significant compute cycles will be needed through the middleware infrastructure.
- 10,000 users by 2007

18

"NCN has been designed to be the lead center of excellence for simulation at the nanoscale in the US."

Mihail Roco
NSF Senior Advisor for Nanotechnology

19

The Biocomplexity Institute at Indiana University

James A. Glazier

http://www.biocomplexity.indiana.edu

Key Participants

James Glazier (Physics, Informatics and Biology)
Geoffrey Fox (Informatics and Computer Science)
Andrew Lumsdaine (Computer Science)
Sima Satayeshgar (Physics)
Rob de Ruyter (Physics and Psychology)
John Beggs (Physics)
Olaf Sporns (Psychology)
Filippo Menczer (Informatics and Physics)
Santiago Schnell (Informatics and Physics)
Alessandro Vespignani (Informatics and Physics)
Alessandro Flammini (Informatics and Physics)
Katy Borner (Informatics and Library Science)
Bill Saxton (Biology)
David Stocum (IUPUI)
Loren Field (IUPUI)
Merv Yoder (IUPUI)
Matthew Grow (IUPUI)
Purdue
Notre Dame

Issues

Contemporary Biomedical problems require input from engineers, molecular, cell biologists, physicists, computer scientists,…

Require Cooperation of Experiment and Theory (Data Organization and Sharing)

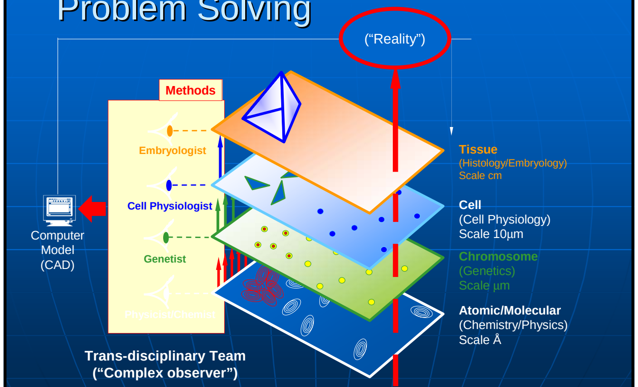
Problems are intrinsically multiscale

Different subdisciplines speak different languages (Workshops, Cross Training, Remote Education, Course Sharing)

Goals

- To understand in a quantitative and predictive way the complex patterns and organization that arise in living organisms at length scales from molecular to organismal and population.
- To develop Integrated Interdisciplinary Approaches to Complex Biological Problems at All Length Scales
- To connect Theoretical, Experimental and Applied Research
- To prepare undergraduate, graduate and postdoctoral students for the challenges of twenty-first century biology, which require a merging of fundamental biological understanding and methods with physical, computational, and mathematical approaches.
- To Develop Software of General Applicability for Addressing Key Biological Problems

Hierarchical Levels/ Multiscale Problem Solving



Slide 4

Projects

Limb Development and Regeneration
Cardiac Development and Function
Systems Neuroscience
Artificial Life and Populations
Bacterial Pattern Formation and Biofilms
Biological Networks
Structure and Function of the Cytoskeleton

Development and Regeneration

- **How** does the pattern of gene expression act through physical and chemical mechanisms to result in the structures we observe? Genetics is just the beginning.
- Same mechanisms occur repeatedly in different developmental examples.
- Begin by using phenomenological descriptions. In many cases very complex pathways have fairly simple effects on cell properties/behaviors under conditions of interest.

Key Goals

- Understand how cellular decisions are made at the molecular level.
- Elucidate the structure and properties of molecular and control networks.
- Understand the development of cell polarity.
- Explain how cells sense and respond to their external environment.
- Determine the mechanisms that give rise to large-scale cell migration and the patterning of differentiation.
- Understand how cell and extracellular matrix (ECM) properties interact with tissue geometry to give rise to specific function.

- Develop general purpose organo(re)genesis simulations which can:
  - Predictively model normal development and regeneration.
  - Help design strategies to enhance healing and regeneration.
  - Predict effects of chemical intervention.
  - Predict effects of mutations.

