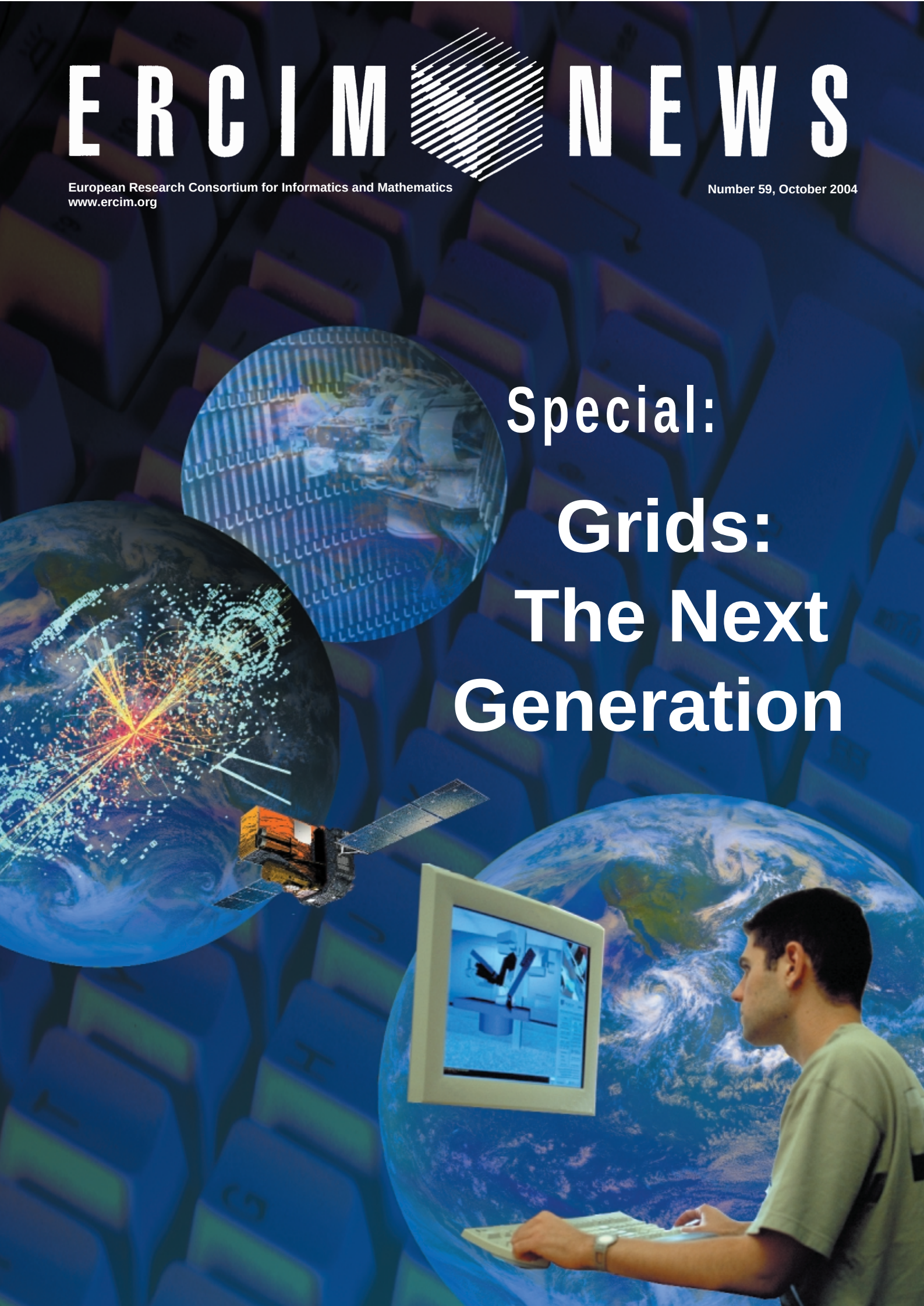
ERCIM NEWS

European Research Consortium for Informatics and Mathematics
www.ercim.org

Number 59, October 2004



Special:

Grids: The Next Generation

CONTENTS

SPECIAL THEME: GRIDS — THE NEXT GENERATION

Grids: A Crucial Technology for Science and Industry

Fabio Colasanti, Director-General European Commission Directorate-General for "Information Society"

Grids are an emerging technology that promises to change the lives of individuals, organisations and society as profoundly as e-mail and the Web have done in the past decade. It is perhaps not surprising that the first Grids were developed for the European Organisation for Nuclear Research (CERN), the particle physics laboratory that gave us the World Wide Web.

By providing everyone with the high-performance computing and knowledge capabilities previously available only to the largest corporations and laboratories, Grids have an immense potential to underpin sustainable growth by improving the competitiveness of existing industries and by helping usher in new markets and services previously thought impossible. It is expected that their impact on our quality of life will be profound, allowing us to monitor and model more successfully everything from global climate change to the way cars behave in collisions.

Grids are both a crucial enabling technology for reaching the 'Lisbon strategy', set in spring 2000, of transforming the European Union into the most competitive, knowledge-based economy in the world by the year 2010, as well a fundamental building block in the realisation of the European Research Area (ERA).

In light of this significant potential, the European Commission has been financing Grid research since early 2000, when the first EU-funded projects were launched under the Fifth Framework Programme for Research and Technological Development (FP5). The Grid research projects under FP5 were focused on both technology development and application pilots. In FP6 (2002–2006) a new technology-driven approach to Grid research is being pursued by the European Commission, substantiated by the longer-term vision of "Next Generation Grids" (http://www.cordis.lu/ist/grids). Research challenges to be addressed for the realisation of this vision include the conceptualisation and/or development of 'network-centric Grid operating systems' as new or enhanced fabrics for future distributed systems and services, scale-free, adaptive and dependable Grid architectures and middleware required to realise the 'invisible Grid' vision. In addition, research will address new or enhanced Grids service middleware empowering individuals and organisations to create, provide and use a variety of services, anytime, anywhere, in a transparent and cost-effective way.

A complementary building block to Grid Technology research are EU-funded activities in the area of 'Research Infrastructures' which include the deployment and testing of Grids over GEANT, the world's most powerful research network which links over 3 000 research institutions across Europe.

The budget for Grid Technology research in FP6 is more than double that in FP5 – a clear demonstration of the field's strategic importance. The first 12 FP6 IST Grid research projects launched in summer 2004 with total EU funding of 52M will create a 'critical mass' of expertise and resources from across Europe. Ambitious in their scope, they aim to deliver, inter alia, new Grid architectures by the end of the decade, to develop generic Grid technologies and to advance significantly our capabilties in vital areas such as security, business models, data integration, programming models, collaboration and knowledge discovery.

One of the largest projects in the portfolio just launched – the Network of Excellence CoreGRID – addresses longer term Grid research and will play an essential role in creating the foundations for next generation Grids. The European Research Consortium for Informatics and Mathematics, ERCIM, leads this significant endeavour. The action will bring together existing, but fragmented, Grid research communities. By doing so it will strive for durable excellence, integration and co-operation in Grid research across Europe, thus contributing to the creation of a European Research Area for Grids.

Considering the significant and sustained European investment in Grid research, Europe's strengths in this area are well established. The challenge is now to create more effective routes for industrial exploitation in order to translate successfully research results into economic benefits. For Europe to capitalise better on its strengths, it is indispensable that collaboration between research organisations, funding bodies and industry at all levels of the value chain and across national borders is reinforced.

More integrated, long-term research visions and effective plans for their implementation need to be established, taking into account industrial needs and commercial ambitions. This will be essential for reaping benefits from the promising area of Grid research, paving the way towards the provision of Grid, software and knowledge services as a utility.

Fabio Colasanti

CoreGRID Kick-off Meeting

by Thierry Priol

The CoreGRID Network of Excellence, in the area of Grid and Peer-to-Peer technology had its first meeting at CETIC, Charleroi (Belgium), on 13–14 September, 2004.

The kick-off meeting of the CoreGRID Network of Excellence (see also article on page 56), which looks at large-scale, distributed Grid and Peer-to-Peer Technology, was recently held at CETIC in Charleroi (Belgium). Almost all the 42 CoreGRID partners were represented by 73 participants. Thanks to the CETIC team, this first meeting ran smoothly.

© CETIC; Photo: Alexis Tamniaux

From left to right: Pierre Guisset (CETIC), Bruno Le Dantec (ERCIM), Franco Accordino (EC) and Thierry Priol (INRIA/ERCIM).

The first day commenced with a welcome from the Director of CETIC, Pierre Guisset, who along with his team had the responsibility of organizing this meeting. Franco Accordino, project officer at the European Commission in charge of CoreGRID, gave the welcome address, describing the EC vision of how to boost excellence in European Grid research through sustainable integration and cooperation, thanks to the Network of Excellence instrument within FP6. He highlighted the importance in the development of Grid technology of CoreGRID within the EC strategy. Following this, representatives of each partner introduced themselves to the general audience and gave a brief overview of their partner's involvement with the joint program of activity. The morning session ended with a general presentation of CoreGRID by the scientific coordinator, Thierry Priol, and a set of presentations of the six scientific areas covered by the network, given by their respective scientific leaders: Domenico Talia (Knowledge and Data Management), Marco Danelutto (Programming Model), Artur Andrzejak (System Architecture), Ludek Matyska (Grid Information and Monitoring), Ramin Yahyapour (Scheduling) and Vladimir Getov (PSE, Tools and GRID systems).

The afternoon session was devoted to a presentation on management issues by Bruno Le Dantec, which covered in detail various aspects of the administrative procedures related to the network: budget, reporting, evaluation, fellowship and mobility programs. This session concluded with a set of meetings of the bodies that govern the network: the General Assembly meeting, chaired by Domenico Laforenza, the Executive Committee meeting, chaired by Thierry Priol, and the Integration Monitoring Committee meeting. For the latter meeting, a call was issued for candidates to chair this committee, and a vote was taken soon after the conclusion of the meeting. Ron Perrott, from Belfast University, was elected as the chairman of the IMC. The first day concluded with a dinner, allowing participants to pursue their discussions over Belgian cuisine.

The second day consisted of a set of parallel sessions, each dedicated to one of the six scientific areas. The purpose of these sessions was to establish a more precise set of expected scientific results, a work-plan and task descriptions, and to review the deliverables and milestones. Each partner involved in these scientific areas gave a short description of its role, and the scientific leader of each area proposed an agenda for the next six months. Dissemination actions were also discussed. A summary of the parallel sessions was reported in a plenary session that afternoon to inform all participants of the future activities related to the six scientific areas. The afternoon concluded with presentations on the activities related to integration (given by Sergei Gorlatch), spreading excellence (Anne Falier) and collaboration with other EU-funded grid projects (Wolfgang Ziegler).

Link:
http://www.coregrid.net

Please contact:
Bruno Le Dantec, ERCIM
Administrative & Financial Coordinator
Tel: +33 4 92 38 50 13
E-mail: Bruno.Le_Dantec@ercim.org

Thierry Priol, INRIA, France
Scientific Coordinator
Tel: +33 2 99 84 72 10
E-mail: Thierry.Priol@inria.fr

8th ERCIM Workshop "User Interfaces for All"

by Constantine Stephanidis

The 8th ERCIM Workshop 'User Interfaces for All' was held in Vienna, Austria, 28-29 June 2004, and Prof. Chris Stary, University of Linz, was the Programme Chair and local organiser. The workshop built on the results of the seven previous workshops.

In the tradition of its predecessors (held in Heraklion, Crete, Greece, 30-31 October 1995; Prague, Czech Republic, 7-8 November 1996; Obernai, France, 3-4 November 1997; Stockholm, Sweden, 19-21 October 1998; Dagstuhl, Germany, 28 November-1 December 1999; Florence, Italy, 25-26 October 2000; and Paris (Chantilly), France, 24-25 October 2002), the 8th ERCIM Workshop 'User Interfaces for All' aimed to consolidate recent work and to stimulate further discussion on the state of the art in User Interfaces for All and its increasing range of applications in the upcoming Information Society. The special theme of this year's workshop was 'User-Centred Interaction Paradigms for Universal Access in the Information Society'. The requirement for User-Centred Universal Access stems from the growing impact of the fusion of the emerging technologies and from the different dimensions of diversity that are intrinsic to the Information Society. These dimensions become evident when considering the broad range of user characteristics, the changing nature of human activities, the variety of contexts of use, the increasing availability and diversification of information, knowledge sources and e-services, the proliferation of technological platforms, etc. The 8th ERCIM Workshop 'User Interfaces for All' focused on the new HCI challenges arising from this evolution, and on how these affect the continuing efforts towards Universal Access in the Information Society.

Keynote speakers were Prof. Ben Shneiderman (University of Maryland, Maryland, USA), who presented 'Interface Design Strategies to Promote Learnability for All', and Prof. Jenny Preece (University of Maryland Baltimore County, Maryland, USA), who talked about 'Online Communities for All'.

The workshop attracted the strongest ever interest with over 140 submissions from all over the world, covering a wide range of topics that include novel interaction paradigms and contexts of use, innovative concepts of universal accessibility and sociability, new modalities and dialogue styles, user-centered design in mobile application scenarios, late-breaking empirical results with respect to assessing universally accessible applications, and standardization efforts. Contributions addressed not only technological solutions, but also design paradigms and empirical methods for evaluation, as well as policy developments.

Photo taken during the workshop. Standing: Ben Shneiderman (first keynote speaker), Simone Stoiber (workshop assistant organiser), Constantine Stephanidis (Working Group chair) and Christian Stary (workshop organiser). Sitting: Noelle Carbonell (Workshop PC member) and Jennifer Preece (second keynote speaker).

The official Workshop Proceedings (http://www.ui4all.gr/workshop2004/publications/lncs3196-toc.html) will be published by Springer as part of LNCS (Lecture Notes in Computer Science) series, in the subline 'State-of-the-Art Surveys', and will be embedded in the LNCS digital library. The workshop adjunct proceedings are electronically available in PDF format at http://www.ui4all.gr/workshop2004/publications/adjunct-proceedings.html.

From the year 2001 onwards, the ERCIM Workshop 'User Interfaces for All' takes place every two years, in alternation with the newly established 'Universal Access in Human-Computer Interaction' (UAHCI) Conference, also held every two years in the context HCI International and affiliated conferences. In 2005, the 3rd UAHCI Conference will be held in Las Vegas, Nevada, USA, 22-27 July (http://www.hci-international.org/).

Links:
8th ERCIM Workshop 'User Interfaces for All': http://www.ui4all.gr/workshop2004/

ERCIM Working Group 'User Interfaces for All': http://www.ui4all.gr

Please contact:
Constantine Stephanidis,
ICS-FORTH, Greece
Tel: +30 2810 391741
E-mail: cs@ics.forth.gr

DELOS Information Day

by Costantino Thanos

The DELOS Network of Excellence on Digital Libraries organized an Info-Day in conjunction with the European Conference on Digital Libraries (ECDL'04) in Bath, UK, on 15 September. The purpose was to present DELOS to the wider European digital library research community in terms of its objectives, organization, activities and results so far achieved.

DELOS is a Network of Excellence on Digital Libraries funded by the Sixth Framework Programme (FP6) of the European Union. The Network is planned as a four-year activity and began in January 2004. It aims to build on the results of a previous DELOS initiative under FP5. The Network is currently constituted by members from 46 research institutions in 13 European countries and is coordinated scientifically by ISTI-CNR and managed by ERCIM.

The Info-Day was opened by Costantino Thanos (ISTI-CNR), the scientific coordinator of the Network, who welcomed the participants (80 attendees) and presented the main directions of DELOS, its internal organization into clusters (work-packages) and the cluster specific topics of interest. The main objective of DELOS is to define and conduct a joint programme of activities in order to integrate and coordinate the on-going research activities of the major European research teams in the field of digital libraries for the purpose of developing the next generation digital library technologies. The Network is organized into eight Clusters (Work-packages):

- Digital Library Architectures
- Information Access and Personalization
- Audio/Visual and Non-Traditional Objects
- User Interfaces and Visualization
- Knowledge Extraction and Semantic Interoperability
- Preservation
- Evaluation
- Dissemination.

Claude Poliart, the DELOS project officer, also welcomed the participants and outlined the Commission's expectations from the Network. DELOS should tackle the fragmentation of European research in the digital library area by integrating a critical mass of expertise around a joint programme of activity thus spreading excellence and establishing European leadership at the global level. The Commission wants to see a joint management of the knowledge portfolio between the Network's members and a sharing of research platforms, tools and facilities. DELOS should become a world reference for the digital library community.

Costantino Thanos, DELOS coordinator, opens the Information Day.

Following these two introductory talks, Yannis Ioannidis (University of Athens) presented the results of a recently organized DELOS brainstorming workshop on 'Digital Libraries: Future Research Directions'. In particular, he presented the DELOS vision for DL as user-centred systems which function not just as passive repositories but as active communication and collaboration tools. He stated that the actual challenge for the DL research community was to develop powerful, generic digital library management systems rather than ad-hoc solutions meeting the particular needs of a specific application community.

The various workpackage leaders then presented their clusters in terms of activities and results so far achieved. So far, in addition to the brainstorming meeting in Covara, Italy, attended by 24 internationally recognized experts in digital-libraries-related research areas, the Network has organized nine scientific events on a wide range of topics, several national and regional events, and an international summer school on digital library technologies.

The two important evaluation initiatives supported by the Network, the Cross Language Evaluation Forum (CLEF) (see workshop report on page 72) and the Initiative for the Evaluation of XML Retrieval (INEX), which aim at promoting research into multilingual information retrieval systems and into content-oriented XML retrieval, respectively, were also presented.

The meeting concluded with a presentation by Vittore Casarosa (ISTI-CNR) illustrating the dissemination and technology transfer activities of the Network and the research exchange programme.

Link:
DELOS Website: http://www.delos.info/

Please contact:
Costantino Thanos, ISTI-CNR, Italy
Tel: +39 050 3152910
E-mail: costantino.thanos@isti.cnr.it

The EU Information Society Technologies Programme Advisory Group – ISTAG

by Michael Wilson

In planning the Information Society Technologies (IST) programme, the European Commission Directorate-General for "Information Society" takes advice and guidance from many sources. These include consultations with past and present project managers, and active researchers through questionnaires and consultation meetings before major changes to the programme. One of the standing bodies which is available to the European Commission for consultation on the IST programme is the IST Advisory Group (ISTAG).

ISTAG has been set up to advise the Commission on the overall strategy to be followed in carrying out the IST thematic priority and related activities of research as well as on the orientations with respect to the European Research Area by helping to stimulate the corresponding European research communities.

In the context of building an IST European Research Area, a key role of ISTAG is to reflect and advise on the definition and implementation of a coherent research policy in Europe. This policy should ensure the mastering of technology and its applications and should help strengthen industrial competitiveness and address the main societal challenges. It is through such an exercise that recommendations on priorities and activities of Community-funded research in IST would be drawn.

ISTAG members are appointed as individuals rather than to explicitly represent the interests of a single group, since formal representation would result in too large a body to reach significant decisions. ISTAG has a Chairman and 30 members who provide a reasonable coverage across the EU member and accession states (including Rumania and Bulgaria), companies (eg Nokia), research institutes (eg DFKI GmbH), standards bodies (eg W3C), universities (eg Open University of Catalonia), national funding agencies (eg EPSRC in the UK) and even politicians (eg a member of the Parliament of Estonia). Among these individuals two are members of ERCIM institutes who can present the interests of ERCIM: the chairman Prof. Encarnacao of the Fraunhofer ICT Group in Germany and Prof Michael Wilson, CCLRC Rutherford Appleton Laboratory, UK. ERCIM also provides all ISTAG members with its publications.

The objectives of the IST thematic priority are not purely to produce the highest quality research in Europe which can be exploited by European companies, but also to meet the broader objectives of the European Union established by the treaties and Council of Ministers. Any advice which ISTAG offers the CEC is only taken as a recommendation on priorities and can be overruled by other factors. For example, ISTAG cautioned that the implementation of the new funding instruments in the 6th Framework Programme would not achieve the objective of freeing research management without more independent financial control, but the powers of the commission under the treaties prevented them from devolving more financial control to consortia, and the resulting instruments were implemented as we now have them. Similarly, ISTAG has not strongly supported proposals for European Technology Platforms (ETP) as a suitable instrument to structure research and innovation in the ICT area in FP7, although they are supported by industries with a different structure, such as aerospace.

ISTAG meets in plenary session quarterly to establish working groups and report on their progress. The topics for working groups are established by the commission in consultation with the chairman, although as ISTAG gains confidence through working together they are beginning to propose topics themselves (eg to address legal issues potentially arising from ICT research).

The three topics addressed by working groups in 2003 reflected three stages in the innovation cycle: IST Research Content which mainly developed the Ambient Intelligence vision to guide the structure of future work programmes; the implementation issues of Human resources, research infrastructures, funding mechanisms and partnership; the exploitation of research results.

In 2004 the topics considered take a different approach, considering the topics for research in terms of a complementary vision of the Grid, but also considering Grand Challenges to structure research objectives rather than using visions and roadmaps alone; to consider a particular instrument for promoting innovation through Experience and Application Research Centres (EARC); and to consider actions that can be taken as EU-wide initiatives to promote and advance European research and technology beyond the existing IST programme instruments and work programme structure.

For further information, past working group reports and drafts of those under development are available on the ISTAG web site.

Link:
ISTAG web page:
http://www.cordis.lu/ist/istag.htm

Please contact:
Michael Wilson, CCLRC
Tel: +44 1235 446619
E-mail: M.D.Wilson@rl.ac.uk

Tim Berners-Lee Knighted by Her Majesty Queen Elizabeth II

Tim Berners-Lee, the inventor of the World Wide Web and Director of the World Wide Web Consortium (W3C) was dubbed a Knight Commander, Order of the British Empire (KBE) by Her Majesty, Queen Elizabeth II during an Investiture in London on Friday, 16 July 2004. The rank of

Photo: Le Fèvre Communications

Knight Commander is the second most senior rank of the Order of the British Empire, one of the Orders of Chivalry.

"I am humbled by this great honor," stated Sir Timothy. "The Web came about through an ongoing collaboration with my fellow inventors and developers worldwide. Everyone in the Internet community should be recognized by this honor."

He continued, "The Web must remain a universal medium, open to all and not biasing the information it conveys. As the technology becomes ever more powerful and available, using more kinds of devices, I hope we learn how to use it as a medium for working together, and resolving misunderstandings on every scale."

While working in 1980 as a consultant software engineer at CERN, the European Particle Physics Laboratory in Geneva, Switzerland, Sir Timothy wrote his own private program for storing information using the kind of random associations the brain makes. The 'Enquire' program, which was never published, formed the conceptual basis for his future development of the Web. Subsequently he proposed a global hypertext project at CERN in 1989, and by December 1990, the program 'WorldWideWeb' became the first successful demonstration of Web clients and servers working over the Internet. All of his code was made available free on the Internet at large in the summer of 1991.

Sir Timothy wrote the book "Weaving The Web" (HarperCollins, 1999) which describes the Web's birth and evolution.

Links:
http://www.w3.org/People/Berners-Lee/
http://www.w3.org/2004/07/timbl_knighted

Celebrating W3C's 10th Anniversary

In 1994, Tim Berners-Lee created the World Wide Web Consortium (W3C) to "Lead the Web to its Full Potential." This year, W3C celebrates its 10th anniversary, and the Consortium is organizing a one-day symposium (1 December) for Members and invited guests to reflect on the progress of the Web, W3C's central role in its growth, and risks and opportunities facing the Web during W3C's second decade.

The symposium program will consist of sessions about the Web to the present and into the future.

Discussion will be led by Web pioneers, prominent technologists, industry leaders and press personalities. The day will end with reflections and projections by W3C Director and the Web's inventor, Tim Berners-Lee.

Link:
http://www.w3.org/2004/09/W3C10.html

COPRAS European Project Kick-off Meeting

COPRAS (Co-Operation Platform for Research And Standards) is a three-year Support Action under the European Commission's Information Society Technologies (IST) Programme, which objective is to improve the interface between research and standards. The members of the COPRAS consortium are the officially recognized European Standards organizations — CEN, CENELEC, ETSI, together with The Open Group and the World Wide Web Consortium (W3C).

The COPRAS project held its first kick-off meeting on 14 October at the CEN/CENELEC Meeting Centre in Brussels, together with a number of other European IST projects from call 1 of the 6th Framework Programme (FP6 IST), some of whose work COPRAS will try to make into international standards. People from various working groups inside the standards organisations that make up COPRAS were present to help those European projects develop a plan (a 'Standards Action Plan') for the standardization of their work and to couple each of them with a particular standards body and possibly couple to them to each other as well.

Visit the COPRAS Web site and subscribe to the COPRAS public mailing list: copras-public-request@w3.org (Subject: subscribe).

Link:
http://www.copras.org/

W3C Spanish Office to promote W3C Standards in Tour of Spain

The W3C Spanish Office organizes its first "W3C Standards Tour" around different Universities in Spain from 3 to 26 November 2004. The tour aims to make W3C work known, to establish contacts with researchers at Universities in Spain and to promote the use of W3C Technologies between students from technical Universities.

'W3C Standards Tour' consists of a route with an environmental friendly bus, which will make stops of one or two days at the following Spanish cities: Gijón (Opening: 3 November), Bilbao (4 - 5 November), Zaragoza (8 November), Barcelona (9 - 10 November), Valencia (11 - 12 November), Sevilla (15 - 16 November), Madrid (18 - 19 November), Salamanca (22 November), A Coruña (24 November) and Oviedo (Closing: 26 November).

The bus has access facilities for people with disabilities, 14 multimedia equipments, where demos of W3C Technologies will be available, projection equipment, videoconference and Internet connectivity via satellite. At the same time, conferences about W3C Technologies and meetings with researchers are organized at the host Universities.

Additionally, the "First W3C Prize on Web Standardization" will be held during the tour. This national prize is an initiative of the W3C Spanish Office to encourage W3C Standards use and adoption within Spanish Universities. The prize will award the best prototype, selected among ten finalist prototypes that, in an innovative way, integrates any of the W3C Technologies in the following fields: Semantic Web, Device Independence, Voice and Multimodal Interaction. The winner will receive an applied research grant to develop a full project based on his/her awarded prototype.

Links:
http://www.w3c.es/gira/info/intro.html.en
http://www.w3c.es/

Latest W3C Recommendation

- **8 September 2004: Speech Synthesis Markup Language (SSML) Version 1.0;** Mark R. Walker, Daniel C. Burnett, Andrew Hunt

An exhaustive list of all W3C Technical Reports:
http://www.w3.org/TR/

Multimodal Web Workshop and Seminars

On 19-20 July, W3C organized a workshop on Multimodal Interaction, under the auspices of the IST Programme's MWeb project.

As the W3C Multimodal Interaction Activity is entering its third year, much progress has been made defining the W3C MMI framework, and the workshop was organised as an opportunity for discussing W3C's current plans and for providing feedback and suggestions for future work. Another goal of the workshop was looking for feedback from multimodal user communities that are currently less well represented in the W3C Multimodal work (eg automotive industry/telematics, home entertainment industry, healthcare, aerospace and manufacturing) as well as feedback and input from the Multimodal research community.

41 people attended the workshop and 19 papers (publicly available) were presented, both from academia and the industry, contributing for a better understanding of each other's experience and requirements. The workshop concluded with separate "break-out" sessions, where participants gathered to discuss problems of more specific interest to them, such as multimodal interaction on mobile devices or advanced academic research. Many attendants also expressed a need for a multimodal interaction authoring language.

The organizers are confident that the results of these discussions as well as the quality of the papers and presentations made this workshop very successful. All papers and slides are available online, but the organizers are still working on editing the minutes of the discussions, for publication in the next few weeks.

Multimodal Web Seminars in Madrid and Berlin

These two Multimodal Web seminars' objective is to inform European research and industry about the results of W3C's Multimodal current work. They are organized in Berlin 13 October 2004 and Madrid 18 November 2004. They will use a combination of speakers from W3C Members and W3C team. Attendance is open to the public.

Links:
MMI Workshop page:
http://www.w3.org/2004/02/mmi-workshop-cfp.html

MMI Workshop papers:
http://www.w3.org/2004/02/mmi-workshop/papers

Multimodal Web Seminar in Berlin, 13 October 2004:
http://www.w3c.de/Events/2004/membermeeting.html

Multimodal Interaction Seminar in Madrid, 18 November 2004:
http://www.w3c.es/Eventos/2004/Noviembre/MMI/

Four W3C Workshops this Fall

W3C Workshop on Semantic Web for Life Sciences, 27-28 October 2004, Cambridge, USA

The W3C is organizing a workshop to discuss emerging and future applications of Semantic Web for Life Sciences (SW-LS). Specifically, how can Semantic Web technologies such as the Resource Description Framework (RDF), Web Ontology Language (OWL) and domain-specific standards such as the Life Sciences Identifier (LSID) help to manage the inherent complexity of modern life sciences research, enable disease understanding, and accelerate the development of new therapies for disease?

