Community Grids Laboratory

January-December 2004 Summary

i) Mission

The Community Grids Laboratory (CGL) was established in July 2001 as one of Indiana University's Pervasive Technology Laboratories. It was funded by the Lilly Endowment and is located in the Indiana University Research Park (Showers) in Bloomington. Its staff includes Director Geoffrey Fox, four Research Associates, two software engineers, one Professional Staff and 18 PhD candidates. There is a visitor from Korea Jai-Hoon Kim funded by his home university, Ajou, working with our PDA research project.

The Laboratory is devoted to the combination of excellent technology and its application to important problems. We believe that e-Business and e-Science will grow in importance and imply global virtual communities. Our technology focus, Grids, is the enabling infrastructure for distributed enterprises and the Cyberinfrastructure for distributed science and engineering. Building communities is an important application of Grids and integrating peer-to-peer network ideas and people into the Grid is a key feature of our work. Much of our innovative research exploits the observation that computers and networks are now so fast that one can use new and more transparent architectures and protocols and move from inflexible hardware to modular flexible software solutions. We also see a blurring of computers and the Network as our systems get more and more distributed.

Cross-disciplinary research and linkage of technologies with applications is an important feature of CGL Research. Here our current major emphasis is in earth science and particle physics but we are developing other opportunities in education, biocomplexity, apparel design, digital film production and sports informatics. We also will be undertaking a new science Grid project to support computational chemistry grid applications.

Outreach to minorities – in particular American Indian and historically black colleges is an ongoing important component of our research.

ii-a) Funded Research

I am not clear what you want here

A highlight this year was substantial funding from the UK plus our new work with the University of Minnesota and Ball Aerospace.

ii-b) Publications and Presentations

In the three years since the founding of the laboratory, members have published over 150 papers and reports. Approximately 50 of these were in reporting period. Further approximately 55 major presentations were also given both as invited talks and as conference contributions. This reflects our emphasis on outreach and on encouraging students to submit conference papers and attend if they are selected. We have a very high acceptance rate due to the interest in and quality of our research.

Fox is writing a new book building on the success of the book published in 2003, "Grid Computing: Making the Global Infrastructure a Reality" co-edited with Fran Berman (head of the NSF Supercomputer center NPACI in San Diego) and Tony Hey (Director of the core UK e-Science program). The new book will be aimed at a broader audience and explain the important applications that Grids can address.

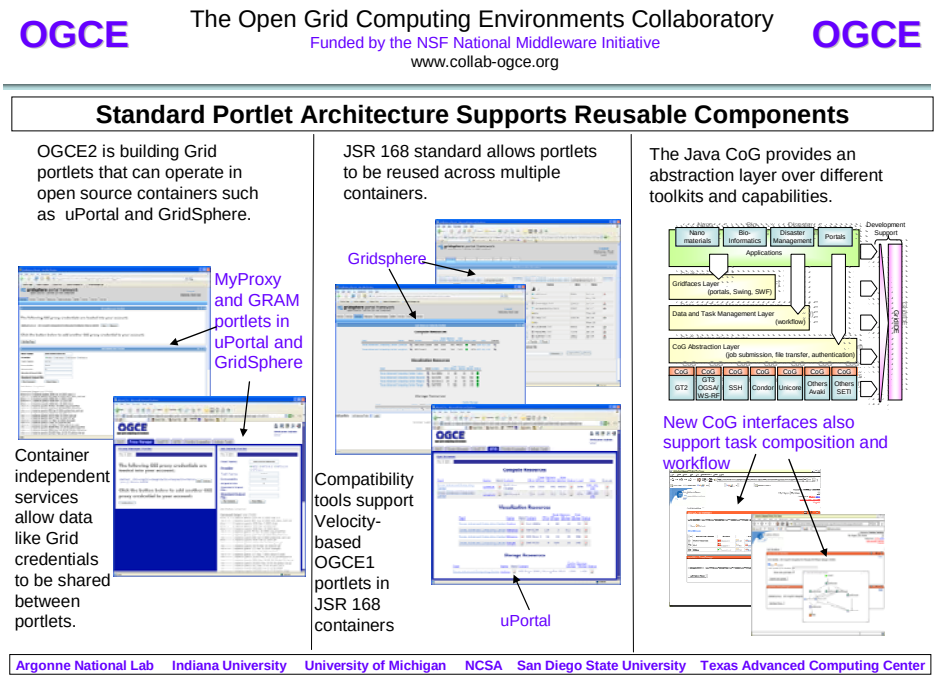
ii-c) New and old Personnel

Korean Visitor; two new staff
Double check number of students
Note Sangmi went to work for Dennis, Youn for SDSC and Sang Boem Lim and Han-Ku Lee got faculty positions in Korea
Bryan went to Grid position (OMII) in UK