Issues

- Many packages for simulating function of single cells: M-Cell, E-Cell, Virtual Cell, Entelos,…
- Finite Element Models Available for Simulating Tissues as Continua
- Much Less Available for Simulating at the Cell Level

Cellular Potts

- Extension of Ising model
  - Instead of just two spins (indexes), introduce as many as there are cells
- Thus, 3 cells $\Rightarrow$4 possible values of spins (each cell identified by its spin/index)

Typical Array Size 1024^3
Corresponds to 2mm cube
Need much larger

				2					
	1	1	2	2	2				
1	1	1	1	2		3			
1	1	1	1		3	3	3		
	1	1		3	3	3	3	3	
					3	3	3		
						3			
E	C	M	=	0					

Cellular Dynamics: CPM

- Energy minimization formalism
  - extended by Graner and Glazier, 1992
- DAH: Contact energy depending on cell types (differentiated cells)
- Extensions:
  - J_cell_cell is type dependent
  - Other terms: Cell volume, Chemotaxis/Haptotaxis

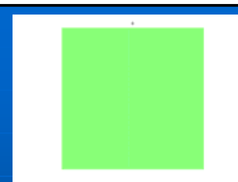
$$E = E_{\text{contact}} + E_{\text{volume}} + E_{\text{chemical}}.$$

$$\begin{aligned} E_{contact} &= \sum J(\sigma,\sigma')(1-\delta(\sigma,\sigma')), \\ E_{volume} &= \sum \lambda_\sigma (v - v_{target})^2 + \sum \lambda'_\sigma (s - s_{target})^2, \\ E_{chemical} &= \sum \mu(\sigma) C(\vec{x}). \end{aligned}$$

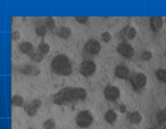
- Metropolis algorithm: probability of configuration change

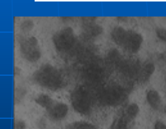
$$\begin{aligned} P(\Delta E) &= 1, \ \Delta E \le 0 \\ P(\Delta E) &= e^{-\Delta E/kT}, \ \Delta E > 0 \end{aligned}$$

Ectopic Fibronectin Transfection Enhances Condensation

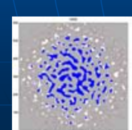


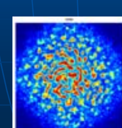
Control

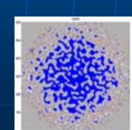


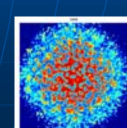


Treated

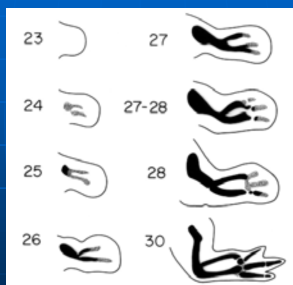








Skeletal Pattern Formation: stages in a chicken limb bud



REGENERATION & DEVELOPMENT

REGENERATION

PHASE I
WOULD HEALING

PHASE II
DEDIFFERENTIATION

PHASE III
REDEVELOPMENT

DEVELOPMENT

PHASE I
LIMB FIELD DETERMINATION

PHASE II
INITIATION

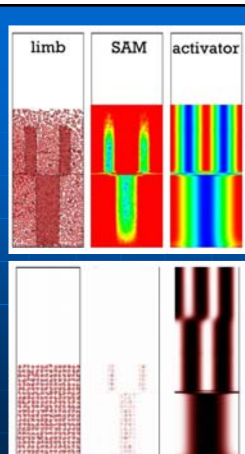
PHASE III
GROWTH & PATTERN FORMATION

Susan B. (Int J Dev Biol. 46(7):887-96, 2002)

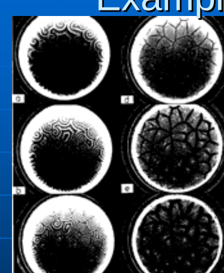
Limb Bud Simulation

Three fields of interest:
- Cells are represented by field of Potts pixels

- Cells:
  - can move around,
  - change type
  - secrete chemicals,
  - grow and mitose
- "Activator" (TGF-β):
  - chemical field governed by RD equations
  - triggers cell differentiation,
  - induces secretion of morphogen SAM (fibronectin)
- "SAM" (fibronectin):
  - secreted by cells
  - Cells haptotax to it

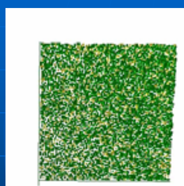


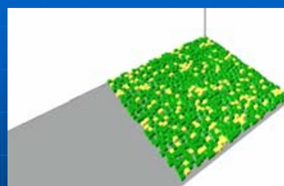
Example: Dictyostelium

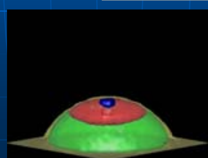


Cellular Automata Simulation

(from *Nick Savill and Stan Maree*)





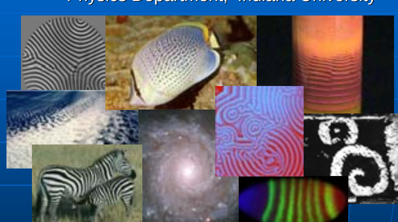


Issues

- Subcellular Models—Coupled ODE models of genetic and metabolic networks, signal transduction
- Tissue Scale Models--Finite Element Models
- How to Connect Models? Propagate Parameters?
- Potts Model (MonteCarlo) doesn't parallelize easily. How to do Large Objects?
- Visualization of 3-D multimode data
- Need Many Runs for Statistics and Parameter Space Searches
- Steering?
- Real Time Calculations?