Link:
http://www.w3.org/2004/07/swls-cfp.html

W3C Workshop on Constraints and Capabilities for Web Services, 12-13 October 2004, Redwood Shores, CA, USA

Various aspects of a Web service may require description. This includes its constraints (eg, "You must use HTTP Authentication when accessing this service") and its capabilities (eg, "I can support GZIP compression"). Likewise, clients accessing Web services have constraints and capabilities of their own to consider when accessing services. This Workshop is being held to discuss the establishment of a framework for the expression of such constraints and capabilities and their association using SOAP, WSDL, HTTP, etc.

Link:
http://www.w3.org/2004/06/ws-cc-cfp.html

W3C Mobile Web Initiative Workshop, 18-19 November 2004, Barcelona, Spain

W3C is thinking of launching a 'Mobile Web Initiative'. The goal of the Mobile Web Initiative would be to make Web access from a mobile device as simple, easy and convenient as Web access from a desktop device. Initial ideas for achieving this goal include developing 'best practices' documents, providing support infrastructures for mobile developers, organizing training programs for Web content providers and creating validation and conformance testing services for Web-access from mobile devices.

Link:
http://www.w3.org/2004/09/mwi-workshop-cfp.html

W3C Workshop on Metadata for Content Adaptation, 12-13 October 2004, Dublin, Ireland

Content metadata is a way to express the meaning of Web content and to map between vocabularies of the XHTML family of markup languages, their attributes and texts, and the underlying conceptual model that they represent. Content metadata promises the ability to go beyond sharing simple tagged content to sharing some of the concepts behind the tags in the content and better device independent Web access.

Link:
http://www.w3.org/2004/06/DI-MCA-WS/cfp.html

Semantic Web Standardization in Europe: SWAD-Europe

by Dan Brickley, Libby Miller and Kate Sharp

The SWAD-Europe project nears completion. SWAD-Europe (Semantic Web Advanced Development in Europe) is a five partner collaborative EU project, part of the Information Society Technologies programme within the 5th Framework. The project has considerably contributed to the evolution of the Semantic Web.

SWAD-Europe originated within W3C's Semantic Web Activity, extending the approach taken in previous SWAD work to focus on collaborative work undertaken alongside W3C's RDF (now Semantic Web) Interest Group. The project was created to complement and feed into 'standards track' work within W3C. SWAD-Europe is a collaboration between W3C Europe, hosted by ERCIM, research institutions at CCLRC and the University of Bristol (ILRT), and the companies Stilo International Plc and Hewlett-Packard Ltd.

The Original Idea

SWAD-Europe was designed to support W3C's Semantic Web initiative in Europe, providing targeted research, demonstrations and outreach. The aim was to support the development and deployment of W3C Semantic Web specifications through implementation, research and testing activities. This involved finding and maintaining a balance between 'in-house' open source tool development, community building, outreach and evangelism, combined with research and analysis to support and field-test Semantic Web standards. In practice, this meant that more time was spent working with members of the Semantic Web Interest Group community than on in-house development work, since this approach helped establish a broader base for Semantic Web development than could be accomplished within a single European project. The Interest Group provided a forum that connected SWAD-Europe (and other SemWeb-themed projects, from Europe and elsewhere) to the more formal standardisation activities of the W3C.

SWAD-Europe's work is best understood in the context of the challenges facing the RDF project in the 1997-2001 period prior to the creation of the Semantic Web Activity at W3C. In RDF's early years, a number of difficulties needed addressing. At that time, RDF technology was perceived by some as complex and overambitious, with poorly formalised specifications. Software tools were relatively immature, with toolkit-specific APIs and query languages, which meant that application code was often tightly bound to a specific RDF toolkit. The relationship between RDF and XML in terms of both tools and standards was unclear. Despite the difficulties faced by RDF in this period, there was nevertheless a lot of enthusiasm from early adopters of RDF which emerged in the Interest Group. The design of the project was an attempt to find a way to use the structure of a European project to support both formal standardisation work at W3C, and the community fora (Interest Groups)

which can provide a home for users of Web technology to share their experiences.

The project's initial workplan included work on reviewing and developing tools to access, store, search Semantic Web data, with the goal of motivating and informing future work on RDF query languages and APIs. Practical case studies and demos were proposed, especially in 'near term' areas such as thesaurus systems and next-generation blogging tools. A fair number of workshops were included, to foster developer discussion on critical 'seed applications', eg image annotation, calendaring, social networking and geo/mapping work. In addition to these efforts, the project tried to emphasise the importance not only of standards-based technology, but on technology integration. We undertook work on linking XML to RDF, on schema languages, SVG, query, XForms and Web Services. In each case, there was the same underlying intent: to enter into dialogue with users of W3C technology and understand how their practical problems might be addressed through combing RDF approaches with other, perhaps better understood, standards-based methods. The goal, in effect, was to help demonstrate the basics of Semantic Web as a fast-maturing technology platform. Reports from each of these activity areas are available from the project Web site.

Before, During and After

Throughout SWAD-Europe, we sought to build on pre-existing work, in particular work which was conducted in public fora and was supported by freely available open source implementations. Previous collaborative efforts (by the project partners and members of the broader community) were supported through SWAD-Europe and subsequently fed into more formal activities within W3C. The project's general approach is to utilize W3C's Interest Group mechanism to bring collaborators together, test their ideas, and present the results of those collaborations in a form that can have a good prospect of adoption within the Working Group mechanisms of W3C.

What difference did SWAD-Europe make?

As always with such projects, some areas proved more rewarding than others, and the emphasis of the project evolved in response to successes and opportunities. A significant achievement of the project has been in the area of outreach and community-building. The SKOS (Knowledge Organisation Systems and the Semantic Web) work has helped to re-engage the digital library/thesaurus community. The ten workshops held as part of the project have attracted diverse participants from multiple countries and specialities, and from the research, open source and business communities. Other successful work in the project has included software development, in particular the leading open source C implementation of RDF in Redland/Raptor; and well-crafted and appealing demonstrators in the areas of Semantic Blogging and Semantic Portals, showing that Semantic Web applications can be simple, practical and easy. The pragmatic 'walk before we run' focus of the project was appreciated both by Semantic Web sceptics and by enthusiasts. Project members have also made substantial contributions to the editing and chairing of the RDF Core standards, and later, helped to establish 'Semantic Web phase two' groups at W3C: the Data Access and the Semantic Web Best Practices and Deployment Working Groups.

The overarching aim of the project was to provide, through all appropriate means, a body of answers to questions that had previously gone unanswered, and to foster grassroots communities within which such concerns are addressed. Amongst its many themes, SWAD-Europe provided detailed answers to developer questions about RDF query and storage (analysis of scalability issues; query languages; APIs), and human-oriented classification (SKOS for thesauri, bookmark sharing, semantic blogging, etc.). The project's final workshop was on the theme of FOAF, Social Networking and the Semantic Web, and illustrated some of the strengths of the project, combining presentations from academic, commercial and open source perspectives with active collaborative work on tools and applications.

One lesson from the project is that it is both important and rewarding to provide an environment where members of the larger community that surrounds W3C can interact and collaboratively explore the practical issues around Web technologies. The formal work of the W3C is based on small, highly focussed Working Groups where individuals commit a lot of time to creating new Web standards. SWAD-Europe's primary contribution was to help create a supportive background environment for such work, by allowing a much larger number of geographically-dispersed individuals to participate (through email, IRC, workshops and the Web) in the Semantic Web initiative. The project was, in the Semantic Web scene, unique in its emphasis of the practical and Web aspects of 'Semantic Web' for a Web developer audience. The support that SWAD-Europe provided to the RDF and Semantic Web Interest Groups was an important exploratory step towards a model for wider participation in Web standardisation work, showing that W3C's successful Working Group-led approach can be complimented by a broader participation model which allows individual researchers and implementors to make real contributions to the deployment of existing standards and to the creation of new ones. The challenge for the future is to work towards a Web in which all European research efforts contribute to the communities which underpin the evolution of Web standards.

Groups interested in collaborating with W3C Europe on future Semantic Web-related projects are invited to contact Dan Brickley at danbri@w3.org.

- *Dan Brickley is on the technical staff of W3C, where he serves as chair of the Semantic Web Interest Group and as SWAD-Europe Project Director.*
- *Libby Miller is the Semantic Web Group Coordinator at ILRT, University of Bristol.*
- *Kate Sharp is Project Manager for SWAD-Europe at ILRT, University of Bristol.*

Links:
SWAD-Europe reports index
http://www.w3.org/2001/sw/Europe/reports/intro.html

SWAD-Europe project homepage:
http://www.w3.org/2001/sw/Europe/

SKOS: http://www.w3.org/2004/02/skos/

Grids: The Next Generation

by Keith Jeffery and Péter Kacsuk

The topic of Grids was last addressed in Issue 45 of ERCIM News, in April 2001. At that time the field was relatively new. In the last few years, however, significant efforts have been made all over Europe to establish production Grids, in order to research every aspect and potential application of Grid systems. The goal of the current thematic issue is to show both selected results from this research and applications of Grid systems, and to show the goals of some newly launched national and EU projects.

The issue contains 32 papers that can be divided into the following main areas:

- Infrastructure
- Architecture
- Middleware
- Programming
- Applications
- New projects

For rigorous research in this field, it is necessary for Grid infrastructure to be developed. Unfortunately, few functional Grid infrastructures have so far been created in Europe. It is therefore encouraging to find four papers reporting on working Grid infrastructures, and three others detailing plans for the creation of infrastructures. The largest production Grid infrastructure in Europe is the LCG Grid developed by CERN and other physics institutes in collaboration with the EDG (European DataGrid) project. This was successfully completed in March 2004.

The first paper describes in detail the UK Computing Grid and its relationship with LCG. LCG is the largest functioning grid in the world, with over 5000 CPUs and almost 4000 TB of storage at more than 70 sites around the world. The second paper reports on the CERN openlab project, a collaboration between CERN and industrial partners to further develop the LCG Grid infrastructure using the latest industrial technologies. While the LCG's middleware had initially been targeted solely for the x86 platform, the current goal is to port the middleware to Itanium. Another major focus of the project is the '10 Gigabit challenge', ie how to achieve data transfer rates of 10 Gb/s both within the opencluster and to other sites. Other important issues include virtualization and automated installation of grid nodes. The Fraunhofer Resource Grid has been running since 2002 in Germany and combines the computing and storage infrastructure of the six institutes of the Fraunhofer Grid Alliance. This infrastructure has been successfully applied to solving engineering problems such as casting process optimization, microstructure simulation and drug design. The Hungarian ClusterGrid is a production grid that has been operating since mid-2002, and currently involves 1100 nodes and provides 500 Gflops of supercomputing power for Hungarian academic users. Its novelty is that the component clusters serve educational purposes during the day and are connected to the ClusterGrid during nights and weekends. The result is a Grid solution with the extremely inexpensive annual operation cost of 40,000 Euro.

Icatis has developed technology similar to the Hungarian ClusterGrid: during cluster usage peaks, a number of idle user machines can be aggregated to the cluster. The system is designed to be transparent to the PC owners, and the machines are only used at times when they are likely to be idle (nights, weekends, vacations etc).

Two special-purpose Grid infrastructures are under preparation and are reported in this issue. The first is the European Learning Grid infrastructure, which integrates service-oriented Grid

ARTICLES IN THIS SECTION

technology with semantic technology for the creation of future learning scenarios. The second is the DILIGENT project, which will result in a digital library infrastructure based on Grid-enabled technology. This will allow members of dynamic virtual organizations to collaborate by exploiting shared knowledge and physical resources.

Grid architecture is another exciting subject and is addressed by four papers in this issue. One of the most traditional types of Grid architecture is the cycle-harvesting grid. Work at IBM Research has focused on the performance issues of such architectures. Their performance-related work led to a definition of quality of service known as Hard Statistical QoS (HSQ) and an accompanying system architecture that can support commercial service contracts. The JGrid project is an interesting and novel approach to building service-oriented Grid architectures based on the Jini technology and Java platform of Sun Microsystems. The JGrid project extends the look-up service of Jini with a generic Grid discovery subsystem and with a compute broker. Interactive computing tasks are executed by compute services, jobs are executed by batch services, and storage services allow users to store and retrieve their data from anywhere in the world. A major obstacle to applying Grid technology in industry is the lack of a Grid-based ASP model. The GRASP project is an industry-driven European research project exploring the use of the Grid services concept for solving this problem. To achieve this, GRASP has developed an architectural framework for Grid-based Application Service Provision and a prototype realization of this framework. Since mobile communication technology is widely used in everyday life, its integration with Grid technology will radically advance the pervasiveness of Grid computing across Europe. Through the integration of these two technologies, the Akogrimo project

intends to create a commercial operator-oriented architecture and platform that support the deployment of Grid services worldwide.

Grid Middleware

Grid middleware design is a central focus of Grid research. Six papers in this issue address this extremely important subject. The first introduces a new component-based middleware approach designed as the next generation of Grid middleware. The new concept will allow automatic deployment of software components on large-scale distributed systems and grids while making the best use of available resources. The project involves the design of a runtime system, called PadicoTM, to allow components to exchange data independently of the underlying networking technologies. The appropriate data access and integration support in the middleware is an important challenge. The UK's OGSA-DAI system provides a framework and set of components operating across grids or Web services that deliver these mechanisms reliably. Using this framework, many projects are building an OGSA-DAI and extending its repertoire. It already handles a range of relational systems, XML databases and collections of files.

Unsolved security issues such as authentication and authorization remain major obstacles to the industrial implementation of grids. The research area of Grid accounting and economy systems has also been neglected. The NIKHEF and UvA teams are jointly developing a combined authentication, authorization and accounting (AAA) service technique. Based on this, site-local authorization and creation of virtual organization will become more dynamic and easier for both users and administrators. A particular advantage of this approach is that rather than starting everything from scratch, they are building on the Virtual Organization Management Service (VOMS) developed by INFN in Italy. This is one of the rare cases in which one project fertilizes another. Unfortunately not many papers report such result transfer between projects.

Resource management is a classic and central issue in Grid middleware design and it is therefore unsurprising that two papers deal with this area of research. Within the framework of the German VIOLA project, new Grid middleware components will be developed to exploit the capabilities of the new underlying optical network infrastructure. These new components will include a MetaScheduler (allowing co-allocation of compute resources, network resources with necessary QoS or other resources like visualization devices) and MetaMPICH communication libraries for distributed execution of parallel applications using MPI for inter-process communication. The GridCat group at UPC is working on integrating Grid technology and software-agent technology in order to create an agent-based Grid Resource Management system (GRM). Grids and agent communities offer two approaches to open distributed systems. Intelligent cooperation, coordination and negotiation are issues handled by software agents and are required by Grid users and Grid resource providers in order to efficiently inter-operate. Another project working on the merging of Grid and agent technologies addresses QoS problems. In their proposed architecture, self-interested agents will interact using electronic market-based techniques with the goal of establishing and maintaining a certain level of QoS in the Grid network.

High-Level Grid Programming Concepts

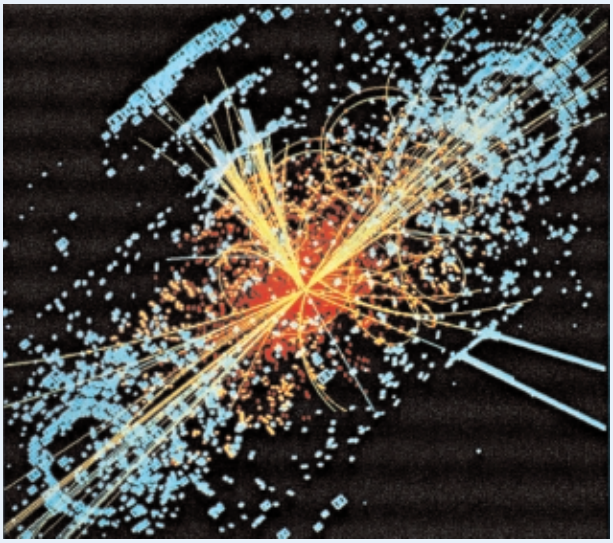
The next field covered by this issue is frequently neglected. Since there is a certain gap between the end-user needs and the access level provided by Grid middleware, in many cases end-users must learn low-level command interfaces if they wish to use the Grid. Moreover they must be fully aware of the underlying Grid middleware concepts and re-learn the concepts and access mechanisms every time a new generation of Grid middleware is introduced.

High-level Grid programming concepts and supporting tools and portals are therefore extremely important for end-users, making access to Grid services more convenient. Four papers of the current issue describe concepts, tools and portals that will make the lives of end-users much easier. The Spanish GRID-IT project introduced a new middleware design that offers an object-oriented, high-level application programming interface, which simplifies the process of remote task execution in Grid deployment. It is intended to give support for parameter sweep applications. The user is provided with a high-level API that exposes a very natural and convenient point of entry to the Grid services.

The well-known workflow concept facilitates Grid programming, and is handled by two papers. Both of these emphasize the potential commercial/business applications of grids based on their workflow concept. The first paper extends the capability of workflow systems in Grid environments, generally focusing on the capacity to enable more reliable, trustworthy and efficient use of Grid resources. One of the biggest obstacles to widespread industrial utilization of Grid technology is the existence of large numbers of applications written in legacy code, which are inaccessible as Grid services. The second paper in this area introduces a new approach known as GEMLCA (Grid Execution Management for Legacy Code Architecture), to deploy legacy codes as Grid services without modifying the original code. Moreover, such legacy code services can easily be applied by the end-user within the P-GRADE portal, providing visual workflow construction facilities. With simple drag-and-drop style programming, users can build complex workflows from GEMLCA legacy code services. These can be run on complex Grid systems such as GT-3 grids, while the low-level details of that grid remain completely hidden to the user.

Another visual service composition system is Jopera, consisting of a visual composition language and a set of integrated software development tools for composing Grid services. Jopera offers an open, flexible platform for Grid service composition that significantly facilitates the work of end-users in accessing the Grid.

The final and perhaps most important class of papers demonstrates that the



Simulation of a particle physics experiment. Read about the Large Hadron Collider Computing Grid in the articles on pages 17 and 18.

Grid really is able to contribute to the realization of e-science. They also show different application areas in which Grid technology can provide significant support for the fast and efficient solution of large scientific problems. The particle physics application has already been mentioned in the context of existing production Grid infrastructures, but other application areas are also covered.

The first paper in this class describes work on Grid-enabled Weka, a widely used toolkit for machine learning and data mining. It is a large collection of state-of-the-art machine-learning algorithms written in Java. Weka contains tools for classification, regression clustering, association rules, visualization, and data preprocessing. In Grid-enabled Weka, execution of these tasks can be distributed across several computers in an ad hoc Grid. The second describes the eMinerals project, one of the NERC e-science testbed projects. The scientific aim is to use Grid technology to advance the capabilities of molecular-scale simulations for the study of environmental processes. The third application area is extremely important and will open new horizons for medical treatments. Medical simulations and visualizations typically require computational power at a level usually unavailable in a hospital. The University of Amsterdam recently demonstrated Virtual Vascular Surgery (a virtual bypass operation), in which large-scale simulation and visualization capabilities were offered as services on a grid. Visualization can be further improved and made photorealistic through the use of Grid technology. The fourth paper reports that at VTT, Grid-based distributed computing was utilized in order to produce photorealistic images in a faster and cheaper way. A new grid known as Grix was developed based on platform-neutral architecture.

Complex problems that can only be solved in non-polynomial time are becoming common in many domains of our lives: economy, industrial environments, bioinformatics, etc. For such a wide spectrum of problems, heuristics come to the rescue, since exact techniques are unable to locate any kind of solution. The fifth paper addresses these issues, showing current research on solving optimization problems with Grid-enabled technology like Globus, Condor, Legion and SGE. Bioinformatics has already been mentioned but the last paper in this group puts it in a new context. Fully annotated data Grid services have enormous relevance in bioinformatics, especially systems biology, due to the vast number of biochemical entities and interactions that need to be kept track of. In order to tackle the problem, an annotation tool for data in Grid services has been developed by the Bioinformatics Department and the Department for Web Applications of Fraunhofer Institute.

Such a theme as this is incomplete without a look at plans for creating the next generation of grids. Four papers are reporting newly launched projects that address a variety of aspects of future research. The CoreGRID Network of Excellence project is probably the largest and most influential of these in Europe. It commenced activity on 1st September 2004 with the aim of allowing European researchers to develop next-generation Grid middleware through a high-level joint program of activities. Another important European project is GridCoord, which is a Special Support Action project. The goal is the coordination of European initiatives and research activities in Grid computing in order to strengthen cooperation between agencies planning future activities, to enhance collaboration between research and user communities, and to develop visionary national and EU programs and roadmaps.

Yet another EU project is TrustCom, an integrated project that aims to develop a framework for trust, security and contract management in dynamic virtual organizations. The framework will enable collaborative business processes to be undertaken securely in self-managed and dynamic value-chains of businesses and governments. A fourth paper describes a national Grid project from Italy. Grid.it has a strong interdisciplinary character and aims to define, implement and apply innovative solutions for networked computing-enabling platforms, which are oriented towards scalable virtual organizations.

Finally, we provide a forward-looking perspective with a short paper on Next Generation Grids, which describes the work of the Expert Group convened by the EC and also provides the URLs to the original reports of the group. Considering the importance of this subject we positioned this paper as the first article of this Special Theme section.

Naturally, one thematic issue cannot cover the entire spectrum of current and planned Grid research in Europe. However, we believe that the 32 papers briefly introduced above give a good overview of the major Grid research directions in Europe. From these papers, readers will gain an overall picture of European Grid research, and can position their own strategies on dealing with Grid systems in the future.

Please contact:
Péter Kacsuk, STZAKI, Hungary
E-mail: kacsuk@sztaki.hu

Keith Jeffery, CCLRC, UK
E-mail: K.G.Jeffery@rl.ac.uk

Next Generation Grids : The Work of the European Commission Expert Group

by Keith Jeffery

The newly-formed European Commission Directorate General 'Infomation Society' (EC DG INFSO) Unit F2 organised a workshop to discuss the future of GRIDs in Europe in January 2003. Keynote speakers included three ERCIM researchers: Thierry Priol (INRIA), Domenico Laforenza (CNR) and the author. The workshop with over 200 attendees was a success, and it was decided to form an expert group — including the 3 keynote speakers — to take the ideas forward providing a report to the EC to assist in formulation of the later part of Framework Programme (FP) 6 and the upcoming FP7.

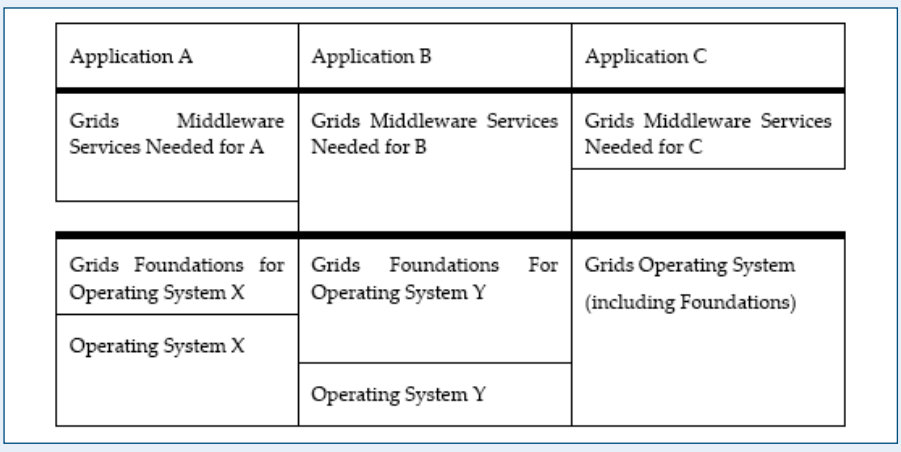
The Next Generation GRIDs first expert group (NGG1) included also from ERCIM: Seif Haridi (SICS) Ludek Matyska (CRCIM) and Jesus Labarta (SPARCIM). The report titled 'Next Generation Grids, European Grid Research 2005-2010' was produced in the summer of 2003 and is available under http://www.cordis.lu/ist/grids. The key concept was 'the invisible GRID'; the GRIDs environment should just 'be there' for use by applications in science, business, healthcare, environment, learning and culture domains.

The FP6 Call2 proposals in the GRIDs area were evaluated and 4 major — CoreGRID, AkoGrimo, NextGRID and SIMDAT — and 8 smaller projects were successful; ERCIM is well represented throughout and CoreGRID (see article in this issue) is coordinated by ERCIM. These projects started officially in summer 2004, synchronized with the European GRID Technology Days, September 2004.

However, by early 2004 it was felt that projects emerging from Call2, while covering many of the issues identified in the NGG1 report, did not cover them all. Also, increasing experience of the use of middleware designed for metacomputing (connected supercomputers) in many other aspects of IT — eg massive data handling — had revealed some inadequacies. In particular, despite the welcome evolution of GRIDs middleware towards service-oriented computing — with the adoption from the World Wide Web Consortium web services the concept of grid services defined by OGSA — the architecture of current middleware did not make it easy to address major requirements of Next Generation Grids such as self-managing, self-repairing, fault-toleration and scalability. Clearly a further tranche of research topics needed to be defined. The EC reconvened an expert group (NGG2) with some of the NGG1 members (including the original 3 from ERCIM) and some new ones with a brief to produce a new report covering – with the increased knowledge of a further year – the R&D requirements as perceived now.

The NGG2 report (published September 2004 available under http://www.cordis.lu/ist/grids) is quite radical; it proposes a new architecture stack for Grids (see Figure). This stack has applications supported by middleware (like Grid services but with many new features) and in turn by foundationware on top of existing operating systems in order to enhance existing operating systems to a common interface suitable for supporting adequately Grids middleware. More dramatically, the report proposes the development of a new Grids operating system with the features of the foundationware — and possibly also the middleware — built-in in a modular, component-based fashion so that for any device (and indeed for any particular role played by a device at any one time) only the components required are loaded. The objective is a common Grids IT surface from RFID tags through sensors, embedded and control system computers, mobile phones, PDAs, laptops, desktops, departmental servers and supercomputers with massive data stores. In fact, the vision combines aspects of traditional GRID computing (metacomputing) with both distributed and P2P (peer-to-peer) architectures with many additional novel features. The key concept is the self-managing GRIDs environment with open standard interfaces to allow multiple commercial offerings.

The vision requires the discarding of many long-cherished computer science



A New Architecture Stack for Grids.

principles. Global state can no longer be maintained and multiple local states with an interconnection and synchronization protocol is required – the so-called 'soft state' or 'relative state'. ACID (atomicity, consistency, isolation, durability) properties of database transactions upon which almost all business depends today are no longer tenable; it is impossible to maintain, eg a lock for update across multiple self-managing nodes and so new transaction models, and new transaction compensation and recovery models, are required. The 'classical' notions of security are also no longer relevant; cooperating systems with different security policies and protocols will have to negotiate trust arrangements in order to provide end-to-end security (identification, authentication and authorization). Synchronisation and real-time operation is a real issue: for some applications it is essential (eg control systems) and so special protocols and guaranteed services will be required. In an environment with millions of connected nodes, intermittent and mobile communications and semantic rich messages (with linguistic and cultural heterogeneity) carried within a formally-defined syntax there are plenty of R&D challenges for information systems engineering.

This vision will permit the realization of applications eg intelligent clothes, the intelligent home, the intelligently-assisted hospital, the remotely-controlled process plant, advisor systems for various aspects of life and business, environmental monitoring and control systems, enhanced lifelong e-learning, intelligent financial management, advanced scientific systems with modeling and activation of detectors, disaster management and many more – all of which are at best only part-realised without GRIDs. More importantly - and subject to trust, security and privacy - these systems will all interoperate as required, and with a common user interface structure, to weave the fabric of the knowledge society.

Please contact:
Keith G Jeffery, CCLRC, UK
E-mail: k.g.Jeffery@rl.ac.uk
Tel: +44 1235 44 6103

GridPP: A UK Computing Grid for Particle Physics

by Sarah Pearce

UK particle physicists have built a functioning prototype Grid, to analyse the data deluge from CERN's next particle accelerator. Over the next three years, this will be scaled up and integrated further with other Grids worldwide, to produce the world's first persistent, international, production Grid.