iii) Activities in January-December 2004

Many of our research projects were started in 2001 and have now reached a stage where they are not "just research or prototyping" but significant systems that can be used in major applications and are becoming quite well known outside the laboratory. Correspondingly we have consolidated some smaller activities into the larger projects. These major CGL projects are

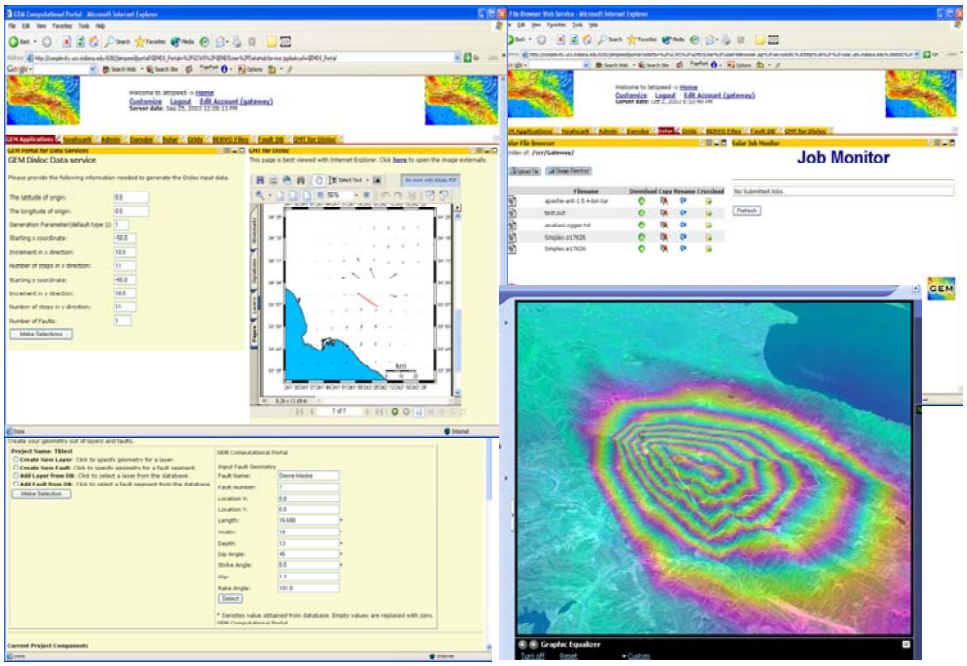
HPJava parallel Java language and mpiJava binding of MPI to Java: This open source project was completed in April 2003 and current work consists of software maintenance in response to user input. . There have been over two thousand downloads so far of the mpiJava system which allows development of portable Java programs that can be executed efficiently on important distributed-memory parallel computers. The HPJava Development Kit releaseincludes a translator (i.e. compiler) for a syntactic extension of Java with support for Fortran-like multidimensional arrays and HPF-like distributed arrays. The lack of these has been an obstacle to uptake of Java for scientific computing. The release package also includes standard libraries for operating on the new data structures.



Grid Portal Technology: This OGCE (Open Grid Computing Environment) activity builds on earlier Web and Grid portal projects and is centered on developing the architecture and support tools for the user interfaces to Grids. This collaboration of six major organizations is led by CGL and is the major such Grid portal activity in the nation. The OGCE is funded by the NSF National Middleware Initiative. Current work is focusing on providing support for computational Grid portals using the new Web Service standard WSRP and the Java standard JSR168 for portlets. The adoption of these technologies will allow OGCE technology to be widely incorporated in a variety of portal systems. Major OGCE deployment activities currently underway include the development of NSF TeraGrid user and system portals. The Community Grids Lab was also part of the team, lead by the University of Minnesota, that was awarded a 2004 NSF ITR project to develop a collaborative computational chemistry Grid, VLab. CGL will develop portals and Web Service infrastructure for this project.