Pattern Formation in Chemical and Biological Systems

Sima Setayeshgar
Physics Department, Indiana University

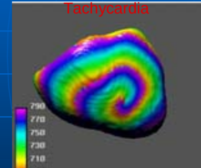


- Common processes underlie natural pattern-forming systems with different underlying microscopic descriptions.
- Theoretical and computational approaches are focused on understanding generic properties of controlled and reproducible experimental systems.
- **Extension to more complicated and physically relevant systems (weather, heart, …)**

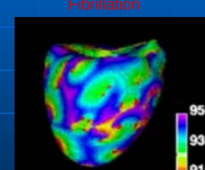
Scroll Waves in Anisotropic Excitable Media with Application to the Heart

Big Picture: What are the mechanisms of transition from tachycardia to fibrillation, and how can we better control them?

Tachycardia



Fibrillation



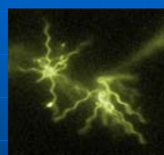
Courtesy of Sasha Panfilov, University of Utrecht

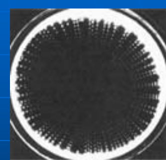
Paradigm: Breakdown of a single spiral to disordered state, resulting from various mechanics of spiral instability (Very Large Scale Simulations)

Chemotaxis

From individual to collective behavior in *E. coli* …

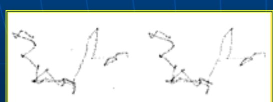
Paradigm for internal and external
biochemical signaling





courtesy of Howard Berg group

Biased three dimensional random walk



Bialek and Setayeshgar, *submitted*, physics/0301001
Setayeshgar, et al., *submitted*, physics/03081001

Access and Compilation of Networks Data

Connection to Potts Model

Attractor states and criticality in a network of living cortical neurons
John Beggs

Examine the fundamental building block of cortex

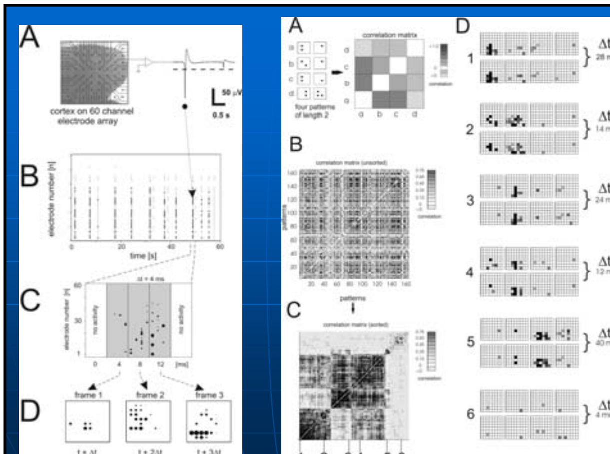
- Study a local network in isolation
- 50,000 to 100,000 neurons (Very Large Data Sets—1000 electrode, future 10,000 electrode arrays)
- Preserve cell types
- Preserve local connectivity
- Record spontaneous activity
- Record from many neurons simultaneously
- Record from neurons that are likely to be connected to each other

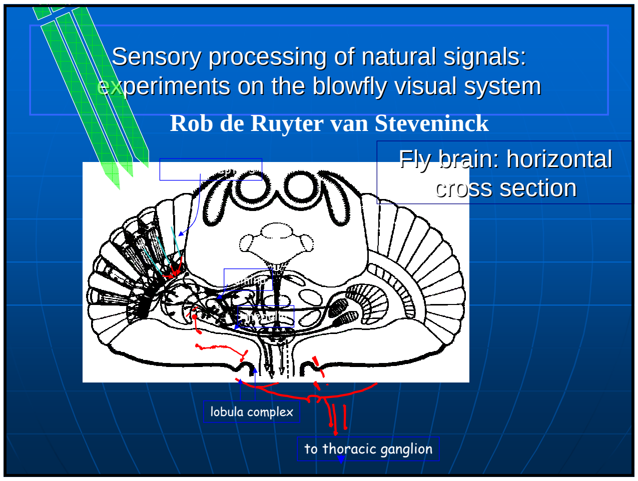
Methods to quantify network performance:

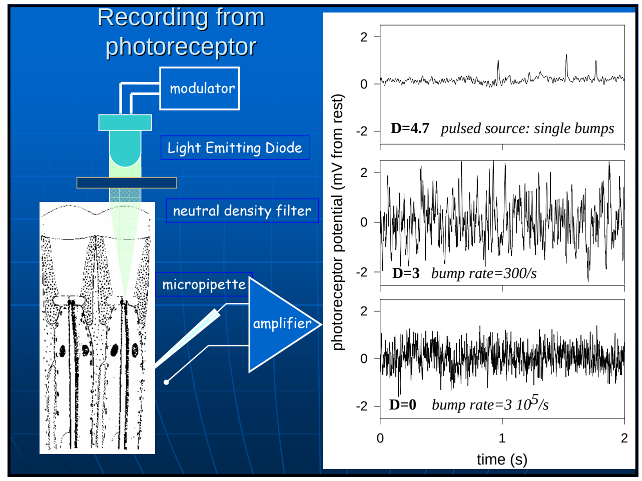
- Number of statistically significant patterns
- Temporal precision of patterns
- Information content of patterns
- Duration of patterns
- Ideal for pre-post manipulations

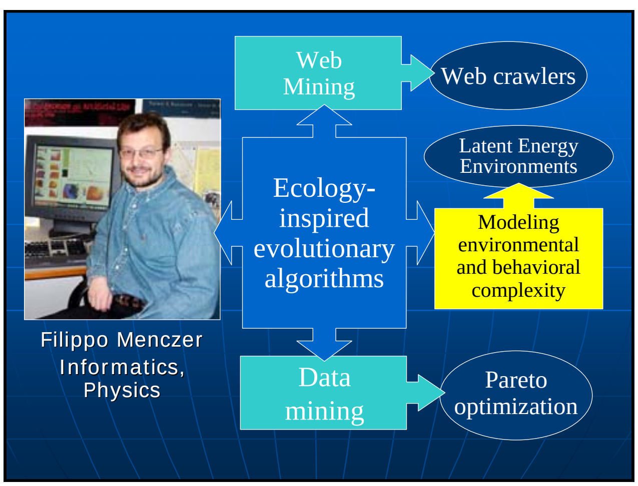
Levels at which to study drug effects:

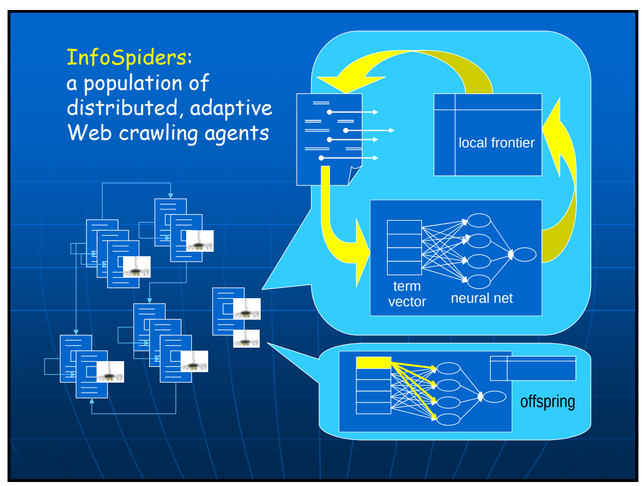
- Synaptic: Many good studies
- Network: Few studies; technology gap
- Whole Animal: Many good studies



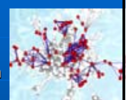

Sensory processing of natural signals: experiments on the blowfly visual system
Rob de Ruyter van Steveninck
Fly brain: horizontal cross section
lobula complex
to thoracic ganglion


Recording from photoreceptor
modulator
Light Emitting Diode
neutral density filter
micropipette
amplifier
photoreceptor potential (mV from rest)
D=4.7 pulsed source: single bumps
D=3 bump rate=300/s
D=0 bump rate=3 10^5/s
time (s)

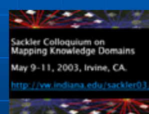

Web Mining
Web crawlers
Latent Energy Environments
Ecology-inspired evolutionary algorithms
Modeling environmental and behavioral complexity
Data mining
Pareto optimization
Filippo Menczer
Informatics, Physics


InfoSpiders: a population of distributed, adaptive Web crawling agents
local frontier
term vector
neural net
offspring

InfoVis Lab Overview (Katy Borner)



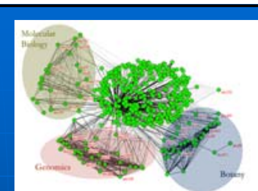
- Process Models of Scientific Structure and Evolution
  Small world & scale free networks, information diffusion
- Visualizing Knowledge Domains
  specific domains & 'backbone of science'



- Mapping Aging Research
  Input – output studies
- Visual Interfaces to Digital Libraries