In 2007, CERN will introduce its Large Hadron Collider (LHC) — the world's largest particle accelerator. LHC will allow scientists to penetrate further into the structure of matter and recreate the conditions prevailing in the early universe, just after the Big Bang. But the four experiments at the LHC will produce more data than any previous coordinated human endeavour — 10 Petabytes each year, equivalent to a stack of CDs twice the height of Mount Everest. Careful analysis of all of this complex data will be required to look for some of the most elusive constituents of current physics, such as the Higgs particle and supersymmetry.

To deal with this data, LHC will use distributed computing. More than 100,000 PCs, spread at one hundred institutions across the world, will allow scientists from different countries to access the data, analyse it and work together in international collaborations. Even today, the LHC Computing Grid is the largest functioning Grid in the world, with over 5,000 CPUs and almost 4,000 TB of storage at more than 70 sites around the world. With up to 5,000 jobs being run on the LHC Computing Grid (LCG) simultaneously, it is becoming a true production Grid.

A Particle Physics Grid for the UK

GridPP is the UK's contribution to analysing this data deluge. It is a collaboration of around 100 researchers in 19 UK University particle physics groups, CCLRC and CERN. The six-year, £33m project, funded by the UK Particle Physics and Astronomy Research Council (PPARC), began in 2001 and has been working in three main areas:

- developing applications that will allow particle physicists to submit their data to the Grid for analysis
- writing middleware, which will manage the distribution of computing jobs around the grid and deal with issues such as security
- deploying computing infrastructure at sites across the UK, to build a prototype Grid.

The UK GridPP testbed currently provides over 1,000 CPUs and 1,000 TB of storage to LCG, from 12 sites in the UK. It is linked to other prototype Grids worldwide, and has been tested by analysing data from US particle physics experiments in which the UK is involved. Several other smaller experiments have also started to use the prototype Grid, and particle physicists are using it to run 'data challenges', that simulate the data analysis needed when LHC is up and running. In this way, UK particle physics has progressed 'from Web to Grid'.

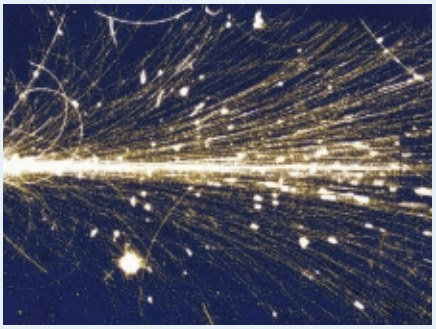
A Tiered Structure

This GridPP testbed is being developed on a hierarchical model, reflecting the overall structure of the wider LCG

testbed. CERN provides the 'Tier-0' centre, where the LHC data will be produced. GridPP has contributed £5m to the CERN for this, which has been used to support staff and buy hardware. The UK's 'Tier-1' centre at Rutherford Appleton Laboratory focuses on data storage and access. In addition there are four smaller, regional, 'Tier-2s' in the UK, with a focus on provision of computing power for generating simulated Monte Carlo data and for analysis of data by individual physicists. In addition, the Grid Operations Centre (GOC), based at RAL, monitors the operational status of resources deployed internationally through LCG and in the UK through GridPP.

GridPP2 – The Next Phase

The second phase of the GridPP project began on 1 September 2004. In the lead up to 2007, this will extend the UK particle physics grid to the equivalent of 10,000 PCs. The infrastructure in the UK will be continually tested, both by current experiments and by the LHC data challenges to ensure that the final system is ready by 2007. By the end of this second phase of GridPP, UK physicists will be analysing real data from the LHC, using the UK Grid for particle physics.



In the 1000 million short lived particles of matter and antimatter are studied each year in the LHCb particle physics experiment. In order to design the detector and to understand the physics, many millions of simulated events also have to be produced.

International Collaboration

As well as working with other international particle physics experiments, GridPP is playing a leading role in European Grid projects. During its first three years, GridPP personnel were integral to the EU-funded DataGrid project, which brought together scientists from Earth observation, bio-medicine and particle physics to create prototype a European-wide Grid. By the time of its final review in March 2004, EU DataGrid had produced around a million lines of code and had a testbed of 1,000 CPUs, which had run more than 60,000 jobs.

GridPP is now involved in the follow-on EGEE project (Enabling Grids for E-science in Europe), which aims to support the European Research Area by bringing together Grids from different countries and different disciplines.

Working beyond Particle Physics

Within the UK, GridPP2 is also collaborating with other parts of the UK's e-science programme, such as the National Grid Service. Many of the tools developed by GridPP could be useful for other disciplines – for example, GridPP is working with clinical researchers on the potential for using its computer security tools in the health service. In addition, GridPP members are collaborating with industry, sharing experience of current Grid development issues and solutions adopted.

Links:
GridPP: http://www.gridpp.ac.uk
The LCG project: http://lcg.web.cern.ch/LCG/
EGEE: http://public.eu-egee.org/
PPARC e-Science:
http://www.pparc.ac.uk/Rs/Fs/Es/intro.asp

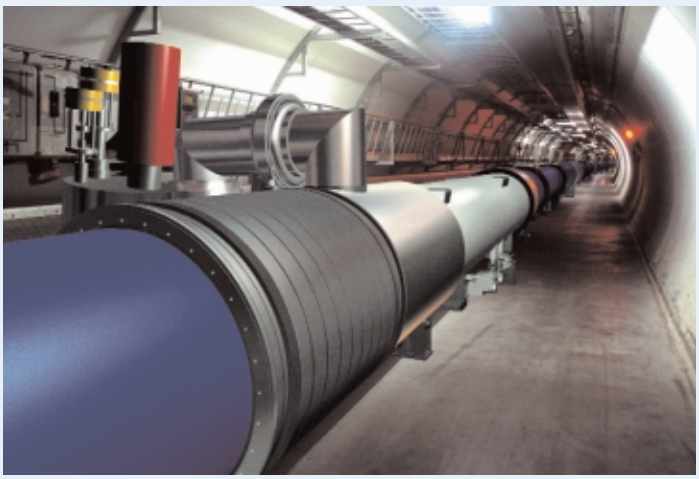
Please contact:
Sarah Pearce, GridPP Dissemination Officer
Queen Mary, University of London
Tel: +44 20 7882 5049
E-mail: s.pearce@qmul.ac.uk

The CERN openlab for DataGrid Applications

by Andreas Unterkircher, Sverre Jarp and Andreas Hirstius

The CERN openlab is a collaboration between CERN and industrial partners to develop data-intensive grid technology to be used by a worldwide community of scientists working at the next-generation Large Hadron Collider.

CERN, the European Organization for Nuclear Research, is currently constructing the Large Hadron Collider (LHC) which will be the largest and most powerful particle accelerator ever built. It will commence operation in 2007 and will run for up to two decades. Four detectors placed around the 27 km LHC tunnel on the outskirts of Geneva will produce about 15 Petabytes of data per year. Storing such huge amounts of data in a distributed fashion and making it accessible to thousands of scientists around the world requires much more than a single organization can provide. CERN has therefore launched the LHC Computing Grid (LCG), with the mission of integrating tens of thousands of computers at dozens of participating centres worldwide into a global computing resource. At the time of writing, the LCG provides about 5200 CPUs and 7.5 TB storage at 68 sites. In this context CERN established the 'CERN openlab for DataGrid Applications', a new type of partnership between CERN and industry. Its goal is to provide CERN with the latest industrial technology in order that it might anticipate possible future commodity solutions and adjust its Grid technology roadmap accordingly. The focus of openlab's research is to find solutions that go beyond prototyping, and thus provide valuable input to the LCG project. In this article we highlight some of our research projects.



Computer-generated image of the LHC tunnel.

Openlab runs a so-called opencluster comprising just over 100 Itanium-2 dual-processor HP rx2600 nodes under Linux. Several machines are equipped with 10 GbE Network Interface Cards delivered by Intel. Enterasys provided four 10-Gps 'N7' switches and IBM delivered 28TB of storage with its 'Storage Tank' file system. Openlab actively participates in data challenges run by CERN to simulate all aspects of data collection and analysis over the grid. Currently one challenge in preparation is the detection of the limits of GridFTP when sending huge amounts of data between CERN and several major Grid sites around the world. In addition we focus on an overall '10 Gigabit challenge', namely, how to achieve data transfer rates of 10 Gb/s within the opencluster as well as to other sites. This involves technological issues like network-switch tuning, Linux kernel and TCP/IP parameter tuning and so on. With a view towards cluster-based grid computing we are evaluating a 12-way InfiniBand switch from Voltaire. This piece of technology allows data to be streamed into the opencluster very quickly with minimal loss of CPU cycles.

One of our key goals has been to fully participate in LCG. However LCG's middleware had initially been targeted only for the x86 platform. We therefore had to port the middleware to Itanium, and for several months this became a major task at openlab. The LCG software stack consists of a special patched Globus 2.4 version provided by the Virtual Data Toolkit (VDT), middleware developed by the European Data Grid (EDG) project, and several tools delivered by the high energy physics community. We are able to provide a fully functional LCG grid node on Itanium systems including elements such as worker nodes, computing elements, the user interface and storage elements and resource brokers. Thanks to this effort, a new level of heterogeneity has been added to LCG. Being the first commercial member of LCG, HP sites at Puerto Rico and Bristol (UK) are contributing Itanium nodes to LCG with the technical help of openlab staff.

CERN openlab core team and contributors in June 04.

Having different platforms poses new challenges to grid computing. Scientific software is usually distributed in form of optimized binaries for every platform and sometimes even tightly coupled to specific versions of the operating system. One of our projects at openlab is to investigate the potential of virtualization within grid computing by using the virtual machine monitor Xen, developed by the University of Cambridge. A grid node executing a task should thus be able to provide exactly the environment needed by the application.

Another area of interest is automated installation of grid nodes. Originating from different sources, the installation and deployment of LCG middleware is a non-trivial task. We use the SmartFrog framework, a technology developed by HP Labs and given open-source status this year, to automatically install and manage LCG nodes. Of particular interest is the possibility of dynamically adding or deleting resources (ie worker nodes and storage elements) to and from a grid node.

Earlier this year Oracle joined openlab as a sponsor. The first project within this collaboration aims at reducing the downtime of LCG's replica catalogue, which ensures correct mapping of filenames and file identifications. The catalogue runs on the Oracle database. With the help of Oracle's technology, catalogue downtime (eg for maintenance reasons) has been reduced from hours to minutes.

Openlab's core team consists of three CERN staff members, four fellows and six summer students. Plans for the future of openlab include increasing the number of nodes and upgrading the high-speed switching environment (both Ethernet and InfiniBand). The software activities will continue to focus on participation in the LHC data challenges as well as supporting LCG on the Itanium platform.

Link:
http://cern.ch/openlab

Please contact:
Sverre Jarp,
IT Department, CERN, Switzerland
Tel: +41 22 767 4944
E-mail: Sverre.Jarp@cern.ch

Fraunhofer Grid Alliance – From Grid-Based e-Science to Industrial Applications

by Franz-Josef Pfreundt, Anette Weisbecker, Jürgen Falkner, Thilo Ernst, Uwe Der, Andreas Hoheisel, Klaus-Peter Eckert and Wolfgang Ziegler

The ICT group of German-based Fraunhofer Society, Europe's largest research organization for applied research, has announced the foundation of the Fraunhofer Grid Alliance, a network of institutes that combines the special Grid-related skills of different institutes within the Fraunhofer Society and extends the Fraunhofer Resource Grid.

In the year 2001 five Fraunhofer Institutes launched a project named 'I-Lab' which was funded by the German Ministry for Science and Education (BMBF).

This project developed an Internet laboratory software on the basis of a Fraunhofer Computing Grid. The objective of the project was to provide a user-friendly computing grid which gives users access to the grid resources and provides appropriate solutions to their computing problems. The Fraunhofer Resource Grid (FhRG) emerged from these efforts, and substantial applications have been running on the FhRG since 2002. The FhRG Web portal can be reached at http://www.fhrg.fhg.de. Today, the FhRG combines the computing and storage infrastructure of the six participating institutes: ITWM (Kaiserslautern), IAO (Stuttgart), FIRST (Berlin), IGD and SIT (both in Darmstadt) and SCAI (Sankt Augustin).

Within the I-Lab project, Fraunhofer developed a middleware package based on, but not strictly dependent on the Globus Toolkit. The result of these efforts was to a large extent made publicly available in the beginning of 2004 and is distributed through the eXeGrid Open Source Project at http://www.exegrid.net.

The software packages developed for the Fraunhofer Resource Grid and which are continually being developed further in the eXeGrid Project cover the areas of grid workflow systems, grid workflow editing, scheduling of jobs, resource brokerage, an intelligent job execution system, a Web portal with user management, management of grid resource descriptions, resource search mechanisms including the mapping of a user's tasks to suitable grid resources, and finally, an advanced security concept including the respective implementation. The basis on which all these subsystems work together is a set of resource and workflow description languages referred to as the 'Grid Application Definition Language' (GADL), and a component model designed for easy integration of legacy applications into complex grid workflows.

The Fraunhofer Resource Grid can be accessed in a number of ways, depending on the user's level of experience. Figure 1 gives a brief overview over the FhRG/eXeGrid architecture and the information flow within it.

As shown, the user can access all the grid's services through the Web portal, then going through a process of 'task mapping' that returns appropriate grid resources to solve the task defined by the user. After having chosen the resources, the user opens the Grid Job Builder – a graphical workflow editing tool – from within the portal. With this tool complex grid jobs can be created, which are then submitted for execution by the Grid Job Handler Web service. This Web service dynamically deploys grid workflows to the hardware resources that are most suitable for the execution of jobs at the time. The Grid Job Handler communicates with resource brokers and schedulers and translates the above grid workflow into GLOBUS commands so it can be run on the underlying grid infrastructure.

As shown in Figure 1, each level can in principle be bypassed, which may be helpful for advanced users or special application scenarios.

In addition to the efforts within the I-Lab project, other institutes in the Fraunhofer Society, namely Fraunhofer SCAI and Fraunhofer FOKUS, have been involved in Grid-related efforts, with SCAI covering supercomputing problems such as the development of meta-message

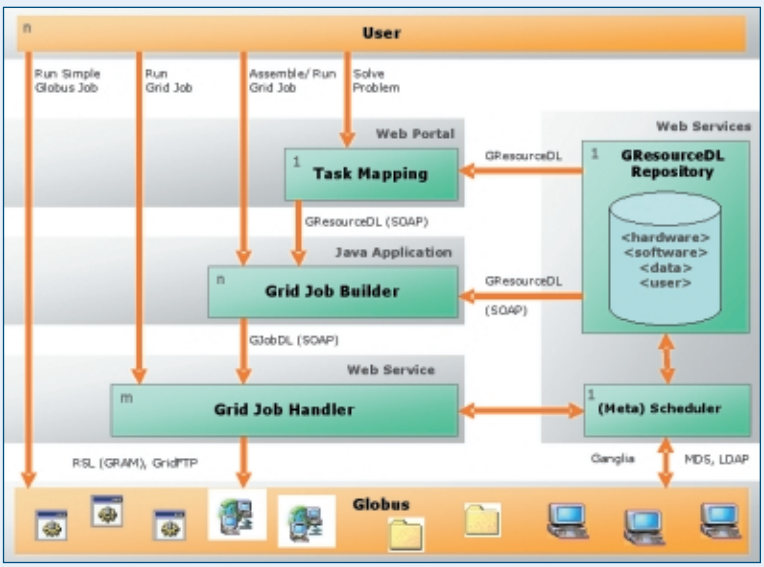


Figure 1: eXeGrid architecture and information flow.

passing and code-coupling algorithms for grid environments and meta-scheduling for co-allocation of various resources including network links or visualization devices. For topics such as the authentication, authorization and accounting of users, the definition and supervision of service-level agreements and service qualities, network monitoring, and the mobile access to Grid services, synergies can be expected between the telecommunication experience of FOKUS and the development of corresponding solutions in the Grid community. The same is true for research in the areas of programming models, component technology and autonomic communication.

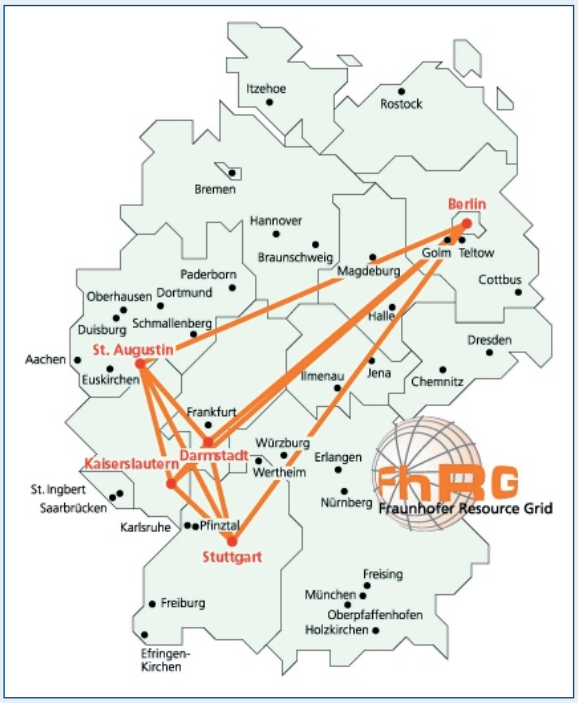


Figure 2: Fraunhofer Grid Alliance – locations connected in the Fraunhofer Resource Grid.

Together, the institutes mentioned above now form the 'Fraunhofer Grid Alliance' (see Figure 2), which coordinates the skills and strengths of the partner institutes to offer combined grid technology and services to public and industrial customers.

The members of the Fraunhofer Grid Alliance currently participate in European-level Grid projects such as EGEE (Enabling Grids for e-Science in Europe), SIMDAT (Data Grid for Process and Product Development Using Numerical Simulation and Knowledge Discovery), K-Wf Grid (Knowledge-based Workflow System for Grid Applications) and the CoreGRID European Network of Exellence. In addition, the Fraunhofer Grid Alliance is strongly involved in the German 'D-Grid' initiative for Grid-based e-Science.

The individual institutes also use technology and experience gained within Alliance activities in smaller, regional-scale projects, often involving SMEs. As an example, one of the most demanding FhRG pilot applications is the Grid-based environmental risk analysis and management system ERAMAS, which is being developed by Fraunhofer FIRST together with two environment consulting companies (http://eramas.first.fhg.de).

FhRG technology has already been successfully applied to solve engineering problems from a broad range of disciplines, for example, casting process optimization, micro-structure simulation and drug design.

Making the Grid and e-science vision a reality – not only in research, but also in industrial sectors such as engineering and life science – will remain the primary objective of the Fraunhofer Grid Alliance in years to come.

Links:
Fraunhofer Resource Grid Portal:
http://www.fhrg.fhg.de
Fraunhofer ICT Group:
http://www.iuk.fraunhofer.de/

Please contact:
Jürgen Falkner
Institute for Industrial Engineering IAO,
Fraunhofer ICT Group, Germany
E-mail: juergen.falkner@iao.fhg.de

The Production-Level Hungarian ClusterGrid Initiative

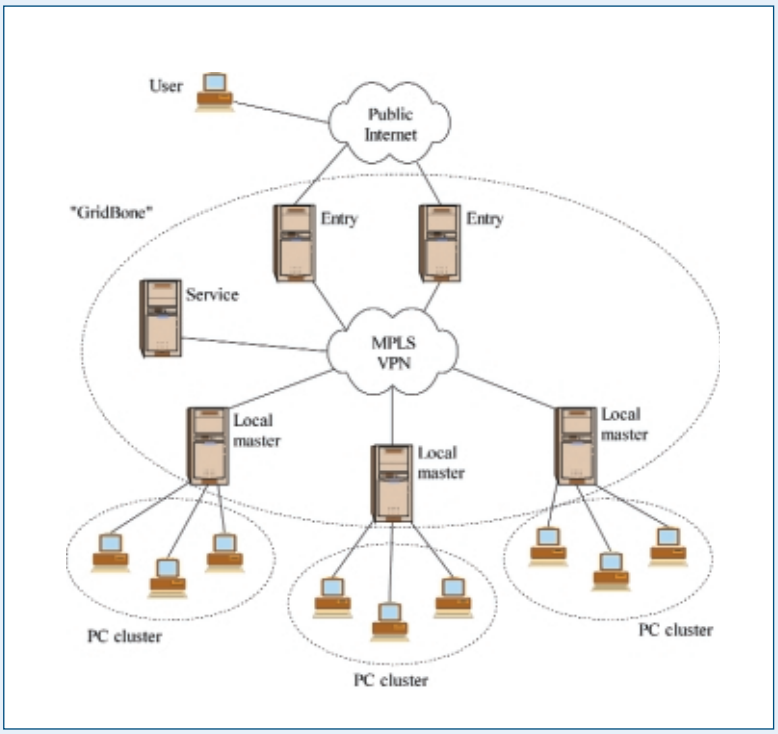
by Péter Stefán

A breakthrough represented by the Hungarian ClusterGrid was achieved when production grid development commenced in mid-2002, involving hundreds of desktop computer nodes. Launching ClusterGrid has been a result of centred research/development and organizational/management efforts. However, the investments have already returned manifold by using the 500 Gflops supercomputing power of a production PC grid in many scientific fields.

The Hungarian ClusterGrid Infrastructure handles the Grid problem in a slightly different way to contemporary grid systems, building up from low-level hardware to the application level by introducing a layered architectural model. As its name suggests, the basic building blocks of the grid are PC labs that perform dual functions. During the day, the labs, which are located at universities, high schools or even public libraries, serve an educational purpose in an office-like environment. Whenever they are fulfilling this function – typically during nights and weekends – they

are used for supercomputing in a Unix environment.

The production grid is used by researchers from many scientific fields and their industrial partners: mathematics, biology, chemistry, and data mining are a few examples of scientific areas that need large amounts of computational power and make use of such facilities with true parallel and parameter study applications. One of the most popular case studies is simulating the radiation process within the heater elements of a nuclear reactor.



Connection of grid clusters via MPLS VPN over the Hungarian Academic Network.

At the technical level, there are many novel technological elements introduced in the ClusterGrid. The first innovative feature compared to traditional grid systems is that the labs are connected to one another through private computer networking (see Figure), using the capabilities of the high-bandwidth Hungarian Academic Network in order to enhance security and user confidence in the whole system.

The second innovative element is the use of dynamic user mapping during job execution. One of the most serious causes of bottlenecks in contemporary solutions is the insufficient separation of user and job credentials, which yields monitoring and authorization problems. Furthermore, it is not necessary in the ClusterGrid architecture to have the submitter's user credentials configured at all clusters or supercomputers in the grid. This gives much more freedom to jobs traversing through different resources (in fact the job becomes an atomic unit on which different operations, such as execution, transfer, store etc can be defined).

The third innovative idea is the Web-service transaction-based, state-full, and distributed resource broker that provides interoperable gateway functionality to those grid systems built on classical disciplines using XML/SOAP interface. The broker itself contains simple implementation of all basic grid services, such as grid information systems, job execution systems, and file transfer subsystems.

The fourth innovative element is the job definition format that allows a job to be defined in both static and in dynamic terms. Jobs are built up in directory hierarchies, ie all pieces of binary, library, input and output files are encapsulated into the structure, and at the same time the job is also a temporal entity, ie a set of operating system processes on different hosts and the relationships (communication, data transfer) between them. The runtime execution structure provides the following features: the job is allowed to take its complete environment to the place of the execution (even a licence file, or organization/virtual organization certificate), the job can be customized, and workflow definitions can be treated as part of the job. Subjobs (ie jobs within the master job) can be defined and executed, and meta-jobs, such as code compilation, can easily be treated as ordinary grid jobs.

The ClusterGrid Infrastructure currently involving 1100 compute nodes has a cumulative performance of 500 Gflops/sec (50 billion floating-point operations per second), which is comparable with that of the top five hundred clusters. The system is also cost effective: the measured performance is achieved at an annual operational cost of 40 000 Euro. The framework works not only for the integration of PC clusters, but also of heterogeneous resources such as supercomputers.

In the future, NIIF/HUNGARNET plans to improve the national production grid in both qualitative and quantitative terms. This means improving the number of compute nodes to two thousand (or more), installing storage nodes and eliminating data network bottlenecks.The introduction of new technical solutions such as job gateways, SOAP-interfaced Web portals and experimental IPv6 grid technologies is also of key importance for the forthcoming development.

Link:
http://www.clustergrid.iif.hu

Please contact:
Péter Stefán, Office for National Information and Infrastructure Development (NIIF/HUNGARNET), Hungary
E-mail: stefán@niff.hu

Effective Aggregation of Idle Computing Resources for Cluster Computing

by Bruno Richard and Philippe Augerat

Modern research and industry projects require extensive computational power, which is conveniently provided by PC clusters. However, typical use of these clusters is irregular and shows some usage peaks. Researchers at Icatis are developing the ComputeMode™ software, which takes advantage of idle PCs on the corporate network and aggregates them to the cluster during usage peaks in order to reduce its load.

Computing clusters can be found in lots of companies and institutions today. Researchers and engineers have high data-processing needs and use them to distribute large jobs to a set of homogeneous machines. The Linux operating system is a de facto standard for cluster management, providing easy administration, good performance and a wide software support basis. Moreover, the set of available libraries and tools makes Linux a good choice for scientific applications.

Past projects have been focusing on how to aggregate user workstations from the enterprise network to the clusters. In this way, the company can take advantage of the availability of the processing power of user PCs. However, the real world shows that most corporate users are running Microsoft Windows™, making it difficult to aggregate user machines to the corporate cluster, which is based on Linux. Other approaches such as SETI@home or XtremWeb use a centralized distribution of computing tasks to Internet machines, but do not offer the smoothness and ease of use of a Linux cluster.

Icatis is developing the ComputeMode™ software suite, which smoothly handles this specific issue and aggregates user machines to the corporate cluster. A server is installed on the customer premises and keeps track of user PCs running Windows. During cluster usage peaks, a number of idle user machines can be aggregated to the cluster. This is done through a transparent switch of the PC to a secondary, protected mode from which it proceeds into a network boot from the ComputeMode™ server, taking advantage of the PXE protocol. This patented technology provides several benefits, such as the full isolation of the PCs' hard disks: these are not accessible while the PCs are dedicated to cluster computing. The OS and system configuration of a computing PC are also the same as a PC from the cluster, hence providing homogeneity and easing administration.

The system is designed to be very transparent to PC owners, and the machines are only used at times when the PC is likely to be idle (nights, weekends and during business trips or vacation). If the user returns unexpectedly while his/her PC is doing a computation, the user can claim the PC back, and it restores to the state in which the user left it within one minute. This includes the user session, open files, on-going programs and the desktop.

On the administration side, ComputeMode™ offers a Web administration interface to register/unregister machines to the system, handle specific system image parameters and the usage schedules for the machines (this can be done automatically), and check usage logs. The Job Management System (JMS) administration for the cluster shows additional machines in the computing pool, and priorities can be adjusted using the standard JMS configuration tools.

Users of the cluster extended through ComputeMode™ do not see much difference when ComputeMode™ is installed. Job management is done in the standard way through the JMS. The only noticeable difference is the boost in reactivity that can be expected when the cluster is heavily loaded. In such cases the PCs that ComputeMode™ aggregates to the cluster provide some extra

A ComputeMode screen shot.

computational power and processing occurs faster for the user.

Icatis is a young company, having been created in January 2004 after several years of investigation and refinement of its offer. It was successful on the commercial side: a contract has already been signed with a major oil and gas company and in June 2004 Icatis was elected a laureate in the 'Innovation-Development' national contest and won a prize from the French Agency for Innovation (http://www.anvar.fr/).

Most Icatis researchers have been working within the ID-IMAG Laboratory (http://www-id.imag.fr/), the Apache project run by INRIA, CNRS, IMAG, and UJF. ID is a French public research laboratory, which for the past twenty years has been researching concepts, algorithms and tools for high-performance, parallel and distributed computing. Successful experiments include the development of a supercomputer from standard hardware components such as those that might be found in a typical large company. An unusual supercomputer built from 225 standard PCs (733 MHz, 256 MB RAM, Ethernet 100) entered the TOP500 contest (http://www.top500.org/) in May 2001 and was ranked 385th worldwide for supercomputing. Other successful experiments such as I-Cluster and NFSp, as well as the Ka-Tools developed in close partnership with Mandrakesoft (http://www.mandrakesoft.com/) for the MandrakeClustering product (CLIC project from ID-IMAG), have built some sound technical bases for Icatis.