SERVO (Solid Earth Research Virtual Observatory) Grid: This NASA funded activity provides the Cyberinfrastructure supporting earthquake science with both data and simulation components running on large parallel machines. It is a collaboration involving 5 universities and is led by the Jet Propulsion Laboratory with CGL responsible for the system architecture and Grid portal and services. It was deployed for use in research and education in spring 2004 and the first tutorial on it was offered in July 2004 at an international meeting in Beijing. This meeting involved a group of earthquake scientists

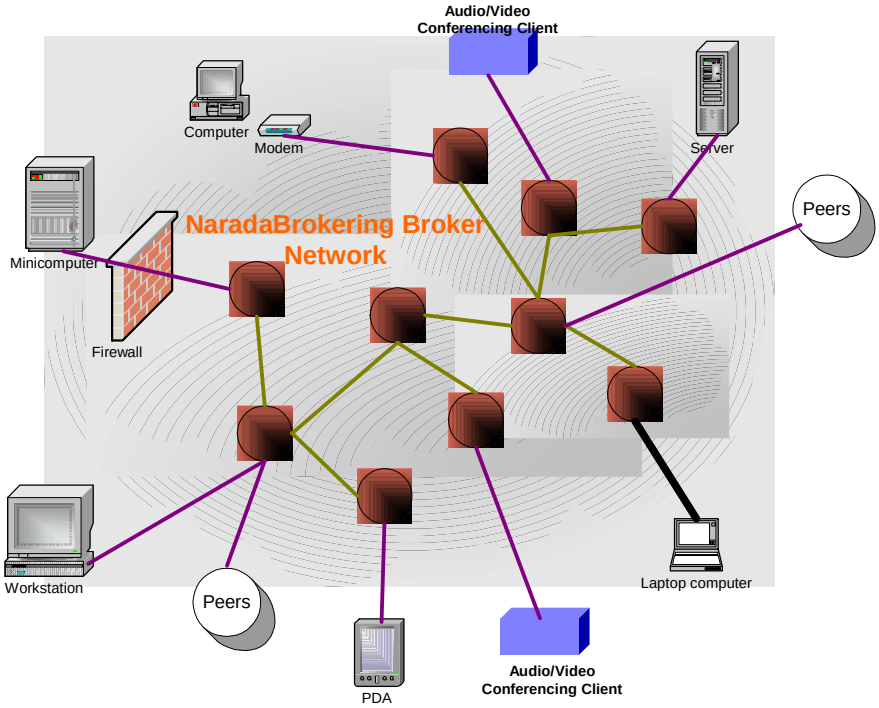
coordinated by the Asia Pacific Economic Cooperation. A major next step iSERVO involves expansion of our Grid to sites in Australia, China and Japan. SERVO's initial funding, to develop and deploy the core Grid and portal technology and support an initial set of earthquake simulation and modeling applications was funded by the NASA Computational Technologies (CT) program. SERVO's current funding comes through NASA's Advanced Information Systems Technology (AIST) program. Current activities build on the core application management technologies to build more sophisticated integration tools for numerous codes and data sources.



Collage of SERVOGrid Portal Views

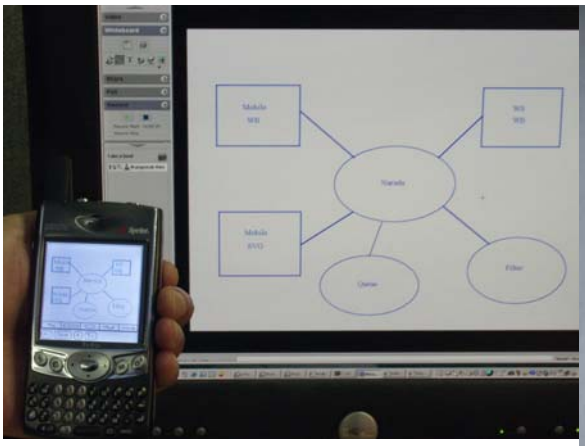
Fusion Grid: We continue our collaboration with San Diego State University (project lead), Texas Advanced Computing Center, San Diego Supercomputing Center, and General Atomics on the DOE SciDAC-funded Portal Web Services project (http://www.doeportals.org/). This project is using OGCE-developed portal technology to provide a computational web portal for the the DOE's Fusion Grid project. CGL activities include the development of Apple OS X-compatible versions of GlobalMMCS collaboration tools and OGCE-compatible GlobalMMCS portlets. We are also using the NaradaBrokering system to develop event-based, collaborative versions of Interactive Data Language (IDL)-based analysis and visualization tools used by the Fusion Grid community (ReviewPlus, http://web.gat.com/comp/analysis/uwpc/reviewplus/manual/). IDL is a popular tool for many scientists, including several of our SERVO collaborators, so we are striving to make our collaborative IDL extensions as general as possible so that this work may be extended to new applications.