Details are available at
http://ella.slis.indiana.edu/~katy/research/

PNAS Document Space in 1995



- **Oncosifter**–Personalized Cancer Information-Filters cancer related news and medical information from Medlineplus and Cancer.gov. Provides hierarchical search & browsing interface.
- **Medical Record Display System**-Provides interaction with the graphical widgets including effective zooming and filtering of patient data.
  http://ella.slis.indiana.edu/~kmane/courses/I579/projects/project4/proj4.html
- **Citeseer**, NSF, NIH, Community of Science Funding, PNAS, Medline-the work of formatting the data and uploading it to an Oracle database is in progress
- **XML Flybase**-A template xml version of the Flybase gene data display
  http://ella.slis.indiana.edu/~kmane/courses/I595_xml/xml_finalproject/fproj_submitted/gene_dpp.xml

The Future

- Industrial and Medical Collaboration
- New Hires in Techniques, Genetic Regulation and Developmental Biology
- Coordination with other IU and Purdue Centers

The MxN Problem: Coupling Parallel Components

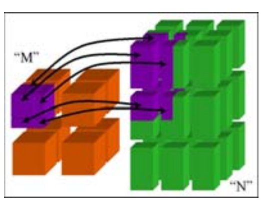
Randall Bramley
Felipe Bertrand
Computer Science Department
IU-Bloomington

If You Don't Do Parallel Programming

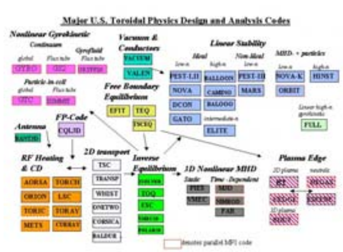
- Parallel programs are ones that partition a coupled problem and solve simultaneously on multiple processors
- Only scalable approach is distributed memory programming: partition data across processors
- Requires sending messages from one processor to another; universal standard interface is MPI (Message Passing Interface)

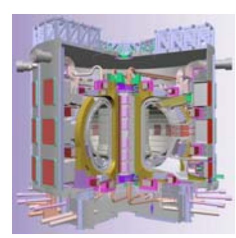
The MxN Problem

- Transfer data from a parallel program running on M processors to another running on N processors.
- M and N may be incommensurate
- May require complex all-to-all communications, data redistribution



Example Application: Fusion Energy Sim

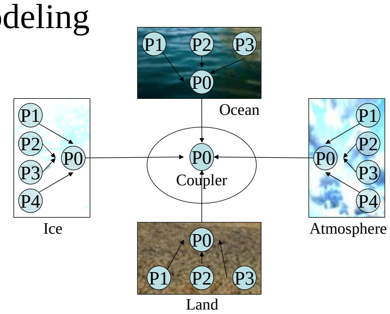




Example Application: Fusion Energy Sim

- Many existing codes for external RF heating, internal MHD, transport, materials properties
- Different groups maintain, evolve those codes
- Finite element, particle-in-cell, finite volume, spectral methods, finite differences all used
- 10^{15} range of time scales, 10^{9} range of spatial scales

Example Application: Climate Modeling



- Models are parallel, but intermodel coupling is serial

Characteristics

- Interacting codes/components are complete simulations of some subphysics of a potential larger simulation
- Must use legacy codes, in variety of languages
- Some codes are tied to particular architecture or machine
  - Architecture dependencies
  - Licensing
  - Research teams that don't play well with others
- Increasing need for data transfers with sizes nearing that of the compute objects
- Must handle parallel case, cannot dictate numbers of processors used for each code/component

Goals

- **Scalably** connect codes to enable new multiphysics simulations
- Support existing and currently evolving codes
  - Created by teams with disparate research interests
  - Spanning multiple time scales, spatial domains, and disciplines
- Rapid prototyping/testing without extensive rewriting
- Use standard APIs and paradigms familiar to most application area scientists (**MPI**)
- Shield user from complex bookkeeping in data shuffles

Existing Methodologies

- Refactor codes: integrate all components into a single large program
  - More efficient data sharing, tighter coupling
  - Large initial time investment in rewriting programs
  - Requires close coordination between teams: more costly in development, testing and deployment

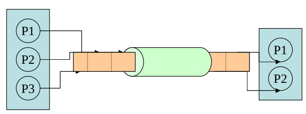
Existing Methodologies

- Use process 0 on all components for data transfer
  - Used by climate models
  - Not scalable (in part leads to new ESMF effort)
  - Worked until now because
    - Used single time stepping "clock", with 10x time scales
    - Interfaces (data transferred) are 2D, simulations are 3D
- Transfer data by reads/writes through files
  - Scalable if parallel I/O used
  - Slow because involves hard drive read/write

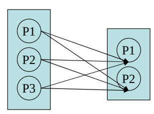
New MxN Solutions: MPI I/O

- Use MPI I/O interface and create *middleware* to transfer data via network
- Provides an easy migration path for existing applications: codes run as before, with some I/O statements added
- Using Argonne's ROMIO, simply substitute a new I/O "test device" that uses net transfers
- Requires M*N connections even if transfer is one to one