Icatis benefits from a strong experience in cluster computing and the Linux System, and has strong links with the high-performance computing community.

Some Icatis customers have already evaluated a ComputeMode™ prototype on their own premises. It has shown good results for peak usage absorption on an application for seismic terrain exploration, with each job running several hundreds or thousands of tasks. The full product will be released from QA in December 2004. Among other features, future versions of ComputeMode will have wider grid capabilities, such as inter-Cluster load balancing, and multiple administration roles/domains. At the same time, Icatis is working on a high-end visualization cluster.

Link:
http://www.icatis.com/

Please contact:
Philippe Augerat, Icatis CEO,
Grenoble, France
Tel: +33 6 76 27 27 92
E-mail: philippe.augerat@icatis.com

The European Learning Grid Infrastructure Integrated Project

by Pierluigi Ritrovato and Matteo Gaeta

A semantic Grid for human learning is the vision behind the European ELeGI Integrated Project for the implementation of future learning scenarios based on ubiquitous, collaborative, experiential-based and contextualized learning through the design, implementation and validation of the Learning Grid.

In recent years, teaching and learning practices have been mainly based on the information transfer paradigm. This focuses on content, and on the teacher as the key authoritative figure who provides information. Teachers' efforts have generally been devoted to finding the best ways of presenting content in order to transmit information to learners. Unfortunately the current generation of 'e-Learning solutions' has adopted the rather narrow pedagogic paradigm of 'information transfer'. This occurred simply because it is an easy way to use the Web's basic facilities. Failures, such as massive drop-out rates, are usually explained by a lack of staff awareness in the use of the Web, rather than critical reflection on the limits of this approach.

So, the question remains – how do we provide better access while maintaining or improving quality of learning through the use of ICT? The aim of ELeGI (European Learning Grid Infrastructure), an EU-funded Integrated Project (23 partners from nine EU countries), is to promote effective human learning by demonstrating and supporting a learning paradigm shift from the current idea of information transfer to one that focuses on the learner and on the construction of knowledge using experiential and collaborative learning approaches in a contextualized, personalised and ubiquitous way.

In our vision, the learner has an active and central role in the learning process. Learning activities are aimed at aiding the construction of knowledge and skills in the learner, rather than the memorisation of information. In keeping the learner at the centre of the learning process, personalisation (exploiting the existing learner's capability and skills) and individualisation (creating and adapting learning paths according to learner's preferences) become relevant aspects to be supported by technologies through the creation of the right context.

The Learning Grid

This new vision has two strong implications with respect to technology: first, teaching and learning will move toward the form of services, and second, the creation of virtual communities that provide services (eg contents, tutoring, searching for learners for sharing experi-

ences etc) will be encouraged. Technology must be selected according to these implications. ELeGI will adopt a service-oriented model, which is deeply intertwined with the use of semantic tagging and aligns itself with the global community, helping to develop trends such as OGSA (Open Grid Services Architecture), WSRF and the Next Generation Grid.

An open distributed service model is based on a concept of service (which in our context differs from a product) as a kind of predefined combination of processes yielding some result (the goal of the service) from distributed, heterogeneous, asynchronously communicating and available resources. The next generation of Grid solutions will increasingly adopt the service-oriented model for exploiting commodity technologies. The goal of this model is to enable and facilitate the transformation of information into knowledge, by humans as well as – progressively – by software agents, providing the electronic underpinning for a global society in business, government, research, science, education and entertainment (semantic aspects). We refer to these efforts as the 'semantic grid'.

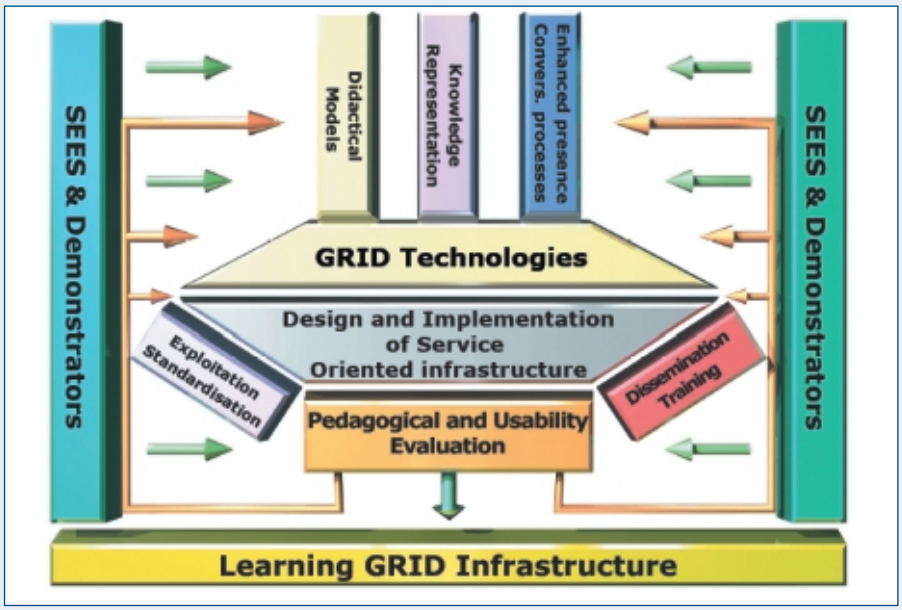
According to this technology scenario, the research at ELeGI should address the creation of a semantic grid for human learning: the Learning Grid.

The Learning Grid is a semantically enriched grid that, bringing together the features of Grid and semantic technology, represents our solution for the creation of future learning scenarios. It is based on the OGSA model, so it inherits all the features of that architecture.

As an example of the potential offered by the Learning Grid, we provide a short description of how it should support all the phases of an IMS-Learning Design specification.

Authoring tools for the production of learning scenarios can rely on knowledge-based decision-making systems to suggest pedagogical models and/or activities to suit the situation. This is done using knowledge of both the context and the people involved in the scenario (eg starting skills, personal profiles etc). Furthermore, experts can use the collaborative features of the Grid to cooperate in modelling the scenario. In this way, the Learning Grid supports the analysis, modelling and development phases of Learning Design documents.

In the delivery phase, the Learning Design document should be understood and its content executed.



ELeGI activities.

The Learning Grid uses its knowledge (represented through OWL-S ontology) to cross-reference learner preferences and the pedagogical model with the tools, resources and activities available on the Grid.

The Learning Grid makes available a learning scenario virtualized as a 'Human Learning Service', with all its 'implicit knowledge' (pedagogical model, learning goals, resources and activities involved, etc…) as a building block for the creation of more complex and interactive learning experiences composed of different scenarios and contexts.

The Methodology and Approach

The project's objectives will be accomplished following a proactive, integrated and iterative approach. Being proactive is essential due to the novelty of the technologies that will be adopted. A proactive approach will also be adopted for risk management. The Figure shows the ELeGI activities and their relationships. We have selected a set of demonstrators and testbeds representing scientific, social, economic and cultural cognate areas that include both formal and informal learning scenarios. The key difference between testbeds and demonstrators is that demonstrators already exist in non-Grid-compliant forms, whereas testbeds are principally new departures, designed to test the ELeGI approach from conception to implementation and evaluation.

Links:
EleGI Web site:
http://www.elegi.org

LeGE-WG Web site:
http://www.lege-wg.org

IMS-LD specifications:
http://www.imsglobal.org/learningdesign/

Please contact:
Pierluigi Ritrovato
Centro di Ricerca in Matematica Pura ed Applicata, Università di Salerno, Italy
Tel: +39 089 964289
E-mail: ritrovato@crmpa.unisa.it

DILIGENT: A Digital Library Infrastructure on Grid Enabled Technology

by Donatella Castelli

Building on the results of its past activity in both the Grid and Digital Library domains, ERCIM has provided the framework for the setting up of an innovative IST 6FP funded three-year Integrated Project, DILIGENT. This project, which is coordinated scientifically by ISTI-CNR, involves 14 European partners and a number of international observers. These research institutions and companies will work together on the development of a digital library infrastructure, based on Grid-enabled technology, that will allow members of dynamic virtual organizations to collaborate by exploiting shared knowledge and physical resources.

Research work today is often a collaborative effort carried out by groups of individuals belonging to different organizations remotely spread worldwide. Motivated by a common goal and funding opportunities, these groups dynamically aggregate into virtual organizations (VOs) that share their resources eg knowledge, experimental results, instruments, etc., for the duration of their collaboration, creating a rich and powerful research environment.

The DILIGENT project aims at supporting this new research operational mode by creating an advanced knowledge infrastructure that will serve the needs of dynamic virtual organizations. This infrastructure will allow members of VOs to access shared knowledge, services and computational resources in a secure, coordinated, and cost-effective way.

DILIGENT will be built by integrating Grid and Digital Library (DL) technologies. The merging of these two different technologies will result in an innovative level of functionality providing the foundations for next generation collaboration environments able to serve many different research and industrial applications.

In particular, the Grid framework will provide the context for implementing the notion of virtual digital libraries (VDLs), ie transient, on-demand DLs based on shared computational, multimedia and multi-type resources. The DILIGENT infrastructure will maintain a network of existing resources on the Grid. A virtual organization will be enabled to dynamically create and modify its own DL by specifying a number of requirements on the information space (eg publishing institutions, content domain, document type, level of replication) and on the services (eg service type, configuration, lifetime, availability, response time). A reliable and secure virtual DL that satisfies the given requirements will be transparently instantiated and made accessible to authorized users through a portal. Many virtual DLs, serving different user communities, will be active on the same shared Grid resources at the same time. The composition of a DL will be dynamic since it will depend on the currently available and registered DL resources and on many other quality parameters such as usage workload, connectivity, etc. This development model will make it possible to avoid heavy investments, long delays and radical changes in the organizations setting up these applications, thus fostering a broader use of DLs as means for communication and collaboration.

The Grid framework will also enable the provision of a number of new functions whose implementation has until now been limited by their high cost in terms of computational, storage and data transfer capacity, such as multimedia document and geographic information processing, 3D handling, spatial data manipulation, etc.

From the technical point of view, DILGENT will exploit the results of the EGEE (Enabling Grids for E-science in Europe (http://public.eu-egee.org/) project, which, in the next two years, will deliver a Grid production infrastructure shared by a very large number of European organizations. DILIGENT will first enrich this infrastructure with the necessary features for creating and handling an open and 'on-the-fly' modifiable set of services as required by the DL application framework. It will then add a number of specific services for:

- providing access to content sources, such as archives, metadata repositories and databases
- implementing typical DL functions, like search, annotation, personalization, document visualization
- supporting existing applications implemented by third-parties, like video summarization, reports generation, media-specific similarity search, etc.

The DILIGENT test-bed will be demonstrated and validated by two complementary real-life application scenarios: one from the environmental e-Science domain and the other from the cultural heritage domain. The user community of the first scenario is composed by representatives of leading organizations that operate in the environmental sector. DILIGENT will be used experimentally as a means for supporting two of the typical activities of this community: the organization of conferences and the preparation of projects and periodical reports on specific environmental topics of concern. We expect that DILIGENT will facilitate the achievement of the goals of these activities and will enhance the quality of their results by improving accessibility, interoperability and

usability of environmental data, models, tools, algorithms and instruments and by integrating the located data sources with specialized data handling services.

The user community of the second scenario consists of scholars, distributed all over the world, who have decided to work in a three year research project in order to set up the basis for a new research discipline that merges together experiences from the medical, humanity, social science, and communication research areas. DILIGENT will be experimented as a means to provide these scholars with a cost-effective instrument for setting up DLs serving their needs despite the limited duration and funding resources available for their project. The system will also be used experimentally for the organization of courses. Specific VDLs that address the knowledge needs of the students will be created by re-using material maintained in the registered archive resources of the DILIGENT infrastructure.

The DILIGENT consortium expects that the Grid-based DL infrastructure will be adopted and extended by several other communities in order to serve different application areas. Some of these communities have already been registered as project observers. Regular training sessions and workshops will be conducted to disseminate the experience and results of the DILIGENT project to both scientists and potential DILIGENT users.

Information on how to become a DILIGENT observer and about the DILIGENT activities and events can be found on the project website.

Links:
http://www.diligentproject.org
http://public.eu-egee.org/

Please contact:
Donatella Castelli, ISTI-CNR
Scientific Coordinator
Tel: +39 050 3152902
E-mail: donatella.castelli@isti.cnr.it

Jessica Michel, ERCIM office
Administrative Coordinator
Tel: +33 4 92 38 50 89
E-mail: jessica.michel@ercim.org

The Cost of Quality for Cycle-Harvesting Grids

by Chris Kenyon

Cycle-harvesting – locating unused on-line computational capacity and making it available – is an important feature of Grid resource management. Work at IBM Research has led to a quality of service (QoS) definition, called Hard Statistical QoS (HSQ), and an accompanying system architecture that can support commercial service contracts.

Business models in Grid computing dealing in buying and selling resources across budget boundaries (within or between organizations) are in their very early stages. Cycle-harvesting (or -scavenging, or -stealing) is a significant area of Grid and cluster computing with software available from several vendors. However, creating commercial contracts based on resources made available by cycle-harvesting is a significant challenge for two reasons. Firstly, the characteristics of the harvested resources are inherently stochastic. Secondly, in a commercial environment, purchasers can expect the sellers of such contracts to optimize against the quality of service (QoS) definitions provided. These challenges have been successfully met in conventional commodities, eg, Random Length Lumber (RLL), traded on financial exchanges (the Chicago Mercantile (CME, www.cme.com) in this case). The essential point for creating a commercially valuable QoS definition is to guarantee a set of statistical parameters of each and every contract instance.

The chance of achieving a good QoS level of the delivered resource slots just by taking whatever comes and passing it on (ie, basically random sampling of the underlying harvested resources) is shown in Figure 1. The error level permitted is taken as 1% of the contract size. As expected, larger contracts (more samples per quantile) make it easier to attain the fixed QoS requirement. However the probability of satisfying the quality requirements with random sampling is basically zero. Our analysis is valid when quantiles are guaranteed for any distribution: the analysis has no dependence on the underlying distribution. Thus we need a systematic way to deliver HSQ – we do this by combining stochastic (harvested) resources with directly controlled (deterministic) resources.

Stochastic QoS is a QoS parameter or set of QoS parameters that is based on statistical properties of some QoS metric. Hard Stochastic QoS (HSQ) is simply a QoS parameter or set of QoS parameters that is based on statistical properties of some QoS metric and is guaranteed with certainty. These metrics are guaranteed for each and every contract. Statistically this is the difference between guaranteeing the properties of a sample versus guaranteeing the properties of the population from which the sample was drawn. In business terms this means that each contract can have a consistent value – not just a collection of contracts or a particular provider. A typical example in conventional commodities is the specification of Random Length Lumber contracts on the Chicago Mercantile Exchange; in the IT world statistical guarantees on network usage and capacity parameters are commonplace.

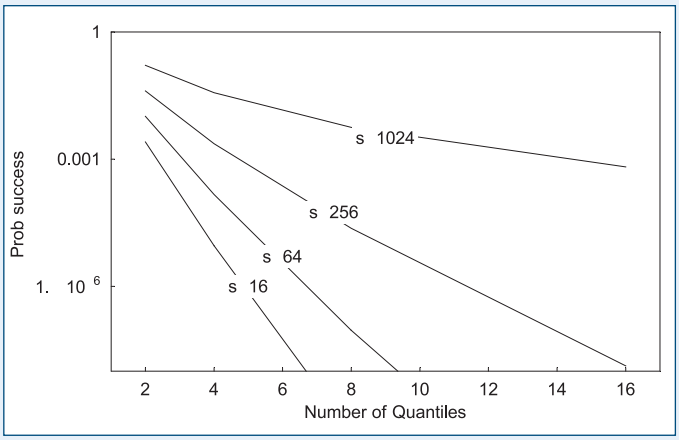


Figure 1: Probability of fulfilling a contract (Prob[success]) for a given distribution of slots using random sampling from the harvested resources relative to the precision with which the distribution is defined (number of quantiles). These results are valid for any distribution.

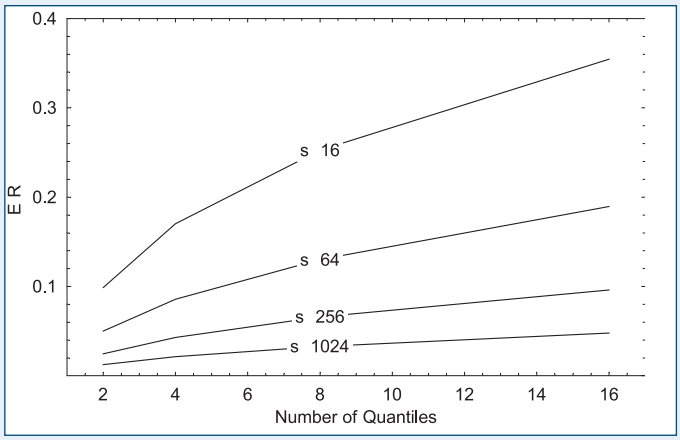


Figure 2: Expected proportion of deterministic resources required (E[R]) to guarantee a given distribution (LogNormal with mean 3 and variance 1.5) to different levels of precision as defined by the distribution's quantiles.

We have defined a system architecture for delivering HSQ. This comprises five things: an HSQ controller; a pool of stochastic (harvested) resources; a pool of deterministic (dedicated) resources; monitoring of the stochastic pool of resources; and control of the stochastic resource pool. The basic idea is that the HSQ Controller monitors the Cycle Harvesting System (CHS) resources, using the CHS Monitor, as applied to each Contract and takes one or more of two possible actions: sends an artificially 'harvested resource end' to a particular CHS resource; or diverts contract obligations from the CHS resources to the Deterministic Resources, under control of the Deterministic Controller, with appropriate instructions for their execution.

Given that resource contracts are being traded and the owner of the HSQ system sees an Ask (ie, a request for resources) with a price and a deadline, what is the cost to support that request? If it is below the price then the HSQ system owner can respond directly to the Ask. If not then the owner can post an alternative price. The cost will consist of the quantity of stochastic resources required and the relative proportion of deterministic resources to meet the quality requirements together with the costs of the two types of resources. In many situations the major cost will be the deterministic resources.

The motivation for this work was the desire to see whether it was possible to turn cycle-harvested resources into something commercially valuable. Was it possible to define a contract that was valuable given that the underlying resources are inherently stochastic and that in a commercial environment we can expect the provider of a QoS guarantee to optimize against it? We wanted to avoid all arguments based on long term reputation because we were interested in having valuable (tradable) individual contracts for cycle harvested resources – not the question of how to create valuable reputations for the providers.

We have shown how HSQ can be implemented for cycle-harvested resources using a hybrid stochastic-deterministic system. We have also described architecture, and algorithms, to support HSQ contracts. We have analyzed the algorithm behavior analytically using a distribution-free approach in terms of the expected proportion of deterministic resources required to support a given HSQ level. Thus our analysis can be applied whatever the details of a particular system happen to be.

For a particular HSQ example (see Figure 2) where time slot lengths were log-Normally distributed we found that to provide hard guarantees on 8 quantiles for contract sizes of 16 to 1024 slots, from 13% to 1% additional deterministic resources were required. Permitting oversampling, for example for a contract size of 64, was relatively inefficient, leading to up to 61% of the stochastic resources being wasted. Including downwards substitution (ie, accepting longer time slots as valid substitutes for shorter ones) reduced deterministic resource requirements by roughly half. These results are robust (change <5%) for changes in shape and scale parameters by a factor of 2 in both directions.

We have concluded that commercial service contracts with hard statistical QoS guarantees (HSQ), based on cycle-harvested resources are viable both from a conceptual point of view and quantitatively with only small (1% – 10%) requirements for deterministic (dedicated) resources.

Links:
http://www.zurich.ibm.com/grideconomics

Please contact:
Chris Kenyon, IBM Research,
Zurich Research Lab., Switzerland
E-mail: chk@zurich.ibm.com

JGrid: A Jini Technology-Based Service Grid

by Zoltán Juhász and Gergely Sipos

The JGrid system is a service-oriented grid framework that enables the construction of dynamic, self-organising grid environments that support global service discovery and provide seamless, secure access to a wide range of grid services on demand.

The aim of the JGrid project is to develop a software framework for next-generation Grid systems. We believe that future grids will be pervasive, enclosing various software and hardware components as services, and will provide simple and seamless access to services from anywhere at any time. Future grids will also be used not only for high-performance computing tasks, but also for carrying out business and assisting society and individuals in many aspects of everyday life. These systems will need to be reliable, to operate with minimal administrator supervision; self-organizing, adaptive and self-healing behaviour is mandatory.

The project is being pursued in Hungary by a consortium of two universities (University of Veszprém and Eötvös University of Sciences, Budapest), one research institute (SZTAKI) and Sun Microsystems Hungary. It is supported by the Hungarian Government under IKTA grant No. 089/2002. The project is also listed as a Jini community project, meaning it is open to the international research and development community (see the links section for addresses). Work started in 2002 and is expected to finish at the end of this year.

Among the most challenging problems posed by grid systems are discovery and coordination, that is, how to find and orchestrate a set of services operating on heterogeneous resources under different administrative control in order to solve a given problem. Grids are dynamic; services may join and leave the system at any time by their own will or due to faults. Current grid frameworks do not adequately support dynamic service discovery. Heterogeneity is addressed by creating a common (XML-based) protocol to achieve interoperability. Next-generation grid applications that use a large set of services need more however; a programming model is required that enables the creation of applications that can execute in such highly dynamic environments. This model needs to support, among other things, state-full service invocation, error handling, notification mechanism based on events and security.

The Java language is a good candidate for creating grid systems and a grid programming model. It is an object-oriented, platform-independent language with built-in security and error handling features. On its own, however, it is unsuitable for creating dynamic systems. Jini Technology – developed by Sun Microsystems in 1999 – provides the missing features for the Java platform. Jini is an infrastructure and programming model for creating dynamic distributed systems. It supports service

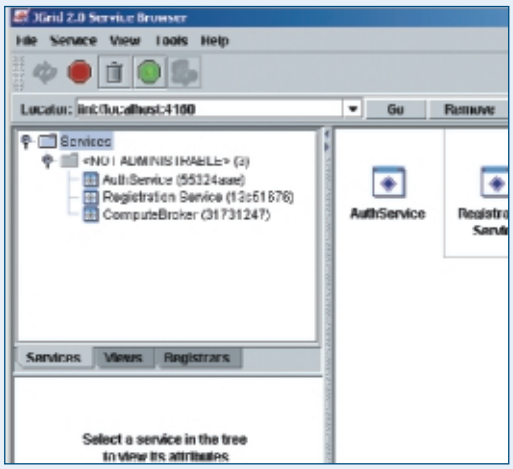


Figure 1: Part of the JGrid Service Browser graphical user interface showing discovered services.

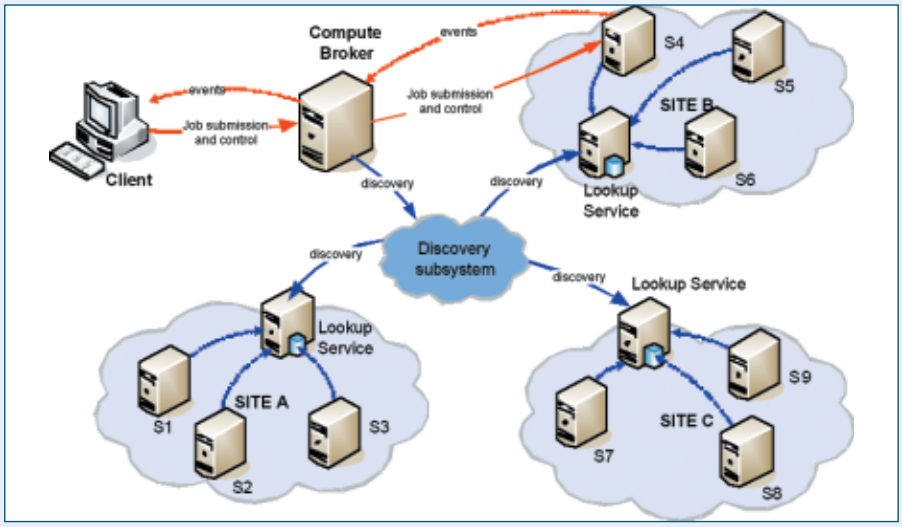


Figure 2: Overview of the job execution process in the JGrid system using a broker. The client contacts the broker and specifies the type of services sought for job execution. The broker discovers available services S1-S9 and selects the most suitable one, S4. It then submits the job to this service. Status events are sent to the broker and then to the client to monitor execution. The broker may choose alternative services in case of a service fault.

discovery, protocol-independent service access, event-based notification, distributed resource management and security, and hence is a natural fit for creating grid systems.

Currently, the main focus of the JGrid project is on computational problems; we have developed key grid services required for the construction of computational grids. JGrid users can create execute-interactive grid applications using Java tasks that execute on Compute services. Tasks can be sequential or parallel. Parallel programs can use MPI-style message passing or remote method invocations for inter-process communication. Non-Java applications such as legacy Fortran and C programs can be executed on Batch services that integrate batch runtime systems (eg Sun Grid Engine or Condor) into the JGrid framework. User data can be stored in the grid using services that allow users to store and retrieve their files and data from anywhere in the world. JGrid provides wide-area discovery that extends the Jini discovery mechanism in order to efficiently find grid services on the Internet using powerful search queries. A service browser is provided to find and display available services to the user. Discovering and controlling services may be a tedious task. The JGrid Compute Broker service relieves users from system-related tasks; a job can be submitted directly to the broker with appropriate resource requirements, and the broker will then discover suitable batch services, submit the job and supervise its execution. On detecting an error, it may take corrective measures on behalf of the user such as re-submitting the job to another batch service; on successful completion, the user receives the result. The broker provides computational users with a largely simplified interaction with the grid.

The JGrid project also provides development support for grid application programmers. The P-Grade graphical program development environment has been extended in order to generate JGrid executables from graphically designed parallel message passing programs. The P-Grade system also supports the creation of grid workflow applications on JGrid. The graphical workflow editor of the P-Grade portal system can be used to define a workflow, which is then executed under the supervision of a JGrid Workflow Manager service.

The JGrid system has been demonstrated on several international forums and can be downloaded from the project homepage. A national testbed is under construction that will allow users to experiment with JGrid services, develop and execute grid applications. Future project activities include the development of non-computational services, eg for media streaming and processing, database access and e-commerce and business applications, as well as providing access to services from mobile devices.

Links:
http://pds.irt.vein.hu/jgrid
http://jgrid.jini.org
http://www.sun.com/jini
http://www.lpds.sztaki.hu/pgrade

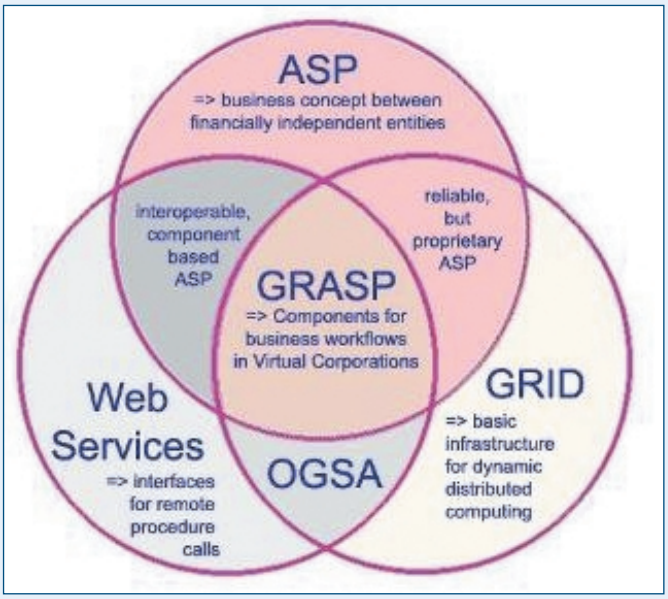
Please contact:
Zoltán Juhász,
University of Veszprém, Hungary
E-mail: juhasz@irt.vein.hu

The GRASP Project: Innovative Middleware for Next Generation Grid-Based Application Service Providers

by Theo Dimitrakos

GRASP is an industry driven European research project exploring the use of the Grid Services concept for providing an effective technological basis supporting the evolution of the ASP market towards a sustainable Utility Computing model. The intention is to improve enterprise application service provision models so as to be able to take full advantage of the flexibility offered by Web Services and the Grid for supporting dynamic resource allocation, life-time management of dynamic service instances, resources integration and efficient distributed computation.