NaradaBrokering Grid Messaging System: This core Grid middleware is being actively developed and provides a software overlay network supporting high performance reliable communication between multiple Grid and Web services and their clients. We have extensively benchmarked the system

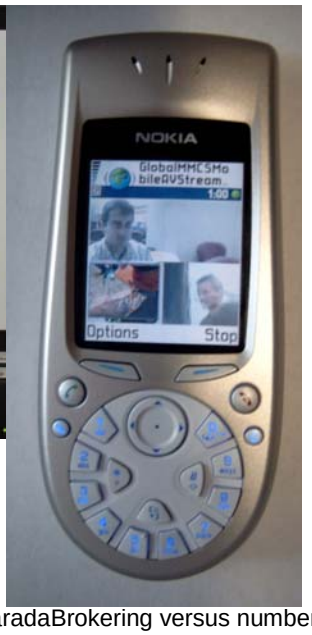


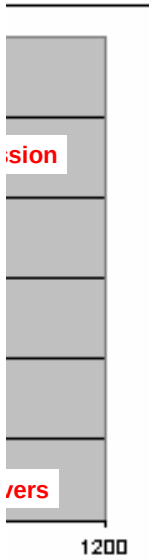
showing that a single broker can support several simultaneous clients. We have linked NaradaBrokering to the peer-to-peer model supporting both JXTA and an enhanced version of the well known file transfer system BitTorrent. We have also prototyped a Web Services reliable messaging version of GridFTP. Other enhancements in this reporting period include the use of NaradaBrokering to support Web Services and substantial additional funding has been obtained from the United Kingdom to deploy our software as core Grid infrastructure supporting reliable messaging and notification services. The latter includes support of the new specifications WS-Eventing and WS-Notification linked to the original Java Message Service. We have added two professional software engineers to our team to lead the transition from NaradaBrokering as a promising prototype to it being robust infrastructure to support multiple Grid and peer-to-peer projects.

GlobalMMCS Service-based Collaboration Environment: This was made available for initial user evaluation in May 2004 with the initial release supporting audio-video conferencing with interoperation between H323, Access Grid and RealPlayer subsystems. It is also features an innovative shared screen capture suitable for dynamic displays such movies. The user interface supports the portlet architecture. Recent work has improved the performance of the system and supported both handheld and Macintosh clients for upload and download. All communication in GlobalMMCS uses NaradaBrokering allowing it reliable transport without the need for hardware multi-cast. We have made substantial progress



GlobalMMCS links Cell Phones and PC's with whilte board and streaming video illustrated





Latency in milliseconds introduced by NaradaBrokering versus number of clients. Latencies of < 15 ms (200-600 clients) are acceptable.

with two new technologies. Firstly a powerful archiving system built in to NaradaBrokering that will enable VCR like capabilities on real-time streams; secondly a video annotation tool aimed at the digital entertainment and e-Sports applications.

These major projects are complemented by two smaller research activities

Carousel Grid Interface to Mobile clients: This project developed a proxy architecture and uses NaradaBrokering to link cell-phones to Grid systems. Work in this period has focused on improving performance and providing fault tolerant communication between Grid services and cell phone clients. We also have demonstrated multi-point video conferencuing linking cell phones with traditional Polycom and Access Grid clients.