MxN via MPI I/O



- From user's viewpoint, the transfer is **logically serialized**

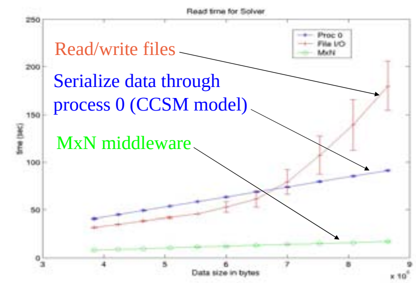


- Actual transfers are carried out in **parallel scalably**

MxN via MPI I/O

- Components live in different MPI instances with MxN device added to ROMIO directory
- Transparency:
  - User need not be aware of the MxN communication
  - Reads and writes data through regular MPI interface
- No change in the source code is required
  - Switch to MxN backend with filename prefix “mxn:”

PDE Solve with Four Components

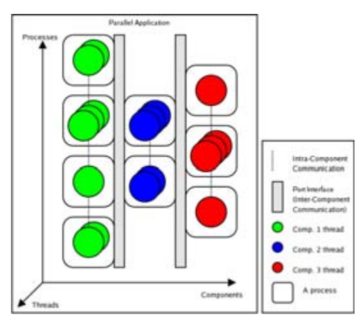


A Second Approach to MxN Problem

- Solve within a DoE **Common Component Architecture** (CCA) framework
- Have the CCA framework provide the necessary data transfer, redistribution
- Protect user from details of parallel bookkeeping
- CCA “design pattern” imposes limitations on remote method invocations.

CCA Components

- Three potential levels of parallelism:
  - Components within applications
  - Processes within components
  - Threads within processes
- These forms of parallelism are independent



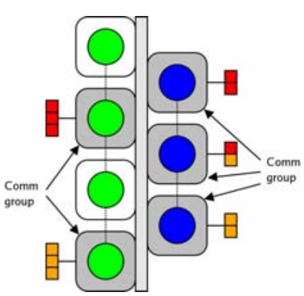
CCA and Distributed Codes

- Existing CCA frameworks are either
  - **Parallel,** but with NxN or Nx1 communications: CCAFFEINE, Dune, Decaff, SciRUN
  - **Distributed**, with serial components: XCAT
- In non-distributed case, one component can "know" data layout and participating processes on the other: use MPI **communicator groups**
- In distributed case, comm group has no meaning at a remote component!

DCA: Distributed CCA Architecture

- Must solve three problems:
  - **Process participation**: which processes invoke a remote function, and which processes must perform it (PRMI = parallel remote method invocation)
  - **Data distribution**: how to describe data distribution on each side
  - **Data transfer and redistribution**: how to shuffle data efficiently between two components
- Synchronization? Deadlock prevention?

DCA Solutions



- **DCA** uses MPI comm groups to establish process participation during remote method calls
  - component making call decides which processes participate
  - on receiving side *all* processes *must* participate.
- This scheme requires **barrier synchronization** at the caller
- These limitations are inherent in the CCA model

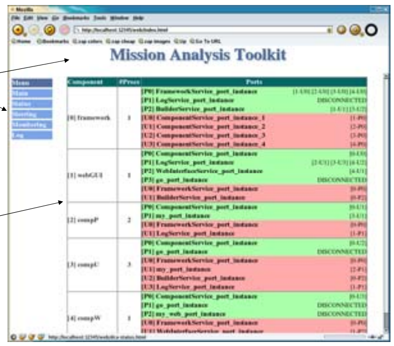
Conclusions

- Scientific computations composed from existing disciplinary codes are increasingly common
- Requires solving the MxN problem: connecting parallel components running on different numbers of processors, at different sites
- All scalable solutions rely on fast networks to carry out transfers in parallel

Web Interface

Application specific
(header and menu)

DCA dynamically
generated frame
(status page)



Measured Response

The Application of Terascale Computing in a Homeland Security Simulation

Prof. Alok Chaturvedi
alok@purdue.edu

Dr. Shailendra Mehta
mehta@purdue.edu

Krannert School of Management

Funded By:
Indiana State 21st Century Research And Technology Fund
National Science Foundation

Measured Response

- Human Aspect in Homeland Security
- Live and Computational Experimentation in Bio-terrorism Response
- Measured Response

People

- Presenter: Tejas Bhatt: **tejas@purdue.edu**
- Prof. Alok Chaturvedi: **alok@purdue.edu**
- Dr. Shailendra Mehta: **mehta@purdue.edu**