Application service providers empowered by GRASP middleware are no longer providing a fixed solution to the customer where the business logic is hardwired in the server side code but large portions of the business logic are achieved by customising and combining private and outsourced services as needed. When properly integrated into a business workflow this bundle can soften enterprise borders, giving way to new, more flexible ways of secure and reliable collaboration. To achieve this, GRASP has developed an architectural framework for Grid-based Application Service Provision and a prototype realization of this framework. GRASP integrates concepts from application service provision and utility computing, from Service Oriented Architectures and Web Services, and a service-oriented view of Grid computing in order to empower this new generation of grid-based application service providers.



Key concepts underpinning the GRASP project vision.

Grasp Architecture

Building on top of existing OGSA compliant middleware the key functionality of Location, Instantiation and Orchestration of Grid Services is extended to handle other enterprise concepts:

- Subsystem "Location" to enable a dynamic search of Grid Services based on functional, performance, technological and business related criteria.
- Subsystem "Instantiation" to attend the properly instantiation of Grid Services performing Load Balancing and invocation of other related subsystems such as Accounting or SLA monitoring.
- Subsystem "Orchestration" to support the ASP in the provision of new business services achieved by means of the composition and coordination of existing Grid Services. It implements the workflow that encapsulates the business logic of the ASP application (described as BPEL business processes).

Further business functionality is provided as part of the GRASP Framework by means of a number of services classified in different subsystems that address core business support operations including:

- Subsystem 'SLA management' to support a SLA management system including the use of SLA templates for negotiating service provision based on QoS criteria, and to monitor the fulfilment of contract associated to each Grid Service instances regarding QoS which are based on technical and business metrics in the vision of GRASP.
- Subsystem 'Accounting' to perform the charging of service usage based on service related contracts between service users and service providers, obviously a must for ASPs. This GRASP subsystem works out the composite account of all services that are being executed on behalf of a client. It takes care of the collection of raw performance/resource data (shared with SLA monitoring); merging different costing models and applying explicitly defined pricing algorithms.
- Subsystem "Security" to offer flexible security perimeters in which GRASP infrastructure secures complete Application Service (multiple distributed component Grid Services). It handles securely life-time and addition and expulsion of component Grid service instances collectively executing a complete application.

Overall, GRASP has been designing and implementing business critical subsystems that have not been addressed as yet in other major Grid projects including EDG, Globus and Unicore.

GRASP Implementation Strategy

The current prototype of the GRASP project middleware leverages on the use of an 'open standards' technology base line including W3C web services standards and Microsoft, IBM, and BEA championed WS-* extensions (notably BPEL4WS, WS-Security, WS-SecureConversation and WS-Trust) and OGSI reference implementations. Continuing its committeemen to pioneering the convergence of Web Services and Grid Technologies for business oriented Grid applications, the GRASP consortium is currently developing a strategy for migrating its implementation base-line from compliance with OGSI standards to compliance with WSRF standards. A preliminary assessment indicates that because of the way that GRASP architecture has been designed (eg loose-coupling between Grid middleware functions and Grid-aware business and generic application components) the cost of such a migration is substantially low and we expect the next generation of GRASP middleware to support WSRF concepts.

Links:
Project Website
http://www.eu-grasp.net

GGF OGSA WG:
https://forge.gridforum.org/projects/ogsa-wg

WSRF:
http://www.oasis-open.org/committees/wsrf

OGSI.NET: http://www.cs.virginia.edu/~humphrey/GCG/ogsi.net.html

Please contact:
Theo Dimitrakos, CCLRC, UK
Tel: +44 1235 446387
E-mail: T.Dimitrakos@rl.ac.uk

Akogrimo — The Grid goes Mobile

by Stefan Wesner, Theo Dimitrakos and Keith Jeffrey

Mobility has become a central aspect of life for European citizens - in business, education, and leisure. Due to rapid technological and societal changes, there has been an astonishing growth of technologies and services for mobile users. Large investments have been made in order to provide the necessary infrastructures across Europe. In 2003 the number of Internet connected handsets or mobile hosts equalled the number of fixed Internet connected. Taking into account this evolution, Akogrimo — by leveraging the large base of mobile users — is aiming to radically advance the pervasiveness of Grid computing across Europe.

Akogrimo will bring together the market orientation and pervasiveness of mobile communication technology in everyday life with the promise of a dynamically concerted use of resources and services provided through Grid infrastructures.

By integrating the widely disjoint worlds of data communication, telecommunication and distributed service architectures, the final result of the Akogrimo exercise will be a commercial operator-oriented architecture and platform, which supports the deployment of Grid services in a worldwide perspective.

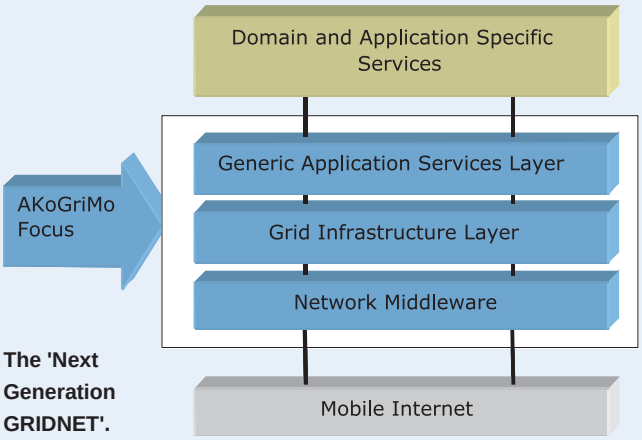
Selected scenarios showing the potentials of the innovative Akogrimo approach for the quality of life of the public citizen, for business opportunities for both small and medium enterprises as well as large companies including network providers will validate the Akogrimo architecture.

The vision of Akogrimo is accordingly a world in which:

- Grid services, pervasively available, are eventually meeting the 'everywhere at any time in any context' paradigm
- Grid services, comprising personalized knowledge and semantics, are allowing for ad-hoc, dynamic, and possibly federated formation of complex problem solving scenarios in everyday life, in business and science - to fixed, nomadic and mobile citizens
- network and service operators are able to develop new business activities and to provide profitable services in such an integrated world based on Grid and Mobile communications concepts.

Scenarios

The realization of a challenging framework merging Grid service architectures with the underlying mobile network technology requires an iterative realization approach. Iterations of the framework need one or more non-trivial scenarios that are able to validate the features and capabilities that are added to the framework. The approach within Akogrimo foresees to define, in the scenario specification phase, performance requirements and to validate early versions through prototypes or 'validation scenarios' testing a limited set of the functionality. Following the framework's evolution, the validation progresses from system component level to the complete demonstrator.



The 'Next Generation GRIDNET'.

The three scenarios foreseen so far for validation impose demands that firstly reflect requirements from 'traditional' Grid applications, namely access to distributed data and compute intensive services. In addition the scenarios demonstrate the need for mobile dynamic virtual organizations (MDVO) exhibiting the ability to dynamically adapt the organizational structure to changing local situations (through context awareness, availability of shared mobile resources). Within these MDVOs, complex workflows based upon Grid infrastructure and Grid services to access data from distributed, sometimes even mobile databases will be to dynamically established and processed. Of particular relevance is the integration of location based services and spontaneous usability.

The IP(v6) capable network supporting the scenarios and their related requirements will be constructed and derived from the infrastructure built in Stuttgart and Madrid in the context of the IST Moby Dick and Daidalos projects. This infrastructure integrates IP-based Mobility, AAA, QoS and network security on a pure Mobile IPv6 platform.

At the moment the following scenarios have been identified:

Akogrimo eLearning

The Akogrimo eLearning scenario will be embedded in the frame of the E-Learning domain. The focus in this scenario will be to build a showcase for new ways of learning that is made possible by the Akogrimo infrastructure. It is planned to liaise with the Integrated

Projects ELeGI and to utilise the Mobile University testbed of the Daidalos project in order to increase the impact.

Akogrimo eHealth

The Akogrimo eHealth testbed explores Grid technology and the Mobility paradigm in the healthcare domain. It builds on previous successful results from a 6-year German priority research program. Target users of a Grid based healthcare information system are the European citizens, mobile pervasive demand for healthcare services (eg chronic diseases), and from the healthcare service suppliers all types of healthcare service providing institutions and stationary or mobile professionals which include healthcare advisors, pharmacies, nursing services, hospitals, emergency service devices and emergency stations.

Akogrimo Disaster Handling and Crisis Management (DHCM)

Akogrimo DHCM involves incidents where various crises or disasters should be handled by rescue services and other mission-critical mobile personnel, who have to collaborate within time-critical and dangerous situations such as large sport events, concerts or special locations such as airports or railway stations.

Conclusion

Akogrimo adds a new dimension to the Grid. Mobility and network integration has not been addressed so far by other initiatives. With a merged infrastructure enabling cross-layer communication between network and Grid middleware new kind of applications are possible that attract not only researchers in specific domains but enable the provision of services for the public citizen making a pervasive Grid possible.

Links:
Akogrimo: http://www.mobilegrids.org
Moby Dick: http://www.ist-mobydick.org/
Daidalos: http://www.ist-daidalos.org
ELeGI: http://www.elegi.org

Please contact:
Antonio Sánchez Esguevillas, Telefónica, Project Coordinator
Tel: +34 983 367 577
E-mail: ajse@tid.es

Stefan Wesner, High Performance Computing Center, HLRS, Stuttgart, Germany, Project Manager
Tel: +49711 685 4275
E-mail: wesner@hlrs.de

A Component-Based Software Infrastructure for Grid Computing

by Christian Pérez

Recent advances made by researchers from the IRISA/INRIA 'PARIS' project-team in the design of runtime systems will allow automatic deployment of software components on large-scale distributed systems and grids while making the best use of the available resources (computers and networks). The design of next-generation grid middleware is based on the concept of software components and a new approach to implementing complex multi-physics applications.

The Padico project, which commenced in 2000, is investigating different aspects of software-component programming to allow the design of middleware and applications for large-scale distributed systems and grids in particular. We concentrated on the challenging problem of designing runtime systems to support efficient execution of software components within a grid. We also proposed extending an existing software component-model to support coarse-grain parallelism within a software component while maintaining scalable connections between components. Two INRIA researchers and three PhD students, within the PARIS project-team at IRISA/INRIA, were involved in this research.

The motivation behind this research was to handle the complexity of grid infrastructures by promoting the use of software components both for the design of the grid middleware itself and the design of applications. However, existing software component models (OMG CCM, Sun EJB, Microsoft DCOM) are unsuitable for grids since they were not designed to handle large-scale distributed systems, to use various networking technologies or to support coarse-grain parallelism. Moreover, automatic deployment of software components on available resources was not possible. The PARIS project-team started several research projects to deal with these problems.

One of the first projects was to extend an existing component model to support coarse-grain parallelism. We selected CORBA from the OMG as the target model simply because it was a standard with several existing open-source implementations. We first proposed an extension of the distributed object model providing distributed parallel CORBA objects aiming at encapsulating parallel SPMD codes easily and efficiently. A parallel CORBA object is just a collection of identical standard CORBA objects managed as a single entity. Data distribution and communication between parallel CORBA objects is being done transparently to the programmers. This extension of the object model implemented in PaCO/PaCO++ was later incorporated in GridCCM, a parallel extension to the component model proposed by the OMG (CCM). It is thus possible nowadays to design multi-physics applications based on the connection of parallel components, each of them solving a particular physics.

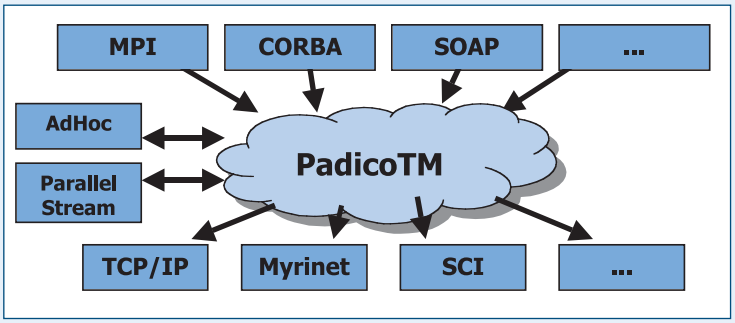


Figure 1: The PadicoTM framework decouples the runtimes from the actual networks.

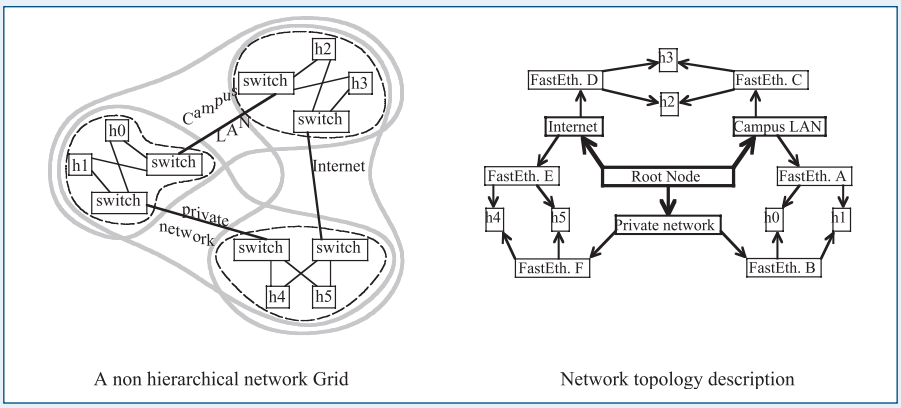


Figure 2: Example of a complex grid and its network topology description.

In parallel, we worked on the design of a runtime system called PadicoTM, to allow components to exchange data independently of the underlying networking technologies. Most of the runtime systems associated with existing component models are only able to communicate through the plain TCP/IP protocol. Efficient networking technologies found in clusters were therefore inefficiently used. Similarly, high-bandwidth Wide Area Networks are underexploited without the use of advanced features such as parallel streams. Thanks to the PadicoTM communication framework, software components can exchange data in a very efficient way whatever the networking technologies (SAN, LAN, WAN).

One of the issues under current investigation is the automatic deployment of software components onto grids. We propose a network topology description model that is capable of describing the available networking resources and their configuration. This model has been implemented in the Globus middleware within the information service. It will allow an efficient mapping of the application component graph to the grid infrastructure, taking into consideration the application requirements.

These different projects are being integrated in a single software platform based on Globus. We are cooperating with different project-teams within INRIA to experiment with real applications. One of them deals with simulation of fluid and solute transport in subsurface geological layers. It requires the coupling of several parallel codes for simulating salt-water intrusion and reactive transport. The software component model provides a convenient model for handling the complexity of the application while hiding the complexity of the grid.

This research received financial support from the ACI GRID – the French national program funding Grid research activities. Researchers involved in the Padico project are A. Denis, S. Lacour, C. Pérez, T. Priol and A. Ribes.

Please contact:
Christian Pérez, INRIA, France
Tel: +33 2 99 84 72 05
E-mail: Christian.Perez@inria.fr

Data Access and Integration

by Malcolm Atkinson

We can make decisions and discoveries with data collected within our own business or research. But we improve decisions or increase the chance and scope of discoveries when we combine information from multiple sources. Then correlations and patterns in the combined data can support new hypotheses, which can be tested and turned into useful knowledge. We see this requirement in medicine, engineering, sciences, economics and business.

An approach is to organise a communal data resource so that the data is maintained in one place, under one organisational scheme. There are three reasons why this doesn't work well for research data:

- Much of the data has already been collected under different regimes chosen by the groups who conducted the work.
- To integrate the collection process itself would lose a valuable asset – the skill and knowledge of the collectors – they understand their domain and know best how to organise it for their work.
- It is incompatible with human nature – people want to 'own' their own work and decide how it should be conducted; in health care and engi-

neering, I have seen the process of trying to agree on a common model go on indefinitely, consuming the time of many experts, but never converging.

An important trend in research is to organise larger shared data resources, but at the same time the number of collections is growing rapidly. Their diversity increases, while multi-disciplinary research increasingly requires integration of data from multiple autonomous and heterogeneous data resources managed by independent communities.

The Data Warehouse Approach

Another approach is the data warehouse – much used in retail and financial decision support for example. Here, data from multiple data resources is copied, 'cleaned' and 'integrated' under a common schema. Typically, more data is added to the data warehouse from a standard set of sources on a regular periodic basis. This works well with a small set of stable data resources serving well-defined goals. In research, there is a large and growing set of resources, an open-ended set of goals and each source changes, both in content and structure, as experiments and surveys progress and as understanding and technology advances. Data warehousing has its place in securing and concentrating data, but encouraging and capturing multi-source evolution is vital. It becomes part of the scientific communication processes and enables rapid use in other domains of new data, new discoveries and new classifications.

A virtual data warehouse (I am indebted to Mark Parsons, EPCC, Edinburgh for this term.) would meet this requirement. It would enable researchers to combine data from a dynamically varying set of evolving data sources. It would accommodate all of the diversity, handling schema integration (the differences in the way data is structured and described) and data integration (the differences in the sets of values used and their representations). It would accommodate the changes in policy of data owners, and the changes in organisation and content of each data source. And it would do all of this while presenting its users with understandable and stable facilities, that nevertheless reflect the new information. It would provide them with confidence and provenance information, so that they could use the resulting data as reliable evidence. At the same time, it would not place restrictions on the autonomy of the data providers, would assure them that their data was not misused and assist in ensuring they gained credit for the quality and content of their data.

A complete realisation of a virtual data warehouse is beyond the current state of the art, instead various aspects of it are realised. Many research teams handcraft a one-off solution. This does not scale well, as the skilled work of building schema and data mappings grows with the number and richness of the data sources integrated. It is vulnerable to being unsustainable as those resources evolve and rarely implements mechanisms to ensure data access policies and data provenance tracking. Projects, such as myGrid and Discovery Net craft workflows to assemble relevant data for each analysis. The Virtual Data Technology (Pegasus, Chimera, and DagMan) developed by the GriPhyN project, encourage high-level expression of combined data integration and analysis, enabling the underlying system to plan and evaluate more optimally. VDT exploits the predominant update pattern in physics data of incremental addition against a constant schema. Projects such as BIRN and GEON at SDSC catalogue data from multiple sources, describing their structure and data representation. For example, each geological survey's rock classification is described with an ontology and its coordinate system is defined. Tools are provided to manage these descriptions and use them to construct schema and data mappings.

Technical solutions to recurring distributed tasks underlie the assembly of data in data warehouses, communal or shared repositories, bespoke solutions, catalogued registries and in advances towards virtual data warehouses:

- data description and discovery
- access, authentication, authorisation & accounting
- data transformation
- data transport
- query, selection and aggregation
- data update, bulk load and archiving
- provenance tracking, audit trails and diagnostic logging.

The UK's OGSA-DAI (Open Grid Services Architecture Data Access and Integration) system provides a framework and set of components operating across grids or web services to deliver these mechanisms reliably. Using this framework, many projects are building on OGSA-DAI and extending its repertoire. It already handles a range of relational systems, XML databases and collections of files.

Perhaps the most important challenges that remain are:

- Automatic adaptation of schema and data transformations in response to changes in a data source.
- Integrated optimisation of computation, data management and data movement.
- high-level presentation of data composition, operations, analyses and transformations.

All three require a formal foundation, based on tractable mathematical models of sufficient scope and realism.

Link:
UK National e-Science Centre:
http://www.nesc.ac.uk/

OGSA-DAI:http://www.ogsadai.org.uk/

Please contact:
Malcolm Atkinson,
National e-Science Centre, Edinburgh, UK
E-mail: mpa@nesc.ac.uk

Building a Multi-Domain Grid

by Freek Dijkstra and David Groep

Joining resources from different administrative domains is currently a nightmare, but the problem may be alleviated with local and global authorization systems developed by the Dutch high-energy physics institute NIKHEF and the University of Amsterdam (UvA). By combining generic authentication, authorization and accounting (AAA) service techniques and site-local authorization, the creation of a 'virtual organisation' can become more dynamic and easier for both users and administrators.

It is expected that as the Grid matures, user communities will become more dynamic, shorter-lived, and based on detailed policies rather than binary access decisions. Policy-based coordination of all resources, and minimal per-resource configuration is therefore essential.

Currently, most Grid infrastructures use as their core authorization system GSI, the Grid Security Infrastructure developed by the Globus Alliance. GSI is based around the Public Key Infrastructure concept, where trusted third parties issue certificates to both users and resources. In plain GSI, access rights are directly linked to the user's identity. The resource provider keeps a list (grid-mapfile) of names of authorized users, and associates local accounts with these users. However, it is necessary to make the mappings in advance, as the grid-mapfile is maintained locally at each resource.

Unfortunately this does not scale well, since the grid-mapfile must be updated continuously as users are added to the community or 'virtual organization' (VO). It is also prone to privacy problems, since it forces one to publish a list of all users in a project to all partners, even if the user never intends to use a specific resource.

Local Credential Mapping

The local authorization services developed by NIKHEF mitigate these problems. The new model is based on the Virtual Organization Management Service VOMS, an authorization-attribute service developed in the DataTAG project by the Italian nuclear physics institute INFN. Access controls are no longer bound only to the user's identity, and decisions can be based on attributes issued to the user by the VO. The only infrastructure needed to establish a VO is this attribute authority, and at the local resource a single line of configuration suffices both to enable access for the new community and to support fine-grained policy controls defined and implemented by the VO.

Recently, the Virtual Laboratory for e-Science project in the Netherlands adopted this model for enforcing access controls to data sets spread across two hospitals, a research institute and the national computer centre.

The site aspect of the system implemented by NIKHEF consists of two parts: the Local Centre Authorization Service (LCAS) and the Local Credential Mapping Service (LCMAPS).

Apart from computing services, users may want to provision dedicated network services at the same time.

LCAS is a policy-decision module for implementing site-access policies, where the policies can be changed at run-time. It is currently a key element in enforcing security policies in many production systems, enabling site security officers to limit damage caused by compromised identities.

However, access control to the resource as a whole is not sufficient. LCMAPS, which can parse VOMS attributes, can dynamically assign specific Unix groups to fine-grained attributes. These groups are then linked to an unprivileged account for execution. The translation from VOMS attributes to Unix groups is performed on the fly. Wildcard VO group specifications give the site control over the number of Unix groups assigned to an individual VO. Sandboxing is thus achieved inside any existing Unix system.

LCMAPS supports a variety of ways to collect VO information from traditional user-proxies, VOMS attributes and grid-mapfiles, and a wide variety of account-enforcement modules for both single machines and clusters of computers.

Multiple Resources

Computing and storage are only part of the picture however, with new resource types such as remote instruments and optical networks joining the grid. This poses new challenges for resource brokers when independent organizations are involved.

Wide-area networks in particular have many different stakeholders. Packets travelling from NIKHEF in Amsterdam to, say, the Laurence Berkeley lab in San Francisco will encounter five different providers. With valuable resources such

as optical networks - your own personal light path from A to B - access control and dynamic authorization are sine qua non in the commercial world. Consequently the user running a calculation in Amsterdam, with the data in San Francisco, will need both of these resources simultaneously.

The AAA server, developed by the University of Amsterdam, fills this gap in traditional resource management. The AAA server is capable of taking authorization decisions across domains, and uses policy files to determine business logic. It can, based on the current policy, contact other services to fulfill users' requests. The attributes themselves may in fact be issued by the VOMS server, thereby seamlessly integrating network, storage and computer access.

The University of Amsterdam (UvA) has already demonstrated that the AAA server is able to make authorization decisions involving the provisioning of a dedicated network connection, even if the connection crosses multiple domains.

Now, the UvA and NIKHEF have joined forces to build a role-based authorization system for the European Grid Infrastructure, deployed by two projects: Enabling Grids for E-science in Europe (EGEE) and the Dutch Virtual Laboratory for e-Science (VL-E).

Links:
http://www.nikhef.nl/grid/
http://www.science.uva.nl/research/air/

Please contact:
Freek Dijkstra, Universiteit van Amsterdam, The Netherlands
Tel: +31 20 5257531
E-mail: fdijkstr@science.uva.nl

David Groep, NIKHEF, The Netherlands
Tel: +31 20 592 2179
E-mail: davidg@nikhef.nl

Resource Management in an Optical Grid Testbed

by Helmut Grund and Wolfgang Ziegler

The German Ministry for Research and Education (BMBF) launched the Vertically Integrated Optical Testbed for Large Applications (VIOLA) project in spring 2004. The project is managed by 'Deutsches Forschungsnetz' (DFN) and will run for three years.

Emerging new technology for optical networks will deliver QoS and bandwidth far beyond the capacity and capabilities of today's networks. For the evaluation of new network components and architecture, the integration of network techniques and applications has proven a success in former testbeds. The technical basis of the testbed is comprised of optical network components connecting compute resources, from which a grid based on the UNICORE system is built up. To allow large complex applications stressing the capabilities of the underlying optical network, additional grid components will be adopted, including a MetaScheduler (allowing co-allocation of compute resources, network resources with the necessary QoS or other resources like visualization devices) and MetaMPICH communication libraries for distributed applications using MPI for interprocess communication.

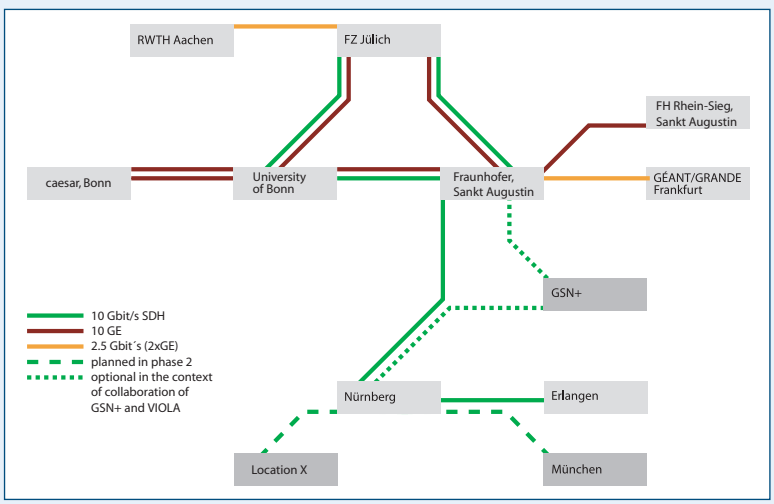


Figure 1: VIOLA Testbed Topology.

The expected outcomes of the project are threefold. First, new network techniques will be deployed and tested in an optical testbed, and along with ambitious applications will provide know-how for future generations of networks, especially the next generation of the German NREN X-WIN. Second, the enhanced Grid middleware originating from the project will become useful in the German e-Science Initiative D-Grid. The third important aspect of the project is the collaboration with other projects on a national, European and international level.

The consortium of the project is led by the DFN, ranges from research institutes and universities to the telecommunication industry, and includes Research Centre Jülich, Fraunhofer Institute for Scientific Computing and Algorithms (SCAI), Fraunhofer Institute for Media Communication (IMK), the Centre of Advanced European Studies and Research (CAESAR), RWTH Aachen University, Bonn University, the University of Applied Sciences Bonn-

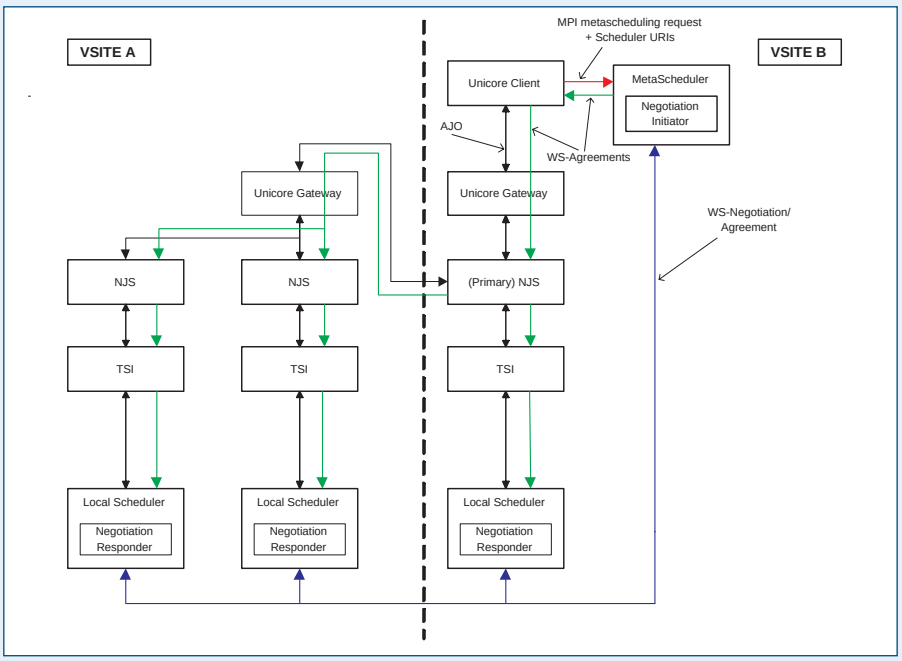


Figure 2: MetaScheduler Integration in UNICORE.