Message-based User Interfaces: This has shown how to adapt the well known MVC (Model View Control) to use explicit messaging to link the model and view components of an application. This gives an elegant linkage of traditional desktop applications with Web services. It should allow easier support of applications on both Windows and Linux clients. This novel M-MVC (Message based MVC) architecture is being used both in our mobile client research and to develop a valuable collaborative visualization tool for the Department of Energy Nuclear Fusion Grid as discussed above..

Dr. Fox is co-chair of the Semantic Grid and Grid Computing Environments Research Groups of the Global Grid Forum; he has just been selected to be a member of the Grid Forum Steering Group. He works closely with the UK e-Science program where he is a member of their TAG – Technical Advisory Group – and helped set up the OMII – Open Middleware Infrastructure Institute which started in January 2004 where Fox serves on their technical advisory board. Fox has continued as lead editor of the well regarded journal Concurrency and Computation: Practice and Experience and is on the editorial board of 3 other international journals in computer and computational science.

iv) Activities in January-June 2005

The discussion above indicated that the Lilly funding had enabled us to develop a number of significant projects of international significance. In the next phase we will continue to enhance the technology but also pro-actively look for new collaborations and projects that can utilize our software and suggest new directions. We are exploring international activities in core Grid technology – especially with researchers in China and the UK which both have very active national programs. Two researchers from China will join the group in 2005 with funding from their home university.

The success of the SERVO Grid project has led to a follow-on iSERVO with international scope including Australia, China, Japan and USA. This new international Solid Earth Research Virtual Observatory will contribute seismic sensors, computers, simulation codes, and data repositories across the four countries with a focus on advancing our ability to predict earthquakes. This will be organized under the aegis of APEC – the Asia Pacific Economic Cooperation. SERVO development in the upcoming period will support more sophisticated integration of data assimilation and data mining applications. This will incorporate GIS-based data grid services for managing access to GPS and seismic activity catalogs that will be bound together with previously developed application execution services into an application coupling framework (www.hpsearch.org). The tools used are generic and can be used in other application grid projects that we are pursuing, such as crisis management and biocomplexity grids.

The OGCE will be building on its current work in two primary fronts: outreach and core development. We will be working with portal developers at NCSA and Argonne National Labs to develop application Web portals using our current OGCE releases. This will include the development of extensive, formal tutorial material along with more informal collaborations. This tutorial material will also be used for presentation material at the HPC 2005 conference in April. Core development work during this period will be the investigation of lightweight, thin, purely Web service-based portal containers that will simplify continuous, evolutionary upgrading. We will also work on messaging and service systems for managing portal content within current portal container frameworks. Our portal collaborations with the Fusion and Earthquake Grids will continue.

The new VLab project with the University of Minnesota, Florida State University, and Louisiana State University will combine CGL expertise in portals, messaging systems, and Web Services to build collaborative computational and visualization Grid systems.

We will be teaching an Internet class on high performance computing during the spring to students at Jackson State University; a Historically Black University. This is part of our outreach program and tests our core GlobalMMCS collaboration environment. More generally we are looking for funding to explore the linkage of our collaboration and Grid projects to support e-learning or Education Grids. Here we are exploring collaborations with the Schools of Education and Continuing Studies at Indiana University.

We are continuing to work with the Undergroundfilm.org non profit organization to develop educational material in the multimedia area. This will prototype a collaborative multimedia tool allowing annotation of real-time video streams.. This tool will be key to a collaboration with HPER – the School of Health Physical Education and Recreation at Indiana University – in the area of e-Sports. The HPER collaborators are funded through the PTL fellows program. We will work with HPER on the support of trainers and coaches interacting with players across the Internet. Trainers will monitor faculty and staff exercising in their home while sports coaches will interact with players in Asia.

We have several application areas where we hope to grow our collaborations and provide testbeds for the laboratory technology. These include particle physics with a new approved experiment looking for "Glueballs", crisis management (with the POLIS center at IUPUI and the Open Geospatial Consortium) and biocomplexity. We continue to work with the start-up Anabas in the e-Sports, apparel and e-learning areas. In particular we have partnered on several SBIR and STTR proposals to commercialize CGL technologies.

1.5 Contact Information

Geoffrey Fox: Phone 8122194643 email: gcf@indiana.edu
Community Grids Laboratory
501. N. Morton Street, Suite 224
Bloomington IN USA 47404-3730

http://www.infomall.org