Place

- Synthetic Environment for Analysis and Simulation (SEAS) Lab
- Krannert School of Management
- Purdue University

Funding

- Indiana State 21st Century Research And Technology Fund
- National Science Foundation

Presentation

- Need: Human Aspect of Homeland Security
- Project: Measured Response (MR)
- Architecture: Tera-scale Computing
- Event: Measured Response 2003
- Updates: New Models
- Future: Visualization, Timelines

Need

- Capabilities to analyze human responses
- Emphasize citizens and policy makers
- Virtual environment to test and analyze
- Experiment with complex coordination

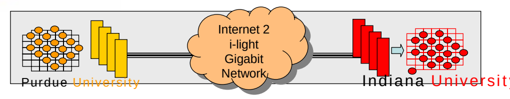
Project: MR

- Virtual Environment
- Emergency Response Systems
- Decision Making Tool

Architecture

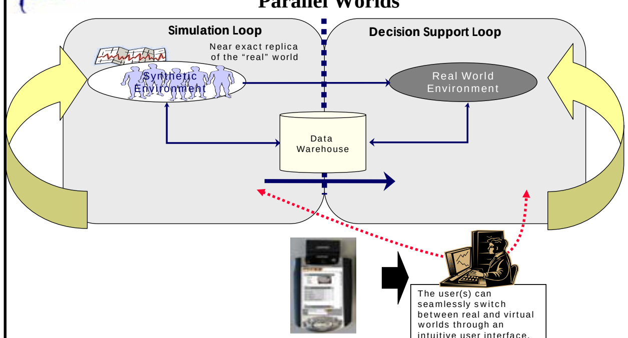
- The Application of Tera-scale Computing in a Homeland Security Simulation
- Connect Palmtops to eventually Teraflops in a Seamless Simulation

Architecture (cont)



- The Real and Virtual World
- Wireless hand-held devices
- Super Nodes at Purdue and IU
- I-Light Gigabit Network
- Distributed Grid Computing

Parallel Worlds



Architecture: Parallel Worlds

- Simulation Loop
- Decision Support Loop
- Common Data Storage

Architecture: Parallel Worlds: Simulation Loop

Simulation Loop

- Replica of the real world
- Behavioral modeling, demographics and calibration
- Time compression
- Supports millions of artificial agents

Architecture: Parallel Worlds: Simulation Loop (cont)

Used to…

- Explore
- Experiment
- Learn
- Analyze
- Test
- Anticipate

Architecture: Parallel Worlds: Decision Support Loop

Decision Support Loop

- Real world environment
- Data collection
- Association
- Trends
- Parameter Estimation

Architecture: Parallel Worlds: Decision Support Loop (cont)

Used to…

- Implement
- Access

Architecture: Parallel Worlds: Common Data Storage

- Real World Data
  - Supply Chain Management (SCM)
  - Enterprise Resource Planning (ERP)
  - Customer Relationship Management (CRM)
  - Data Warehousing
- Simulation Data

Event: MR 03

- Simulation
- Participants
- Demo
- Value

Event: MR 03: Simulation

- Bio-terror Attack
- Population
- Disease

Event: MR 03: Simulation (cont)

- Facilitated Workshop
- Computer Based Simulation
- After Action Review

Event: MR 03: Real-Time Simulation

- Real-Time
  - Information
  - Analysis
  - Judgment
  - Communication
  - Decisions

Event: MR 03: Real-Time Information

Information

- What do we know?
- What more do we need to know?
- Where can we acquire it?

Event: MR 03: Real-Time Analysis

Analysis

- How reliable is our information?
- What does it mean?

Event: MR 03: Real-Time Judgment

Judgment

- What should we do about the situation?
- What can we do about it?
- What resources are available?

Event: MR 03: Real-Time Communication

Communication

- Whom should we inform?
- How?
- Within the government?
- Outside the government?

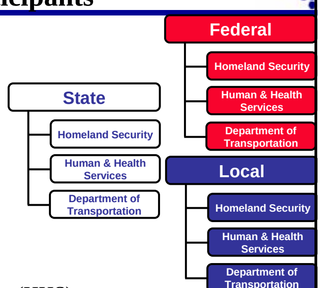
Event: MR 03: Real-Time Decisions

Decisions

- Who makes them?
- What range of options are available?
- What resources are available?