Rhine-Sieg (FHRS), Alcatel SEL AG, Siemens AG and T-Systems International GmbH.

VIOLA-Support for Set-Up and Operation of the Grid

The initial testbed providing 10 Gbit/s connection between the sites is located in the region of Cologne-Bonn-Aachen, with a connection to GEANT/GRANDE in Frankfurt and an extension to Bavaria (see Figure 1, Testbed Topology). All nodes of the testbed clusters may interconnect with their Gbit interfaces through the underlying optical testbed network. The dynamic reservation and allocation of bandwidth upon user request makes the network technology a substantial building block of evolving grids. Nevertheless today's middleware lacks the functionality to handle this and make it available to the end user and his/her applications. Enhancing existing grid middleware 'UNICORE' to create a grid infrastructure based on optical network technologies is a substantial task in VIOLA. Research Centre Jülich, Fraunhofer Institute SCAI and the University of Applied Sciences Bonn-Rhine-Sieg are responsible for this task.

Interaction with network resources is needed on two levels. The first is scheduling of network resources through dedicated (local) resource management systems in a similar way to compute resources today, whereas the resource management system for (optical) switches of the testbed implements the necessary protocol. The implementation of the resource management system is done by Fraunhofer IMK and Bonn University. The second is co-allocation of network and compute resources for a given task of the user. This will be done through implementation of WS-Agreement and WS-Negotiation as proposed by the Global Grid Forum working group GRAAP (Grid Resource Allocation Agreement Protocol).

Two requirements for local scheduling systems must be satisfied: local systems must handle the negotiation protocol, and must be able to do advance reservation of resources for a negotiated timeslot. These requirements also guarantee maximum local site autonomy, since the negotiation respects the local scheduling mechanisms and site policies remain fully effective. User authentication and mapping is handled by the local UNICORE system as usual.

Current State of the Project

A Web-service-based version of the MetaScheduler is now ready and co-allocation has been tested locally. This first version interacts with the UNICORE client, receiving the job requirements and sending back the results of the negotiations with the local scheduling systems, ie simple agreements. The UNICORE system then takes responsibility for the user's job. At the time this article will appear, the necessary modifications of the current version of UNICORE will be almost complete, and evaluation of co-allocation across multiple sites is expected to start by the end of 2004 (see Figure 2, MetaScheduler Integration in UNICORE). By the end of 2004, a stable version of MetaMPICH will also be available and initial tests running applications in the VIOLA middleware environment will take place in early 2005. Once MetaMPICH becomes available, parallel I/O will be implemented on top of it. The basic optical network connectivity will also be operable by the end of 2004.

Future Plans

In the last third of the project, a substantially revised UNICORE version based on OGSA and WSRF will be made available as a middleware layer for the VIOLA testbed. The MetaScheduling service as mentioned above will interact directly with the new UNICORE system: the UNICORE sends a request for a WS-Agreement based on the job requirements submitted by a user, the MetaScheduler negotiates the agreements through WS-Negotiation and delivers the resulting WS-Agreements back to the UNICORE system.

Finally, another objective of VIOLA is to connect to similar projects on a European and international level; these include the Canadian CANARIE, the Czech CESNET, the Polish PIONIER, the British UKERNA and the Dutch SURFnet. Efforts to establish these connections and create joint projects will be made starting in 2005.

Links:
VIOLA: http://www.viola-testbed.de

WS-Agreement/WS-Negotiation:
http://forge.gridforum.org/projects/graap-wg

UNICORE: http://www.unicore.org

Please contact:
Helmut Grund, Wolfgang Ziegler,
Institute for Scientific Computing and Algorithms SCAI, Fraunhofer ICT Group, Germany
Tel: +49 2241 14 2298/2258
E-mail: Helmut.Grund@scai.fraunhofer.de, Wolfgang.Ziegler@scai.fraunhofer.de

Grid Resource Management using Software Agents

by Isaac Chao, Ramon Sangüesa and Oscar Oardaiz

Grids and agent communities offer two approaches to open distributed systems. Grids focus on robust infrastructure and services, and agents on the operation of autonomous intelligent problem solvers. The interests of grids and agents converge when efficient management of the services provided by grids comes to the fore. Cooperation, coordination and negotiation are issues that Grid users and Grid resource providers need to address in order to successfully interoperate. Here we present work developed by the GridCat group in order to provide an agent-based Grid Resource Management (GRM) system compliant with adopted standards both for both Grid and Agent computing.

In spite of the success of Grid computing in providing solutions for a number of large-scale science and engineering problems, the problem of GRM remains an open challenge. GRM problems include the management of user requests, their matching with suitable resources through service discovery, and the scheduling of tasks on the matched resources. A high level of complexity arises from the various issues GRM must address: resource heterogeneity, site autonomy and security mechanisms, the need to negotiate between resource users and resource providers, the handling of the addition of new policies, scalability, fault tolerance and QoS, just to cite the most relevant.

Software agents are emerging as a solution for providing flexible, autonomous and intelligent software components. Multiagent Systems (MAS) are collections of software agents that interact through cooperation, coordination or negotiation in order to reach individual and common goals. There are five main dimensions that software computing is presently trying to address and that are shared by the Agent perspective:

- ubiquity
- interconnection
- intelligence (and learning)
- delegation (autonomic computing)
- human orientation.

Agents and MAS meet these requirements, and many computer scientists consider Agent-oriented programming to be the next paradigm in software computing. Grid computing pioneers, as the developers of the 'de facto' Grid middleware Globus Toolkit, agree with agent experts on the advantages that the convergence of the two approaches may bring (see their recent paper: I. Foster, N. R. Jennings and C. Kesselman, (2004). Brain meets brawn: Why Grid and agents need each other. Proc. 3rd Int. Conf. on Autonomous Agents and Multi-Agent Systems, New York, USA)

The GridCat group at the Polytechnic University of Catalonia (UPC, http://www.gridcat.org) has been working for the last two years on a Grid computing domain. This will provide Grid services to the research community formed around the i2CAT project and Foundation (http://www.i2cat.net), addressing several aspects of Grid computing. For example, Globus Grid middleware has been tested and extended, and a Grid testbed has been created on UPC resources. We are also involved in various research activities, including deployment of applications on grids, the construction of a generic Grid Framework and Portal, as well as the current work we present here to provide an agent-based architecture for GRM. Aside from these, GridCat is involved in a joint project with Sun Microsystems Spain, to develop a nation-wide Grid Infrastructure for collaborative science projects. The business and industrial sectors are also actively cooperating in the design, development and deployment of grids for the media sector in the Barcelona area.

We are currently looking at the intermingling of Grid and Agent technologies. We propose an environment based on an agent marketplace, in which agents representing Grid users can detect and meet agents representing Grid services, negotiate with them for the conditions of the service, and execute the tasks on the Grid node corresponding to the service (see Figure 1). Agents allow flexible negotiation, exchanging messages on behalf of the entities they represent. In our system, agents can use an extra information source for negotiations, consisting of a table containing the possible grid configura-

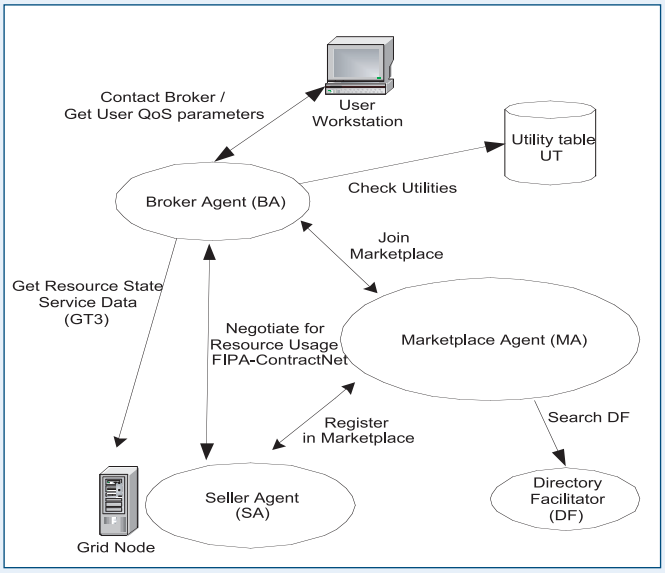


Figure 1: The Overall Architecture.

tions ranked by estimated utility. For each possible configuration, this utility table stores the expected time of task completion, given values for parameters used to model the state of the resources (eg available resources, system load, network traffic) and for execution parameters of the task (eg granularity and time between each submission of a single job). The Broker Agent (BA) representing the Grid user can check this utility table to decide which of the bids proposed by Seller Agents (SAs) is the best, and also decide how to prepare the task for the effective execution in order to minimize the time to completion.

The negotiation between agents for usage of resources is based on the FIPA-Contract-Net protocol, a common protocol used to dynamically create relationships in response to current processing requirements embodied in a contract. The contract is an agreement between a manager (the Broker) and a contractor (the Seller Agent acting on behalf of a Grid resource), resulting from the contractor successfully bidding for the contract.

The agents are implemented using JADE Framework supporting FIPA standard for agent architectures, and they behave in a coordinated manner by means of the sequential execution of their behaviours (see Figure 2). We use Globus Toolkit 3 as Grid middleware for the interaction between agents and the Grid services and for the effective scheduling of jobs into the available resources.

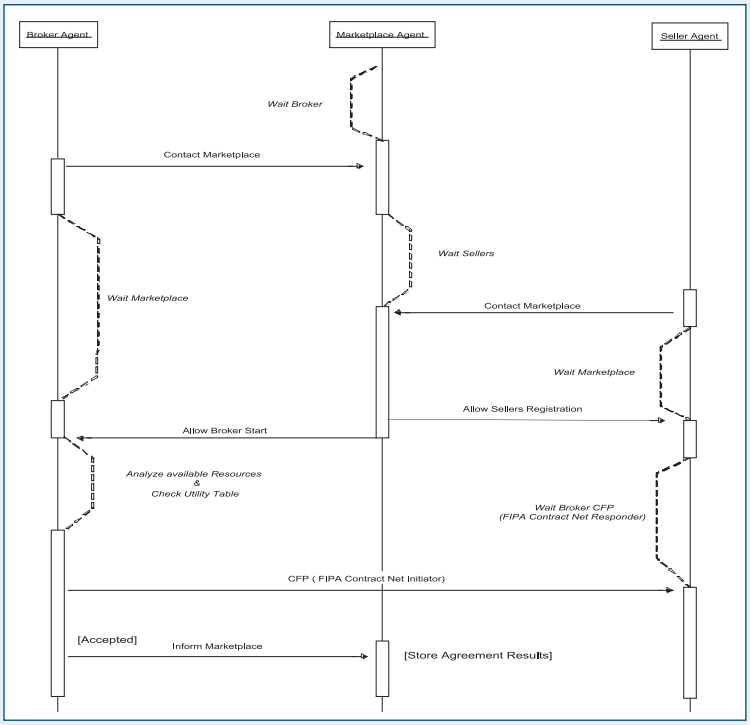


Figure 2: The Agents operation Sequence Diagram.

Future work will cover the inclusion of economic models in the negotiations, as well as an additional layer of evolutionary learning embodied in the agents in order to efficiently build the utility table, thereby allowing the architecture to be ported to real-world operative grids.

Link:
http://www.gridcat.org

Please Contact:
Isaac Chao, Ramon Sanguesa, Oscar Oardaiz, Polytechnic University of Catalonia, Barcelona, Spain
Tel: + 34 93 401 71 11
E-mail: ichao@lsi.upc.es, sanguesa@lsi.upc.es, oardaiz@ac.upc.es

Multi-Agent Systems for Efficient Quality of Service Routing in Grids

by Eric Ayienga , Bernard Manderick, Okello-Odongo and Ann Nowe

Research carried out in cooperation between University of Nairobi and the Vrije Universiteit Brussel proposes to use multi-agent systems (MAS) for provision of QoS at the network level in a Grid environment. In the proposed architecture, self-interested agents will interact using electronic market-based techniques with the aim of establishing and maintaining a certain level of QoS in the Grid network.

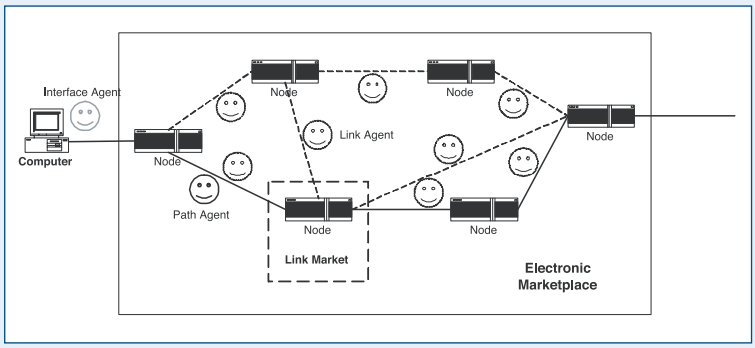
The Computational Grid is a piece of infrastructure formed from a variety of computational resources interconnected by wide area network (WAN) links. The current Internet cannot adequately support Grid services, however, because it is only able to offer 'best effort' service with no resource reservation. This type of service is adequate for traditional applications like FTP, Telnet and email but is intolerable for emerging real-time, multimedia and Grid applications (eg Internet telephony, video-conferencing, video on-demand etc), which require high throughput (bandwidth), low latency (delay and jitter) and low packet-loss rate (reliability). This calls for the need to provision Quality-of-Service (QoS). Provisioning of QoS in a network requires its definition and deployment using various mechanisms. Network QoS can be defined as 'a set of service requirements to be met by a network while transporting a flow'. The requirements can be qualitative or quantitative. Qualitatively, QoS is defined through user perceptions and require-

ments. Quantitatively QoS is defined through the specification of requirements in terms of constraints on various quantifiable metrics such as bandwidth, delay, delay jitter, reliability and cost. Its deployment is done through admission control, scheduling, routing, flow-control and policing. Attempts to satisfy the service requirements are being made by IETF, IEEE etc, through Bandwidth Overprovisioning, QoS Architectures and Traffic Engineering.

Since 'best-effort' will continue to be dominant however, all QoS mechanisms are layered on top of the existing Internet rather than replacing it with a new infrastructure. The reason for this dominance is that the infrastructure already exists, and the routing protocols and algorithms are reliable and stable for the applications it was designed for.

Two of the key issues in supporting these QoS campaigns in communication networks are QoS specification and routing. QoS specification specifies the requirements needed for QoS and quantifies them as accurately as possible. QoS routing on the other hand is concerned with routing traffic such that the QoS requirements of the carried traffic are met. In trying to find the 'shortest path' through the network, the current Internet routing protocols (OSPF, RIP, BGP) have two limitations. First, they use single objective optimization algorithms which consider a single arbitrary metric such as hop count or administrative weight during the path selection process. This may lead to congestion on this path while alternate paths with acceptable but non-optimal conditions are relatively free. Second, 'best-effort' algorithms tend to shift traffic from one path to a 'better' path whenever such a path is found. QoS-based routing should be designed to avoid these problems and in addition should take into account the applications' requirement and the availability of network resources. However, QoS routing poses several challenges that must be addressed to enable the support of advanced services in the Internet, both at intra and inter-domain levels.

Contemporary trends are moving towards the provision of QoS through intelligent resource allocation while routing traffic using machine-learning techniques. This research will approach this challenge through the use of multi-agent systems (MAS). The motivation the use of MAS is that grids are open systems. In these systems, control is distributed, characteristics are not known in advance, and are dynamic and heterogeneous. MAS can address this problem by computing solutions locally based on limited information from isolated parts of the system, and use this information in a social way. Such locality enables agents to respond rapidly to changes in the network state, while their social nature can potentially enable their actions to be coordinated to achieve some wider and socially desirable effect. MAS have been known to solve problems that have the property of inherent distribution (physically and geographically). They also solve problems requiring the interconnection and interoperation of multiple autonomous, self-interested legacy systems.



The MAS Architecture.

In the proposed architecture, agents will interact in the Grid network infrastructure with the aim of establishing and maintaining a certain level of QoS. In allocating resources to implement QoS in Grids, the MAS properties to be exploited include the following:

- Agency: agents act as representatives for other entities with the express purpose of performing specific acts that are seen to be beneficial to the represented entity.
- Autonomy: agents will make autonomous decisions based on their intelligence and social ability.
- Interaction: agents will interact with users, other agents and the environment through coordination (this can be cooperation or competition), communication, and negotiation.

The proposed MAS-based solution will be decentralized and the predominant MAS structure will be the market; in this case the Grid environment in which there will be buyers and sellers of resources. Link agents will be the producers of the resources in this economy. They will wish to maximize the income they get by selling network resources to path agents. Path agents will be the buyers of link resources and sellers of path resources. Negotiations between buyer and seller agents will be based on the market economy and will be modelled using evolutionary game theory. The MAS architecture is as shown in the Figure.

When a request for a service is made, the Interface Agent will map the service request to the QoS parameters and come up with a quantitative value for the service. This value will be used to derive another quantitative value for the resources. which the Path Agent will use to competitively negotiate for a link with the Link Agents. Upon agreement, the Path Agent will use this link to move to the next node and repeat a similar process. In this way a route with the right QoS parameters for a service request will be formed through the network.

Initially, the system will be modelled through simulation. A simulator representing the current Internet will be built, and the MAS will be built on top of this simulator as an additional layer of control. Challenges include finding suitable paradigms for agent autonomy, multi-agent communication techniques, agent communication platform and language, and the many layers involved that will lead to a high communication and protocol overhead.

Please contact:
Bernard Manderick, Vrije Universiteit Brussel
E-mail: Bernard.Manderick@vub.ac.be

Enabling High-Level Access to Grid Computing Services

by José Miguel Alonso, Vicente Hernández, Germán Moltó

The Globus Toolkit is considered the de facto standard for Grid computing. However, its steep learning curve has discouraged its widespread adoption. In order to overcome this disadvantage an object-oriented software has been developed, which provides an intuitive interface that completely hides the direct interaction with Globus.

Grid computing technology enables the collaborative usage of computational resources across different organizations. Nowadays, the Globus Toolkit represents the industrial standard for Grid computing, and is involved in a large number of research projects. Even though Globus offers all the necessary tools to perform Grid-based executions, its steep learning curve means an intense training process is required before its benefits can be fully exploited in each particular area of interest.

In the framework of the project GRID-IT (TIC2003-01318), funded by the Spanish Ministry of Science and Technology, a middleware has been developed by the Networking and High performance Computing Research Group (GRyCAP) of the Valencia University of Technology in order to provide easy and transparent access to Grid computing facilities.

Purpose and Description of the Middleware

This middleware offers an object-oriented, high-level application programming interface which allows the process of remote task execution in a Grid deployment to be simplified. It is intended to give support to parameter sweep applications, which involve the execution of multiple instances of a task. These sorts of executions are typically resource-starved and thus Grid computing offers significant benefits by enlarging the computational capabilities of a single organization with new computational resources from abroad.

As represented in Figure 1, the middleware was developed on top of the Java Commodity Grid Kit, allowing the user to completely avoid direct interaction with Globus services such as GRAM (for job execution), GridFTP (for high-performance file transfer), or MDS (for resource characteristics discovery). Instead, the user is provided with a high-level API that exposes a very natural and convenient point of entry to the Grid services.

It is important to point out that, with the direct interaction with Globus, Grid users must combine all the services provided by the toolkit in order to achieve their purpose, which generally represents remote task execution. This implies that the users must concentrate on how to perform the execution, rather than on what to execute, which is in factwhat they are interested in. Since it is the purpose of the Grid to offer advantages to users, the easier it is to use this technology, the sooner the benefits are achieved.

Figure 2 is a diagram showing some of the most important classes of the middleware. The user directly interacts with instances of these intuitive classes. For example, a GridTask represents an abstraction for a task that must be executed in the Grid, a GridResource represents an abstraction for a computational resource in the Grid infrastructure and the BasicScheduler provides scheduling capabilities for the allocation of GridTasks to GridResources.

With this middleware, the user need only describe the application and the supporting classes involve the entire infrastructure to achieve the remote task execution.

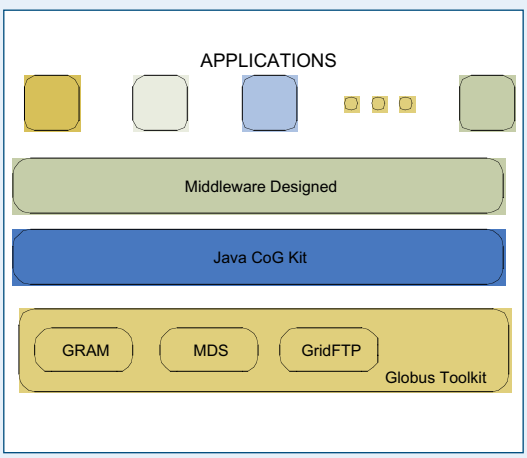


Figure 1: Middleware location with respect to the applications and Globus Toolkit.

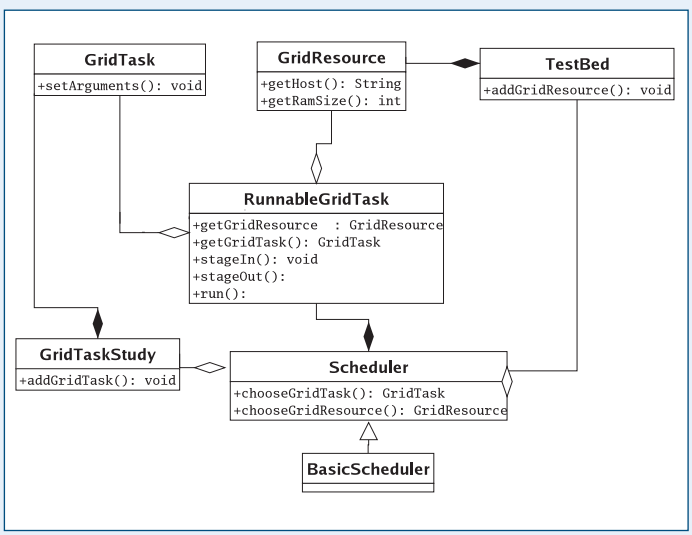


Figure 2: Simplified class diagram of the middleware.

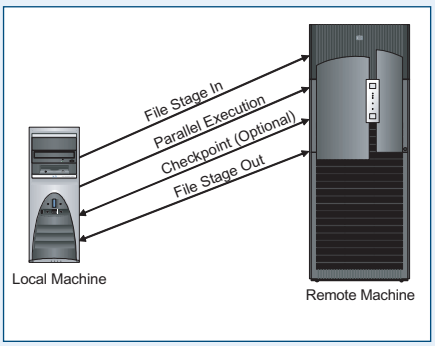


Figure 3: Principal steps performed to achieve remote task execution.

For each task, the middleware allows the dependent input file set to be described and appropriately staged into the resource before execution. The output files can also be specified, and they will be transferred back to the local machine after execution. The task may also specify an a priori quality of service, requesting a minimum amount of available RAM or a minimum number of processors in the remote resource, and refusing to be executed if these are not fulfilled.

The middleware also provides fault-tolerant dynamic scheduling capabilities for the allocation of tasks to resources. Fault tolerance is achieved by the application-dependent checkpointing support, periodically retrieving the checkpoint files to the local machine, so that an execution can be resumed in another resource in the case of failure.

Applications

This middleware is currently being used by the GRyCAP in a number of computational fields. First of all, in the area of biomedical applications, the study of the cardiac electrical activity, especially under pathological conditions, requires the execution of several parametric simulations in order to analyse, for example, the influence of certain anti-arrhythmic drugs.

On the other hand, the structural analysis of buildings in civil engineering design can require the simulation of a large number of different structural alternatives under diverse load conditions, in order to find that which best accomplishes with all the economic limitations, design aspects and safety requirements.

In conclusion, the development of high-level interfaces to Grid computing facilities greatly enhances the usability of and interest in this technology, allowing wider adoption of the Grid by non-experts.

Link:
http://www.grycap.upv.es

Please contact:
José Miguel Alonso, Vicente Hernández, Germán Moltó, Universidad Politécnica de Valencia/SpaRCIM, Spain
Tel: +34 963 87 73 56
E-mail: {jmalonso,vhernand,gmolto}@dsic.upv.es

Trusted Grid Workflows: Assigning 'Reputation' to Grid Service Providers

by Omer Rana, Luc Moreau and Julian Padget

Three Projects (PASOA, GENSS, SERENITI) extend the capability of workflow systems in Grid environments — generally focusing on the capacity to enable more reliable, trustworthy and efficient use of Grid resources. Such issues are also significant to encourage greater use of Grid infrastructure by commercial/business users.

Creating applications by combining independently-managed (often distributed) services forms a key theme in Grid computing. Such services may be hosted by different institutions, and utilise a variety of compute and data resources. Workflow tools play an important role in helping compose these services, and considerable effort has been put into creating workflow tools that combine ease-of-use with support for the widest range of Grid middleware. Most Grid-centric workflow tools assume the existence of a static set of services, available at the time of launching the tool. To exploit fully the dynamic aspects of Grid infrastructure, it is necessary to allow for dynamic services to be deployed and discovered.

In automated discovery techniques, services have a rich semantic description and a matchmaking algorithm attempts to match the service request with service advertisements. The match is then classified as an exact match, or several flavours of 'relaxed' match. It is important to note that often 'automated' discovery also requires having a human in the loop — as it is seldom possible to specify user needs in an exact way (often users are also not fully aware of their requirements). It is intended that the identification of 'sufficiently similar' services to a query is often likely to lead to a discovery of more suitable services. The 'Grid Enabled Numerical and Symbolic Services' (GENSS) project is an EPSRC-funded project between the University of Bath and the Welsh eScience Centre (WeSC) at Cardiff University. The project builds on the development of Mathematical Web Services in the European Commission funded MONET (http://monet.nag.co.uk/) project. The GENSS project addresses the creation of matchmaking techniques to allow users to discover Mathematical services described in terms of an OpenMath-based ontology. Mechanisms to allow such services to be delivered as part of a workflow for mathematical problem analysis is also being undertaken in the project. Such service discovery, by necessity, needs to be problem-specific. By focusing on mathematical services, we intend to address the requirements of a much wider range of users in science and engineering —than perhaps by focusing on a particular application domain (such as BioSciences, for instance).

Once suitable services have been discovered, it is often difficult for the user to ascertain if the service selected will perform as expected. This introduces the issues of trust and reputation in services made available over Grid infrastructure. Trust issues may be considered from the service provider's or the service consumer's perspective. For the provider, trust has always been associated with security -- focusing on authenticating and authorising users accessing the provider's resources. For the consumer, trust is considered as a predictability issue addressing concerns such as which service provider is truly reliable in delivering a service, and how much credence should be given to that provider to deliver the service as advertised. The concepts of trust and reputation are complex and many-faceted issues — and associated with themes such as risk, competence, beliefs/perceptions, utility/benefit and expertise. Two projects at WeSC are attempting to address these concerns:

- PASOA (Provenance Aware Service Oriented Architecture), led by the University of Southampton (EPSRC-funded)
- SERENITI (SERvice-ENabled Interactive Trust Informatics), led by WeSC.

The PASOA project is investigating how provenance data may be associated with a workflow process (process provenance) and with each service (service provenance). The problems of determining the origin of a result, especially when it involves processing through a number of stages, or deciding when results of analysis are no longer valid become important concerns in open Grid environments and in establishing the quality and reproducibility of scientific applications. The PASOA project therefore is aiming to define execution and service provenance in relation to workflow enactment, to enable better use of this data for reasoning about the 'processes' involved in a scientific collaboration over the Grid.

The SERENITI project on the other hand is exploring how the concepts of trust and reputation are related to each other, and especially how a reputation index' may be assigned to a group of services, based on the reputation of each service within the group. The key theme in the project is the observation that in a workflow session there is often a set of 'critical' services which must perform successfully. These critical services are also only available from a restricted set of providers (such as a particular numeric solver that can only execute on a restricted set of platforms). Allocating a reputation-index to such critical services, and also determining how this value changes over time is a key factor in assessing risk for the entire workflow.

Links:
GENSS: http://genss.cs.bath.ac.uk/
PASOA: http://www.pasoa.org/

Please contact:
Omer F. Rana, Cardiff University, Wales, UK
Tel: +44 29 2087 5542
E-mail: o.f.rana@cs.cardiff.ac.uk

Julian Padget, University of Bath,
E-mail: jap@cs.bath.ac.uk

Luc Moreau, University of Southampton, UK
E-mail: l.moreau@ecs.soton.ac.uk

A High-Level Grid Application Environment to Grid-Enable Legacy Code

by Péter Kacsuk, Tamás Kiss, Ariel Goyeneche, Thierry Delaitre, Zoltán Farkas and Tamás Boczkó

One of the biggest obstacles in the widespread industrial take-up of Grid technology is the existence of a large number of legacy code programs that are not accessible as Grid services. We introduce a new approach – Grid Execution Management for Legacy Code Architecture (GEMLCA) – to deploy legacy codes as Grid services without modifying the original code. By integrating GEMLCA with a workflow execution-oriented Grid portal, the P-GRADE portal, Grid services based on such legacy code can be applied in complex business processes.

There are many efforts worldwide to provide new Grid middleware concepts for constructing large production grids. As a result, the Grid community is at thestage of producing third-generation Grid systems represented by the OGSA (Open Grid Services Architecture) and WSRF (Web Services Resource Framework) standards. On the other hand relatively little attention has been paid to how end-users can survive in the rapidly changing world of Grid generations. The primary goal of the research teams of University of Westminster and SZTAKI within the UK e-Science OGSA Testbed Project was to construct a high-level Grid application environment where end-users can:

- deploy and use any legacy code program as a Grid service when creating Grid applications
- easily and conveniently create complex Grid applications by accessing a Grid portal.

GEMLCA is a Grid architecture developed by the research team at the Centre for Parallel Computing, University of Westminster, and aims to make legacy code programs available as Grid services with user-friendly interfaces and without re-engineering the original code. Using GEMLCA, legacy code written in any language (Fortran, C, Java etc) can easily be deployed as an OGSA Grid service without any user effort. GEMLCA has also been integrated with the P-GRADE

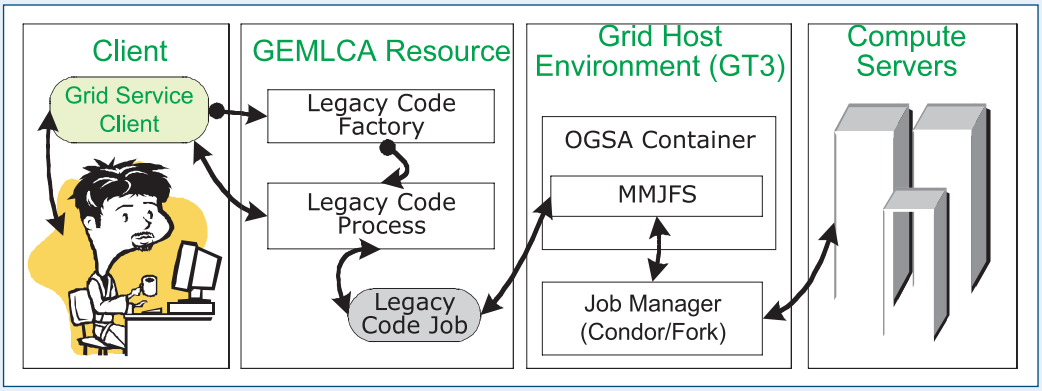


Figure 1: GEMLCA conceptual architecture.

Grid portal and workflow solutions (developed by SZTAKI), allowing the end-user to construct and execute complex Grid workflows, using a user-friendly Web interface, on which legacy components of the workflow show up as Grid services.

GEMLCA is a general solution for the deployment of existing legacy code applications as Grid services without code re-engineering. The only access point for a client to the GEMLCA architecture is the front-end layer composed of a number of Grid services. These services offer interfaces in order to deploy, query, submit, check the status of, and get the results back from legacy computational jobs. As with any other Grid Service, the front-end layer is described in Web Services Description Language (WSDL) and can be used by any Grid service client to bind and make use of its functionality through Simple Object Access Protocol (SOAP). GEMLCA clients can be easily created using either the provided universal GEMLCA Stubs for Java or the WSDL file. The conceptual architecture of GEMLCA is shown in Figure 1.

Our goal was to enable end-users to apply legacy code as OGSA Grid services in workflows and to provide a user-friendly interface through which to do this. In order to achieve our goal GEMLCA has been integrated with the P-GRADE Grid portal solution. The workflow editor of P-GRADE portal enables the connection of component jobs into complex workflows through an easy-to-use graphical environment. The Workflow Manager of P-GRADE portal takes care of executing such workflows on various Grid systems in a user-transparent way. The integration of GEMLCA and the P-GRADE portal allows the user to construct and execute workflows from legacy codes deployed as OGSA Grid services.

In order to demonstrate the capabilities of the solution, a complex workflow simulating and analysing urban car traffic was created and executed on five different sites of the UK OGSA Testbed. Simulators run in parallel at SZTAKI, CCLRC Daresbury Laboratory, and Westminster, Portsmouth and Reading Universities. The simulators were fed by a road-network generator and the results were analysed by a program also running at Westminster. The applications were legacy C codes transferred into Grid services using GEMLCA. The workflow and visualization of execution is shown in Figure 2.

Future work will include the support of other service-oriented Grid architectures like WSRF and pure Web services. It is also envisaged that plug-ins for application-specific visualizers to the P-GRADE portal will be developed.

Links:
GEMLCA: http://www.cpc.wmin.ac.uk/ogsitestbed/GEMLCA/

P-GRADE: portalhttp://www.lpds.sztaki.hu/pgportal/

Please contact:
Péter Kacsuk, SZTAKI, Hungary
Tel: +36 1 279 6064
E-mail: kacsuk@sztaki.hu

Tamás Kiss, Cavendish School of Computer Science, UK
Tel: + 44 20 7911 5000, ext 3542
E-mail: kisst@wmin.ac.uk

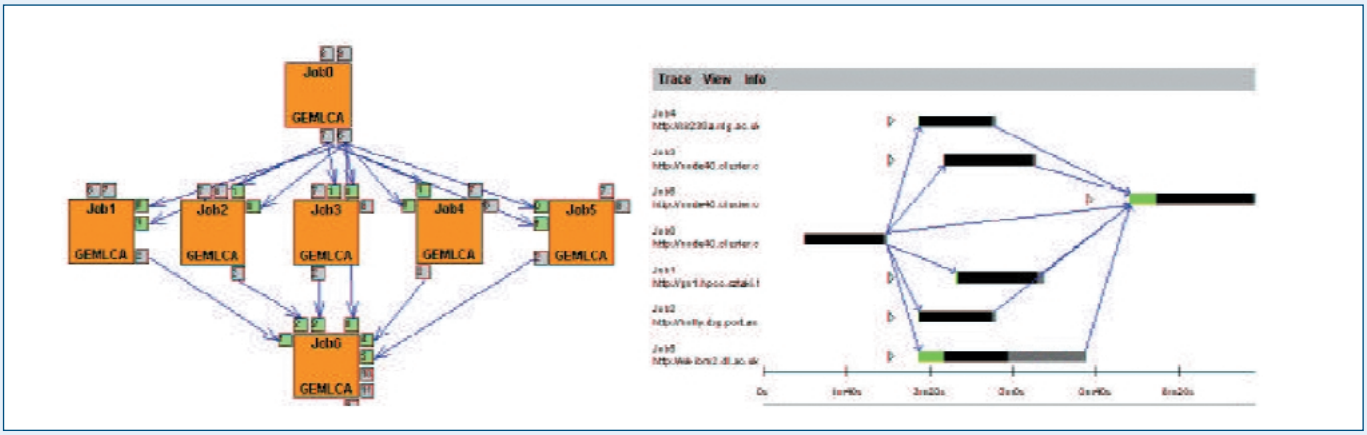


Figure 2: Complex workflow and execution graph for analysing road traffic.

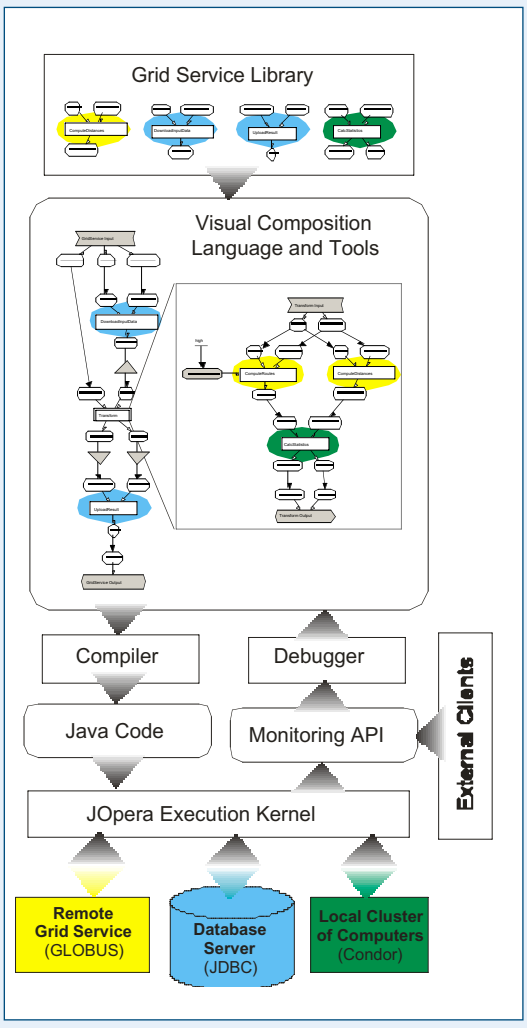
JOpera: Visual Composition of Grid Services

by Cesare Pautasso

The recent shift to service-based Grids enables the use of service composition tools for rapidly building and efficiently running distributed computations. At ETH Zurich, we have developed JOpera, a visual composition language and a set of integrated software development tools for composing Grid services. The system can be freely downloaded and has been successfully tested with data-intensive bioinformatics applications as well as large-scale, Monte-Carlo network simulations.

Early Grid middleware and tools emphasized low-level resource management aspects. Today, Grid computing has adopted the notion of 'services' as a basic building block for large-scale scientific computations. The focus has moved from CPU sharing to defining computational resources and functionality as composable services. Thanks to standards such as Web services, an increasingly large number of basic Grid services are being published. They encapsulate both algorithms and data sources, hiding the complexities and peculiarities of the underlying platforms. With services such as basic building blocks, it becomes possible to rapidly build Grid applications by composing these services at a high level of abstraction.

At the Information and Communication Systems Research Group at ETH Zurich we have developed JOpera, a system for rapid service composition based on a visual composition language. With JOpera, the data transfers between services (data flow), their order of invocation and the necessary failure handling behaviour (control flow) can be specified with a simple, graph-based, visual syntax. As opposed to an XML-based approach (eg BPML, BPEL4WS, XL), a visual language can greatly en-hance the understandability of Grid computations, as they can be specified by literally drawing them. Furthermore, at run-time, they can also be debugged and monitored using the same visual syntax. In addition to nesting, recursion and reflection, the JOpera Visual Composition Language offers several iteration constructs specifically targeting Grid computing. For example, large computational jobs are typically partitioned into smaller units that can then be executed in a pipelined, parallel or sequential fashion. JOpera includes simple primitives for these operations and allows the programmer to precisely control what happens in the case of failures. JOpera can also suggest the most optimal partitioning strategy based on the available resources.



Architecture of the JOpera system.

JOpera offers an open, flexible platform for Grid service composition. It maintains a library of re-usable components, which can be dragged, dropped and connected into a data and control-flow diagram. JOpera makes very few assumptions on the properties of these components, so that the user can freely choose to compose the most appropriate kind of services in terms of performance, reliability, security and convenience. Such components may represent the invocation of basic, remote Grid services, but also, for example, job submissions to external resource management and scheduling systems or the execution of local applications under a variety of operating systems. Additionally, composite Grid services can be re-used in two different ways. On the one hand, Grid computations can be decomposed hierarchically into different modules, which can be invoked and re-used independently. On the other hand, composite Grid services can define re-usable patterns of interaction between services, which can then be customized and tailored to specific applications. To do so, service interfaces are bound to actual Grid sites providing compatible services at deployment or invocation time.

For efficient execution, the JOpera Visual Composition Language is compiled into Java code. To ensure the required level of scalability, reliability and flexibility, the resulting code is dynamically plugged into a runtime container: the JOpera kernel. The flexible architecture of the JOpera kernel can be deployed in different configurations, eg standalone or embedded into other

systems, such as development environments or application servers. An open API is a key feature for providing access to composite Grid services in a variety of ways. In the simplest case, once Grid services are composed, the resulting computation can also be published as a Grid service. To do so, the JOpera kernel manages the execution of the composite service behind a Grid portal. Similarly, JOpera can be embedded into end-user applications, thereby supporting the development of application-specific features built by composing existing Grid services. Additionally, we have developed generic, visual monitoring tools for keeping track of the progress of a Grid computation. In order to handle large workloads, the JOpera kernel can scale across a cluster of computers. We are currently investigating how to apply autonomic techniques in order to extend the system with a controller in charge of determining the optimal configuration of the system based on the current workload.

Link:
The latest version of JOpera can be freely downloaded from http://www.iks.ethz.ch/jopera

Please contact:
Cesare Pautasso, ETH Zurich/SARIT, Swizterland
Tel: +41 1 632 0879
E-mail: pautasso@inf.ethz.ch

Grid-enabled Weka: A Toolkit for Machine Learning on the Grid

by Rinat Khoussainov, Xin Zuo and Nicholas Kushmerick

In the end of the day, Grids are about Grid-enabled applications. While a number of general-purpose libraries for scientific computing have been adapted for Grids, many more opportunities remain in various specialist areas. We describe here an ongoing work on Grid-enabled Weka, a widely used toolkit for machine learning and data mining.

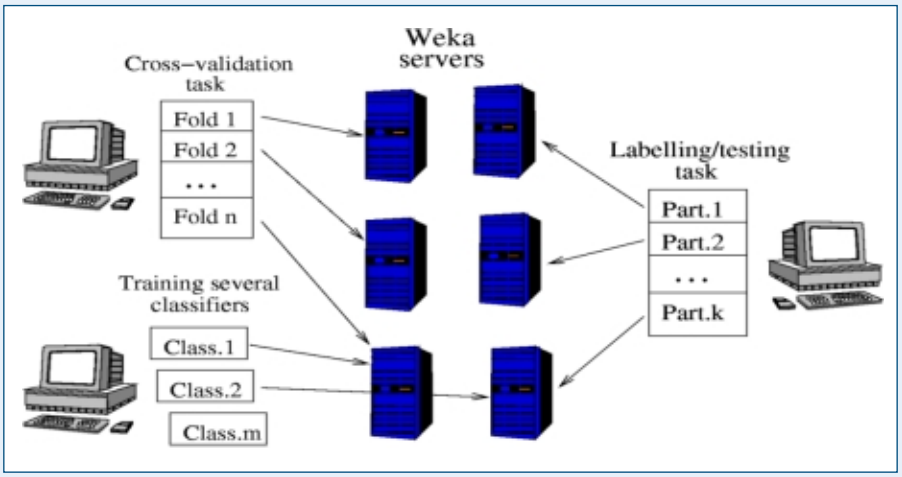
Weka is a widely used toolkit for machine learning and data mining originally developed at the University of Waikato in New Zealand. It is a large collection of state-of-the-art machine learning algorithms written in Java. Weka contains tools for classification, regression, clustering, association rules, visualisation, and data pre-processing. Weka is open source software under the GNU GPL. It is easily extensible, which allows researchers to contribute new learning algorithms to Weka, keeping it up-to-date with the latest developments in the field. As a result, Weka has become very popular with academic and industrial researchers, and is also widely used for teaching purposes.

The main focus of Weka is on classifier algorithms. Simply put, a classifier maps a set of data instances onto a finite set of classes. Each data instance is described by its attribute values. For example, predicting whether it is going to rain based on observations of sky, air temperature, humidity, and wind can be viewed as a classification task. Each data instance includes values of the observation attributes, eg (sunny, warm, humid, strong), and the available classes are {rain, dry}. The goal of classifier learning (or training) process is to derive a classifier from a set of labelled data (ie a set of data instances together with their correct labels). The idea is that a classifier learned on a labelled data set can then be used to predict class labels for future (unlabelled) data instances.

Learning a classifier and using it to label data can be time consuming and require significant amounts of memory, especially for large data sets. Unfortunately, parallelising and distributing classifier learning is a difficult research problem on its own. Nonetheless, there are a number of simpler functions and usage scenarios in Weka that can still benefit from distributing the work on a Grid. The most obvious are labelling, testing, and cross-validation functions.

Labelling involves applying a previously learned classifier to an unlabelled data set to predict instance labels. Testing takes a labelled data set, temporarily removes class labels, applies the classifier, and then analyses the quality of the classification algorithm by comparing the actual and the predicted labels. Finally, for n-fold cross-validation a labelled data set is partitioned into n folds, and n training and testing iterations are performed. On each iteration, one fold is used as a test set, and the rest of the data is used as a training set. A classifier is learned on the training set and then validated on the test data.

In Grid-enabled Weka, which is being developed in University College Dublin, execution of these tasks can be distributed across several computers in an ad-hoc Grid. The labelling function is distributed by partitioning the data set, labelling several partitions in parallel on different available machines, and merging the results into a single labelled data set. The testing function is distributed in a similar way, with test statistics being computed in parallel on several machines for different subsets of the test data. Distributing cross-validation is also straightforward: individual iterations for different folds are executed on different machines. The quality of a classifier often depends on various algorithm parameters. The same classifier may need to be trained with different parameters to obtain better results. In our system, the user can run a training task on a remote machine. This allows the same user to train several classifiers in parallel by launching multiple Weka tasks from the user's computer.



Grid-enabled Weka, usage scenarios.

The two main components of our system are Weka server and Weka client. The server is based on the original Weka. Each machine participating in a Weka Grid runs the server component. The Weka client is responsible for accepting a learning task and input data from a user and distributing the work on the Grid. The client implements the necessary functionality for load balancing and fault monitoring/recovery. It also allows the user to specify resource constraints for a given task and takes these into account when allocating the jobs to servers. The server translates client requests into calls to the corresponding Weka functions. It also provides additional functions like data set recovery from local storage after a crash. The same server can be used by several different clients, which allows users to share resources of the same machine.

The system uses a custom interface for communication between clients and servers utilising native Java object serialisation for data exchange. The obvious next step is to convert it to an OGSA-style service using, for example, a Globus API toolkit. An important advantage of the current implementation is that in a trusted and centrally controlled environment (eg a local computer lab) it allows for utilising idling computing resources with minimal set up, configuration, and maintenance efforts. This is especially convenient for machine learning practitioners who may not be proficient in parallel computing or Grid technologies. The Grid-enabled Weka is currently replacing the original Weka in Elie, a machine learning-based application for information extraction developed in University College Dublin.

Links:
Weka:
http://www.cs.waikato.ac.nz/~ml/weka/

Elie: http://www.cs.ucd.ie/staff/nick/home/research/download/finn-ecml04.pdf

Please contact:
Rinat Khoussainov, Nicholas Kushmerick
University College Dublin, Ireland
E-mail: rinat@ucd.ie, nick@ucd.ie

The eMinerals Minigrid: An Infrastructure to Support Molecular Simulation Scientists

by Martin Dove

The eMinerals project is one of the UK Natural Environment Research Council (NERC) e-science testbed projects (formal name 'Environment from the Molecular Level'). The scientific aim is to use Grid technologies to push forward the capabilities of molecular-scale simulations for the study of environmental processes.

Examples include issues such as nuclear waste encapsulation, adsorption of pollutants on the surfaces of soil particles, and the interaction of mineral surfaces and fluids. The simulations we perform range in their complexity. In some cases we need to simulate systems containing millions of atoms, and for this we represent the forces between atoms by simply parameterised models. In other cases we need to have more accurate representations of these forces, and we use quantum mechanical methods, but with smaller numbers of atoms.

At its heart, the challenges faced in scaling up these calculations are computational, but with larger simulations comes an increased need manage data in new ways. In addition to the fact that larger simulations generate larger data files, we also have the problem that with higher throughput of calculations comes the need for intelligent file management within an individual study. The eMinerals minigrid has been developed as an integrated compute and data grid with the aim to equip the scientists to work with more complex simulation studies than before.

The eMinerals minigrid is built around a heterogeneous set of computing resources, quite deliberately so because different types of molecular simulations have very different requirements. The compute resources include 4 clusters of PCs, an IBM p-series parallel computer,

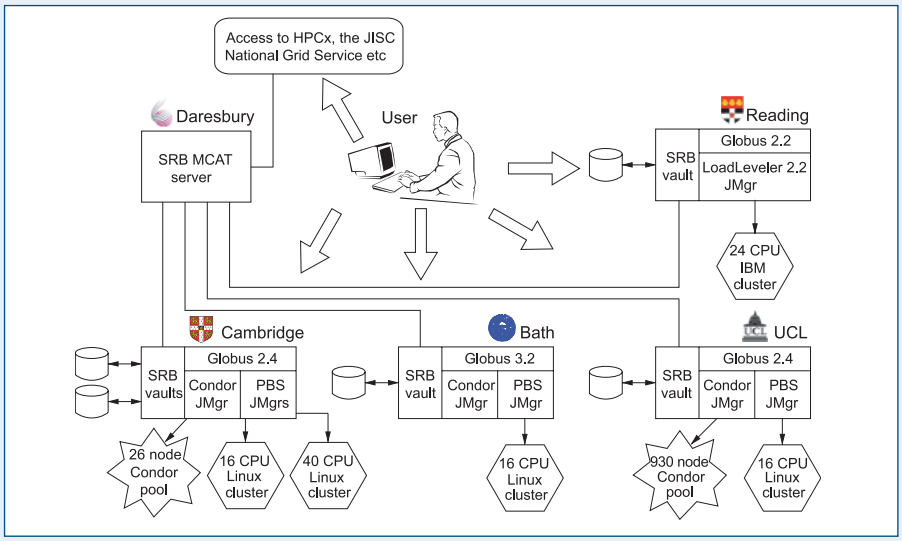


Figure 1: The structure of the eMinerals minigrid, showing both the hardware and middleware configurations for the integrated compute and data structures and their institutional configurations.

a Condor pool containing a range of machines but primarily designed to allow test jobs with high memory requirements, and a large (930 machine) Condor pool containing the teaching PCs of University College London. The latter is the largest Condor installation in the UK, and is designed to allow high-throughput calculations that do not have high memory requirements. Access to these machines is through Globus, with each member of the team having a UK digital certificate to ensure authorisation and authentication. Moreover, we expect the scientists to submit jobs using the Globus toolkit commands rather than via direct logins to the resources. The only logins are permitted to allow code developers to compile and test.

One immediate problem with this regime is that of file transfer. Both Globus and Condor present problems for users in this regard. Our solution has been to directly incorporate the Storage Resource Broker (SRB) developed by the San Diego Supercomputer Centre. The SRB provides a means to distribute files across a number of file systems (called 'SRB vaults') but to be seen by the user as a single file system. The physical location of files is transparent to the user, and is seen only as a file attribute. We have set up 5 SRB vaults across the eMinerals minigrid. The integrated compute and data grid that constitutes the eMinerals minigrid is shown in Figure 1, with the data component shown in more detail in Figure 2.

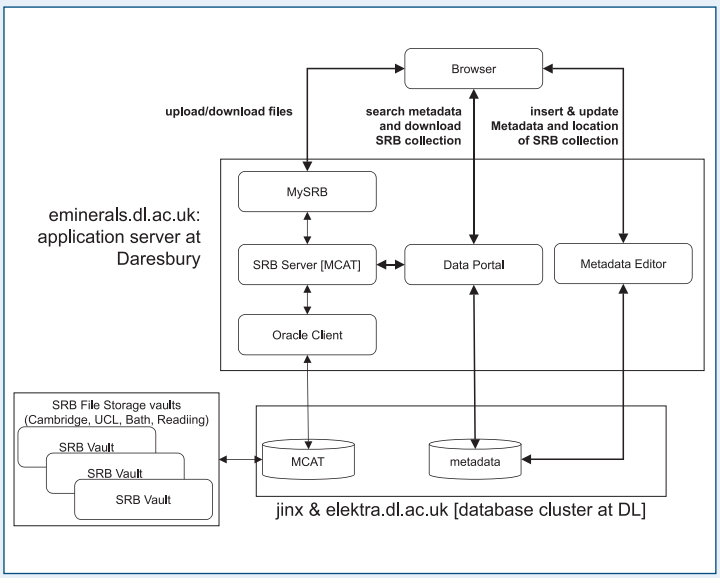


Figure 2: The structure of the data component of the eMinerals minigrid, showing how the various components map onto the infrastructure of the application server and database cluster.

The integration of the SRB within the eMinerals minigrid facilitates the following workflow. When a user wants to run a job, the first task is to upload the relevant files to the SRB. The user then executes a three-stage job using Condor-G, which provides a Condor-style wrapping for the Globus commands. The first stage is for these files to be downloaded from the SRB onto the compute resources being used. The second stage is to run the job. The third stage is to put all output files into the same directory within the SRB for easy access to the user. These three stages have been wrapped within a single script written by one of the project members, although we are working on a compute portal to make job management easier for the scientists.

The final stage in the process is longer-term management of the data. The Daresbury project partners have developed a data portal that has a direct interface with the SRB. Data that require long-term storage are uploaded to the data portal from the SRB, and provided with metadata that enables use to be made of the data by the scientist or collaborators at a later date. In order to ensure interoperability of data, we are making use of the Chemical Markup Language, one of the established XML standard languages.

This report represents work in progress. Components of this work, particularly the eMinerals minigrid as an integrated compute and data grid, are now beyond proof-of-concept and are in production mode. Immediate tasks are to provide the tools such as the compute portal to give the scientists more help in their use of the full minigrid. The eMinerals team consists of scientists, applications code developers and computer scientists/grid specialists from the Universities of Cambridge, Bath, and Reading, University College London and Birkbeck College, and the CCLRC Daresbury Laboratory.

Link:
http://www.eminerals.org

Please contact:
Martin Dove, University of Cambridge, UK
E-mail: martin@esc.cam.ac.uk

Virtual Vascular Surgery on the Grid

by Peter Sloot and Alfons Hoekstra

Medical simulations and visualizations typically require computational power not usually available in a hospital. The University of Amsterdam recently demonstrated Virtual Vascular Surgery (a virtual bypass operation), where large-scale simulation and visualization capabilities were offered as services on the Grid.

Arteriosclerosis is a widespread disease that manifests particularly in developed countries. Treatment often involves surgery, such as the placement of bypasses that lead the blood around clogged arteries to restore normal blood flow. A surgeon plans these interventions on the basis of 3D images obtained from MRI or CT scans. Apart from considerations such as accessibility, the attainable improvement in the blood flow will determine what type of treatment is appropriate. Improvements in the support for planning these procedures are expected to improve their success rate. We have developed a prototype grid-based system for virtual vascular surgery, which may be used during pre-operative planning or as a valuable tool in the training of novice vascular surgeons. The prototype uses advanced distributed simulations and visualizations to support vascular surgeons in making pre-operative decisions.

The Virtual Radiology Explorer

The Virtual Radiology Explorer (VRE), developed at the University of Amsterdam, is a grid-based problem-solving environment for virtual vascular surgery. The VRE contains an efficient mesoscopic computational haemodynamics solver for blood-flow simulations based on parallel cellular automata. To convert the medical scans into computational meshes, raw scanner data is first segmented so that only the arterial structures of interest remain in the data set. The segmented data is then converted into a computational mesh. The patient's blood flow is simulated using grid resources. The VRE system uses a desktop virtual reality environment where the patient's data, obtained from a scanner, is visualized as a 3D stereoscopic image together with the graphical interpretation of the simulation results. A user can then manipulate the 3D images of the arteries. This would be the virtual surgical procedure, eg the addition of a bypass. Blood flow in this new geometry is also computed, and the user can then assess the effectiveness of the proposed treatment, and might try other alternatives to optimise the procedure.