Event: MR 03: Participants

- Three Levels
  - Federal
  - State
  - Local
- Three Agencies
  - Homeland Security (HLS)
  - Health and Human Services (HHS)
  - Department of Transportation (DOT)



Event: MR 03: Demo

Demonstration

- Briefing Screens
- Situation Monitors
- 2D Simulation
- Video

Event: MR 03: Value

The Value: The value of any simulation is not the technology used for the simulation, but the lessons learnt from its outcome

The Outcome: We found better cohesion among participating first responder agencies

Updates

- Modeled Smallpox
- Social Contagion Model
- Leadership Behavior
- Modeled Media effects

Future

- 2D and 3D Simulations
- Multiple Simulations
- Time Branching
- Population Granularity
- Adjacency and Proximity Problem

Summary

- Need:Human Aspect of Homeland Security
- Project: Measured Response (MR)
- Architecture: Tera-scale Computing
- Event: Measured Response 2003
- Updates: New Models
- Future: Visualization, Timelines

Measured Response

Thank You

Prof. Alok Chaturvedi
alok@purdue.edu

Dr. Shailendra Mehta
mehta@purdue.edu

Tejas Bhatt
tejas@purdue.edu

Krannert School of Management, Purdue

Introduction to Virtual Environments - A Collaborated Classroom Offering between PU and IU

Laura Arns
Purdue University

- Eric Wernert, Ph.D.
  - School of Informatics (w/ technical support from School of Informatics and UITS Advanced Visualization Lab)
- Laura Arns, Ph.D.
  - Envision Center (ITaP) and Computer Graphics Technology Dept. (School of Technology)

Course Background

- Introduction to Virtual Reality/Virtual Environments
  - What is VR?
  - How/why to use VR?
  - Hardware and software for VR
- Grad level class
  - PU TECH 519V
  - IU I590

Course Format

- Joint Lectures
  - 2 per week, each 75min
  - Concentrate on theory, background
  - Instructors alternate lectures
- Independent Labs
  - 1 per week, approx 1 hour
- Student presentations
  - Each student gives one 20 min paper presentation during a lecture

Envision Center | Indiana University School of informatics | PURDUE UNIVERSITY

Access Grid

- www.accessgrid.org
- Collection of resources to support group-to-group interactions such as teleconferencing
- Software created at Futures Lab at Argonne National Labs. Runs over Internet2.

Envision Center | Indiana University School of informatics | PURDUE UNIVERSITY

Images from AGFocus, Fall 2003.

Envision Center | Indiana University School of informatics | PURDUE UNIVERSITY

WebCT

- Web-based software to create on-line courses or complement traditional classrooms
- www.webct.com
- Lecture notes and other materials
- Discussion boards, email, etc
- Grades and assignments (PU only)
- Need guest accounts for IU

Envision Center | Indiana University School of informatics | PURDUE UNIVERSITY

Similar Course Offerings

- Previous “Intro to VR” at both IU and PU
- Other distance learning/collaborative VR courses such as Iowa State University “Virtual Environments and Applications”
  - Some offerings included multiple sites such as Virginia Tech and Old Dominion
  - Did not use AG, much more expensive

Advantages

- Less expensive
- AG can run on PC, more students from remote locations
- Larger class size
- Less instructor overhead
- Students hear from multiple instructors
- Opportunities to bring in experts with less expense
- Projection-based allows VR demos in class
- VHS/DVD to make up missed class
- Occasional remote teaching

Advertisement

- Today 2:30-3:45
- PIG in the exhibits/demos
- Join in the class

Student Response

- Independent of technology, course has lots of good info. Wide variety of areas for VR use.
- Course materials very interesting.
- One of my favorite classes.
- Like having 2 instructors with real experience to cover more topics.
- Glad I'm finally getting to take a class that actually uses technology instead of old boring classroom.

Student Response-2

- Audio problems are very annoying
- More difficult to pay attention to non-local lectures. Audio related?

Techie Complications

- Multicast problems
- Choppy audio
- Video codecs
- Slow VNC update
- No back of room screen

Lessons Learned

- Scheduling difficult
  - Time zones will be worse
- Have tech backup (polycom, distribute PPT) & node operator
- Make students introduce themselves first day
- Remind students of microphones
- Having 1 outgoing student helps a lot!
- Extra pauses for questions
- Check "special stuff" in advance (video)
- Have time for short local discussion before or after

Lessons Learned -2

- Department/School support important
- Instructors must work well together & be flexible
  - Realize course content must meet many different guidelines/standards/desires
- Technical content may need adjustment for students of different backgrounds and different access to technology. Must be more general but still practical.
- Choose the right class
  - Not an important prerequisite for other classes
  - Technology related helpful

Envision Center

Indiana University School of informatics

PURDUE UNIVERSITY

Future Courses

- Fall 2004
  - PU
  - IU
  - Iowa State
  - Others?

Envision Center

Indiana University School of informatics

PURDUE UNIVERSITY

References & Contact Info

- Course website: www.avl.iu.edu/I590_VR/
- Access Grid: www.accessgrid.org
- Laura Arns, arns @ purdue . edu
- Eric Wernert, ewernert @ indiana . edu

I - L I G H T