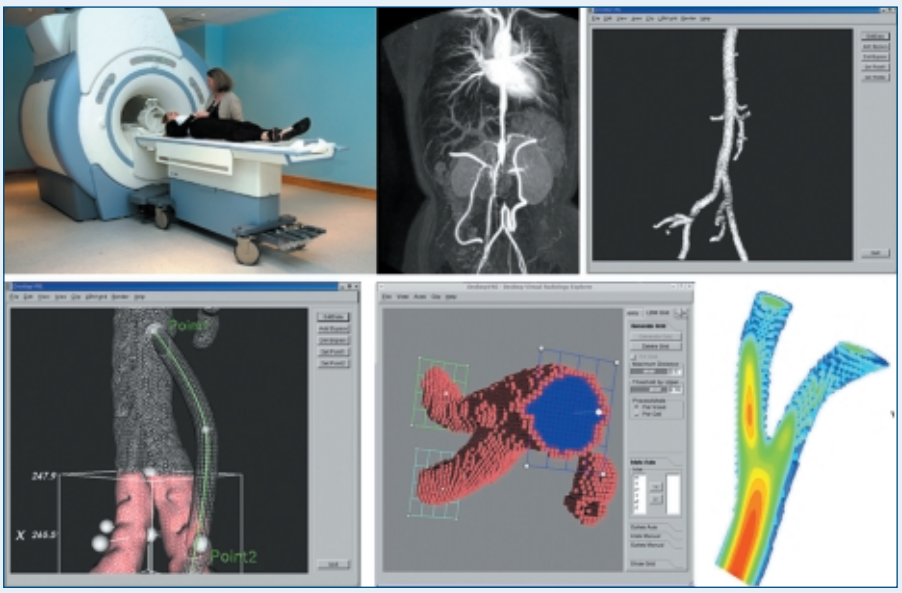


Figure 1: Distributed image-based blood-flow simulation and visualization on the Grid. From top left to bottom right: A patient is scanned in Leiden, The Netherlands; this results in a raw image stored in a Storage Element on the Grid (eg Poznan, Poland); this is segmented, filtered and cropped, using a Grid service; a bypass is added; a computational mesh is generated (on the local machine) and given to the parallel flow solver, running on Compute Elements in the Grid (Amsterdam and Spain); the resulting flow fields are displayed on the local machine using visualization services offered by the Grid Visualization Kernel (Linz, Austria).

The medical scanners, the visualization environment and the computational resources required for the flow computations and visualizations are usually located in disparate geographical locations and distinct administrative domains. The need for transparent access to these resources, high efficiency and strict data security triggered the development of our advanced grid-based simulation and interactive visualization for virtual vascular surgery. The middleware to support this was developed in the CrossGrid project.

CrossGrid

The CrossGrid project is oriented towards compute- and data-intensive applications characterized by the interaction with a person in the loop. The CrossGrid pan-European distributed testbed shares resources across sixteen European sites. One key component of Crossgrid is the Migrating Desktop (MD) grid portal. The MD produces a transparent user work environment, permitting the user to access grid resources from remote computers. Users can run applications, manage data files, and store personal settings independently of the location or the terminal type. We have incorporated our VRE system into the Grid via the MD grid portal. We achieved secured grid access,

node discovery and registration, grid data transfer, application initialization, medical data segmentation, segmented data visualization, computational mesh creation, job submission, distributed blood-flow visualization, and bypass creation. VRE runs on a local machine but is launched and initialized through the MD. The input takes the form of segmented or non-segmented medical data produced at the Leiden Medical Centre (LUMC); the CrossGrid testbed provides access to this data from a medical image repository acting as a Grid Storage Element in Leiden.

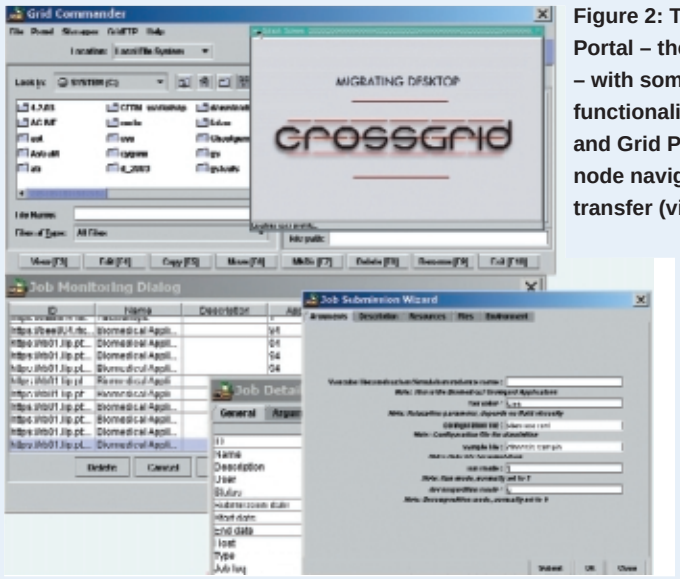


Figure 2: The CrossGrid Grid Portal – the 'migrating desktop' – with some of its functionalities, eg grid log-in and Grid Proxy creation, virtual node navigation and Grid data transfer (via the Grid Commander), submission of blood flow simulations to the CrossGrid testbed (via the Job Submission Wizard) and monitoring of jobs running in the testbed (via the Job Monitoring Dialog).

A Virtual Bypass Operation on the Grid

We have recently demonstrated the following scenario: The abdominal aorta area of a patient is scanned, and the resulting image is stored in a Radiology Information System repository. Later, a physician (user) logs into the CrossGrid Portal using his Grid certificate and private key. The user checks if there are segmented or non-segmented medical data ready for analysis in one of the virtual nodes, and securely transfers a few to his local machine. The user then starts the VRE from within the MD, loads the segmented medical data, selects a region of interest, crops the image, adds a bypass, and creates a computational mesh. The user selects the Biomedical Application icon within the MD (with parameters and files being taken from the user's profile), and submits the job to the CrossGrid testbed, to the nearest or most adequate Computing Element in the Grid. The user may then check job submission or progress via the MD. After the job has been completed, the calculated velocities, pressure, and shear stress are transferred to the local Storage Element or to the Grid Visualization Kernel to be rendered and reviewed by the user. We will give a live demonstration of this process during the upcoming European Grid Conference in Amsterdam, February 2005.

Overall, the successful deployment of grids requires the re-design of algorithms to support loosely coupled computer resources. This should be driven by relevant and challenging applications. With the experience gained within CrossGrid from the virtual vascular surgery application (and three other applications), European Grid Technology leads the way in new Grid developments worldwide.

Links:
http://www.science.uva.nl/research/scs
http://www.crossgrid.org/

Please contact:
Peter Sloot, University of Amsterdam,
The Netherlands
Tel: +31 20 525 7462
E-Mail: sloot@science.uva.nl

Photorealistic Visualisation with Grid-Based Technology

by Mika Pennanen and Markus Ylikerälä

The rapid development and low price of personal computers make them an interesting choice for conveying ideas through visualisation. Although real-time 3D graphics are adequate for many applications, a single computer is insufficient if photorealistic images are required, not only with today's technology, but even in the near future. At VTT Information Technology, grid-based distributed computing has been utilised to take advantage of the processing power of off-the-shelf office computers that are idle most of the time. As a result, using the rich resources of this existing distributed and networked environment, photorealistic images were produced faster and more cheaply.

Over the years, bandwidth has increased and become cheaper, computers have become more powerful, and the Internet has evolved into an accepted way of connecting computers and computing devices. As resources are rarely utilised to the fullest extent, companies have begun to promote the concept of grid computing, which enables the integra-

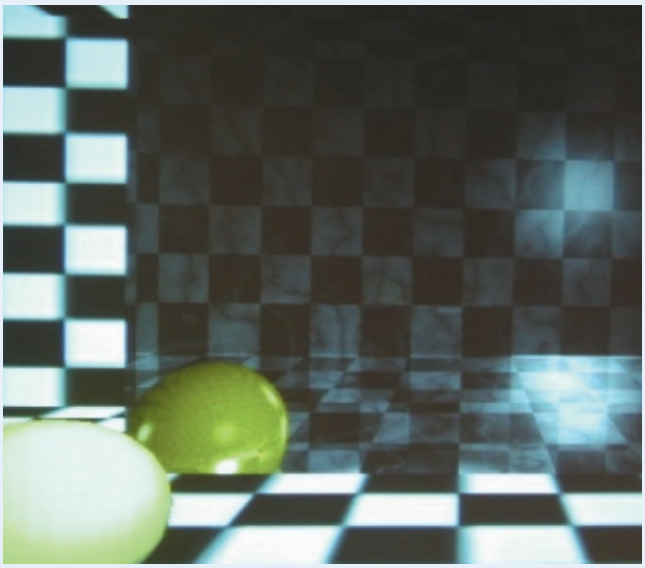


Figure 1: Ray-tracing from an arbitrary point of view in an OpenGL application.

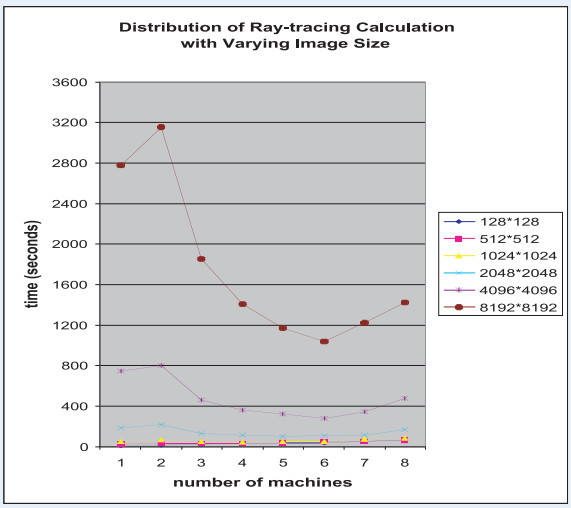


Figure 2: Distribution of ray-tracing calculations.

tion of computing power from anywhere in the world as easily as users now log on to the Internet, or even use electricity. This also opens up the possibility, for example, of increasing the rendering capability of low-capacity devices via the Internet, on demand. Imagine being able to plug your computer into the wall and tapping into as much processing power as you need. The promise of such utility computing is very attractive: it means obtaining the right amount of computing power at the right time, employing the latest hardware, software and services, and paying only for the computing power used. The savings could be substantial.

VTT Information Technology has participated in several national and international research projects related to grid technology. During these projects we utilised two of our research areas: utility computing used for calculation purposes, and visualisation and virtual reality environments. The objective was to use the grid for visualisation purposes to convey ideas to the interest group, and research has focused on distributed calculations for visualisation purposes. The Globus Toolkit has been used, and a grid called Grix has been developed based on platform-neutral architecture. During our research on distributed calculations for visualisation purposes, we utilised the processing power of off-the-shelf office computers that are idle most of the time. With the rich resources of this existing distributed and networked environment, photorealistic images were produced more quickly and more cheaply. To obtain photorealistic images, techniques such as ray-tracing can be used. Ray-tracing simulates the travel of rays of light from the observer's viewing angle to the light sources. It should be noted that the simulation is made contrary to physical laws in order to reduce the necessary number of calculations. This is based on the fact that the observer can sense only some of the rays. The program currently in use is POV-Ray, a well-known ray-tracing program. Unfortunately a single computer – not only current machines but those of the near future – has insufficient power to render large or complex photorealistic images. A grid is therefore used in order to gain more computing power. An example of our research on ray-tracing with our OpenGL application is shown in Figure 1. With this application the user can move within a room that has chessboard surfaces, a yellow ball and a canvas. The observer takes a snapshot from an arbitrary viewpoint, and a photorealistic image is rendered with a grid and attached to the canvas.

To measure the effect of the grid with varying numbers of computers, we measured the time it took to render a photorealistic image. The image was from a model of a ski lodge containing two light sources and approximately 55,000 polygons. The complexity of this rendering task varied with the image size; Figure 2 shows the results of these tasks. Depending on the image size, the processing time with one computer (PIII processor) varied from nine seconds to 47 minutes. With multiple computers, one computer acts as the controller, while the rest perform the actual ray-tracing calculations. As a result, the performance of two computers is worse than one alone, since control data merely increases overhead. Using the increased power of a moderate number of computers does produce significant improvements, particularly with large images. However, for large numbers of computers, the control data overhead increases such that the results again deteriorate. Significant benefits can therefore be obtained even with a small number of computers. Performance is not improved by continuously increasing the number of computers.

Our future work will concentrate on extending these solutions and demonstrating them on Grix. Furthermore, it is planned to extend and implement the business scenario with payable services for device users.

Please contact:
Mika Pennanen, VTT Information Technology, Finland
Tel: +358 9 456 5623
E-mail: mika.pennanen@vtt.fi

Markus Ylikerälä, VTT Information Technology, Finland
Tel: +358 9 456 6082
E-mail: markus.ylikerala@vtt.fi

Solving Optimization Problems with Grid-Enabled Technologies

by Enrique Alba and Antonio J. Nebro

Research in the 'Networking and Emerging Optimization Research Line' at the University of Málaga aims at solving optimization problems through the utilization of grid–enabled technologies and large computer networks.

Grid computing is a recent discipline related to the utilization of large-scale distributed systems for a given purpose by taking advantage of the rich infrastructure provided by the Internet. To understand the basic idea that motivates our research, consider a grid of computers as a huge virtual multicomputer ready for processing, storage and communication. Since a grid can be made up of a set of geographically separate networks, enormous computer power is available for solving complex problems that are limited in CPU and that require long delays if solved in modern computer LANs.

Complex problems that can only be solved in non-polynomial time arise in most fields of research and are becoming common in many areas of our lives: telecommunications, economy, bioinformatics, industrial environments,and so on. For such a wide spectrum of problems, heuristics come to the rescue, since exact techniques are unable to locate any kind of solution. This area is known as Networking and Emerging Optimization (NEO), and the GISUM group at the University of Málaga (Spain) is working on just these techniques.

In short, using Grid technology, our aim is to solve optimization problems that are otherwise out of the scope of researchers dealing with parallel algorithms. Five types of applications are usually identified as related to Grid computing: distributed supercomputing, high-throughput computing, on-demand computing, data-intensive computing and collaborative computing. At this stage of our research we are primary involved with the two first topics, ie distributed and high-throughput computing. However, we plan to enter the other three domains by developing an open optimization service for the Internet, solving data-mining problems and facing software agent applications respectively. Figure 1 is an example of the goal of the present work: computing an exact Pareto front.

$$
\begin{aligned}
\text{Minimize } F &= (f_1(x,y), f_2(x,y)) \\
f_1(x,y) &= 4x^2 + 4y^2 \\
f_2(x,y) &= (x-5)^2 + (y-5)^2 \\
0 &\geq (x-5)^2 + y^2 - 25 \\
0 &\geq -(x-8)^2 - (y+3)^2 + 7.7 \\
0 \leq\ & x \leq 5 \\
0 \leq\ & y \leq 3
\end{aligned}
$$

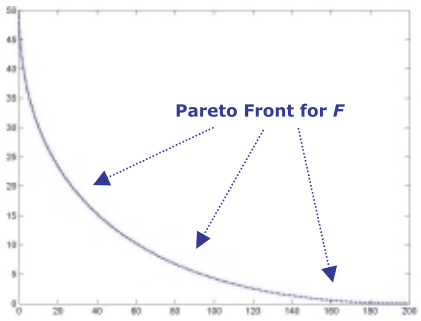


Figure 1: A simple example of a multi-objective optimization problem with the constraints (left), and its Pareto front (right).

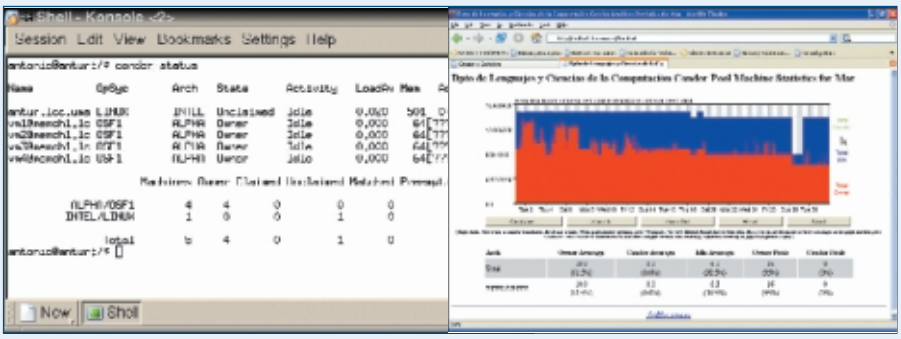


Figure 2: Tracking and managing the optimization algorithms in the grid with Condor.

Multi-criterion optimization in which several (non-dominated) optimum solutions must be found is a promising research field. At present, this field lacks algorithms that could ensure the computation of the exact Pareto front for a general problem. We avoid this inconvenience by using an enumerative-like search that computes all the non-dominated solutions in a grid. Later, researchers can use these results to find efficient heuristics achieving similar (optimal) results. We are in the first stage of research, and have solved problems at Málaga using a modest grid of around 110 processors. As simple as it sounds, finding the optimum Pareto front is extremely hard, even for small problems, and with most algorithms it is not guaranteed that the optimum front will be located for arbitrary problems as we do with our grid exact algorithm.

We are evaluating the performance of several grid-enabled technologies for our applications. Concretely, we have tested Globus, Condor, Legion, and Sun Grid Engine. Of these, the first two seem the most suitable for optimization in the grid. The simplicity and powerful process management of Condor were greatly appreciated in setting up a grid from scratch within a few weeks. Globus is in this sense more complex to use,

because users must deal with certificates, installation details and production tools that require a time-consuming learning curve. However, Globus actually eases the next step of our research in connecting to foreign grids. In addition, we have tested the network usage and execution times of heuristic algorithms and tools like MPI on Globus with very satisfactory results, and this suggests that we could achieve fruitful research with Globus in the future.

Most of this work has been undertaken within the on-going TRACER project (http://tracer.lcc.uma.es), which is developing advanced algorithms in close connection with the Internet. Two successful case studies in TRACER relating to Grid computing can be cited. Firstly, a grid algorithm was developed to compute the exact Pareto front of several hard multi-objective problems by using Condor in a grid with more than 110 computers (http://neo.lcc.uma.es/Software/ESaM). At present, we are extending the grid to other sites in Spain and Europe. Secondly, grid-aware multi-objective heuristics, initially run on Globus, were developed. We are currently testing new grid algorithms derived from well-known standard techniques like PAES or NSGA II.

We invite everyone to look at our results and contact us with any comments or suggestions for collaboration.

Link:
http://neo.lcc.uma.es

Please contact:
Enrique Alba, University of Málaga/
SpaRCIM, Spain
Tel: +34 952 132803
E-mail: eat@lcc.uma.es

Bioinformatics in the Semantic Grid

by Kai Kumpf and Martin Hofmann

One of the major challenges in Grid computing is the semantics-driven retrieval of Grid services and distributed data. DB-Annotator is an annotation tool for data in Grid services developed in a collaborative project between the bioinformatics department and the department for Web applications at the Fraunhofer Institute for Scientific Computing (SCAI). Fully annotated data in the Grid is particularly important in bioinformatics.

DB-Annotator was conceived for the Resource Description Framework (RDF) annotation of structured information sources such as relational data or XML-based service descriptions. Only semantically annotated Grid services (GS) provide a means of finding data or compute-services that are suitable for the task at hand. Furthermore, they make the building of workflows (coupled services within the grid) a realistic goal. DB-Annotator will support several levels of semantic annotation to enrich Grid services, ranging from the services themselves to the data within the Grid.

The project DB-Annotator (see Figure 2) was designed for universal annotation of data that can be represented in tabular form, ie data that can be qualified by unique keys.

The main thrust for the development came from the realization that most interesting data within biological databases resides within free-form text fields and is therefore not easily accesible for data mining. Relational databases (RDB) are usually the only choice for data retrieval within distributed organizations. Categorial description of the typed data residing in RDBs can be derived from the tables (entities) themselves, and from the columns (entity attributes) and the cell contents (instances of attributes). Still more fine-grained information comes from the notorious 'description' fields, but putting this implicit information to use for machine-reasoning is unrealistic at this point.

The goal was thus to provide easy navigation through existing ontologies and data-source-independent RDF annotation of RDB data via drag and drop (see Figure 2). This data categorization can

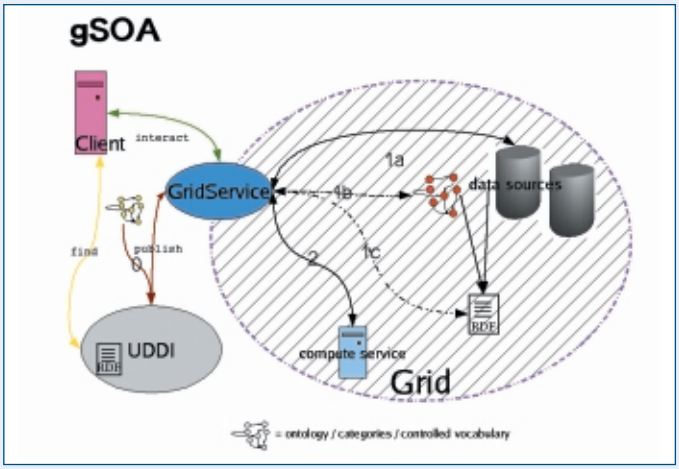


Figure 1: Grid Service- Oriented Architecture. The multiple levels of semantic annotation within a Grid-Service Oriented Architecture (gSOA). Path 0 corresponds to the semantic enhancement of the UDDI service registry via RDF. If no additional semantic glue is needed, direct access to either data- or compute-services is provided (1a, 2). When retrieving data, the semantic description can either stem from UDDI or, more finely -grained, from semantic annotation of structured (most often relational) data. 1b/c correspond to retrieval of annotated data; both are equivalent when there exists a central RDF annotation repository that stores n:m ontology-class -– (data) - relationships. Whereas 1a provides a complete view ofn the data via full-text search or keys, 1b/c allow querying by content.

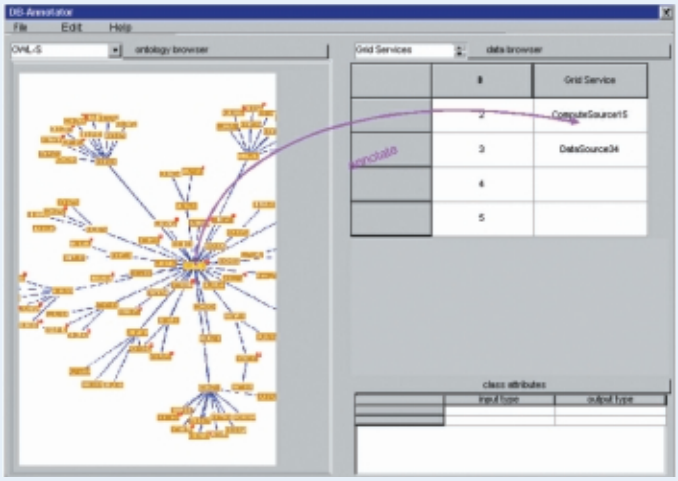


Figure 2: DB-Annotator was designed to allow RDF-annotation of structured data. Navigation through a chosen ontology is facilitated by a DAG view on the classes and their relationships (left). Any data that can be represented in tabular form and is accessible via unique keys can be annotated (upper right). Class attributes have to be filled in on annotation (lower right). In the case shown here the user has chosen a listing of available Grid Services that have to be semantically described. The major field of application for DB-Annotator however, is the annotation of tabular data from RDBs.

take place on three levels within a database (see Figure 3), not counting the database itself. A unique data key that can be linked to an ontological category will therefore appear as 'database:schema:table[:column:row]', the latter two key parts being optional depending on the depth of the annotation.

This concept of assigning ontology classes to data was inspired by existing annotation software for the Semantic Web. Once a data source has received ample annotation, data mining will be possible by employing standard methods on the categorized data. With annotated RDB, access to data services in the Grid will improve significantly, since querying by categories is enabled across several RDBs at once. Further, every result set will be enhanced with the respective category information, a data basis that could be used to support further automated annotation of newly inserted data.

Since DB-Annotator produces standard RDF (subject-predicate-object triples like 'entity_xyz is-a category_abc'), its output can also serve to enhance UDDI, a proposal made for the introduction of content-searchable Web services and a possible starting point for GS-workflow construction. Instead of the database unique keys, the RDF would contain qualifiers for the available Grid services. Eventually a semantically enhanced Grid Service-Oriented Architecture (gSOA – SOA being a term borrowed from the Web services world) would look like Figure 1. Full categorization of both GS and the exchanged data leaves no more room for ambiguities – a precondition for data exchange and manipulation within the Grid network of virtual organizations (VO).

Given the bulk of biochemical entities that go under various names and interact on different levels, the relevance of fully annotated data Grid services in bioinformatics, especially systems biology, is enormous.. Until now, ambiguities in synonyms and interactions had to be detected by human intelligence,– meaning full-fledged systems biology services won't be successfully automated without proper categorization. Since DB-Annotator draws on a common set of agreed ontologies (eg gene ontology) for the annotation of both data services and data sets or subsets, it can provide the semantic glue on all levels required.

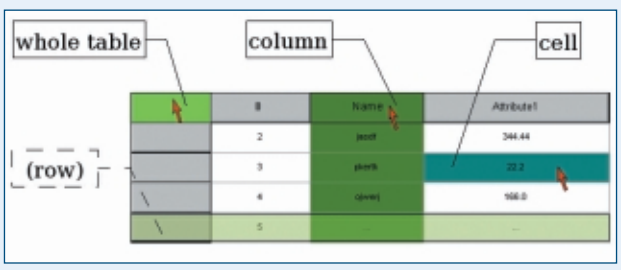


Figure 3: Selection modes in DB-Annotator. Data entities for annotation can either be the whole table, one column or one cell. The data key for annotation is constructed depending on scope of the selection. Only (multiple) annotation of a single cell corresponds to an extension of the data model. Selection of a row corresponds to selection of a record from the extended data model on the basis of table and ontology keys and as such need not be considered separately.

The project was launched in the department for bioinformatics at Fraunhofer SCAI. It is a collaborative project between the bioinformatics department and the department for Web applications. Preliminary work commenced late in 2003 with the aim of producing a prototype that would capture the basic functionality of the future system. The prototype, featuring 1:n mapping of a database to several ontologies and RDF flatfile storage, was completed along with a detailed requirement analysis this summer. A new structured development from scratch has now started and is expected to be finished early in 2005. Among the things that will be included in the new phase are n:m relationships between data and ontology classes and storage in a single, structured relational repository. The product is currently intended to provide support for semantic annotation of biological and chemical databases. The possibilities for Grid services annotation within UDDI are explored within the semantic grid research group under the auspices of the Global Grid Forum (GGF).

Links:
Globus Toolkit: http://www.globus.org

Semantic Web Services: http://swws.semanticweb.org

Global Grid Forum: http://www.gridforum.org

Gene Ontology: http://www.geneontology.org

Please contact:
Kai Kumpf, Institute for Algorithms and Scientific Computing (SCAI), Fraunhofer ICT Group, Germany
Tel: +49 2241 14 2257
E-mail: kai.kumpf@scai.fhg.de

CoreGRID: European Research Network on Foundations, Software Infrastructures and Applications for Large-Scale, Distributed Grid and Peer-to-Peer Technologies

by Thierry Priol

A Network of Excellence in the area of Grid and Peer-to-Peer technologies commenced on 1 September 2004, and will allow European researchers to carry out a joint program of activities at an unprecedented level with the aim of developing next-generation grid middleware.

The CoreGRID Network of Excellence (NoE) aims at strengthening and advancing scientific and technological excellence in the area of Grid and Peer-to-Peer technologies. It is funded by the European Commission through a grant of 8.2 million Euro for a duration of four years. To achieve its objective, the Network brings together a critical mass of well-established researchers (119 permanent researchers and 165 PhD students) from 42 institutions — several of which belong to ERCIM — who have constructed an ambitious joint program of activities. This joint program of activity is structured around six complementary research areas that have been selected on the basis of their strategic importance, research challenges and existing European expertise to develop next-generation grid middleware.

Virtual Institutes

To ensure progressive evolution and durable integration, the Network is operated as a European Research Laboratory — known as the CoreGRID Research Laboratory — having six institutes. Each of them is dedicated to the particular domain identified as of strategic importance to ensure a durable development and deployment of GRID infrastructures:

- Institute on Knowledge and Data Management
- Institute on Programming Model
- Institute on System Architecture
- Institute on Grid Information and Monitoring Services (WP5)
- Institute on Resource Management and Scheduling
- Institute on Problem Solving Environment, Tools and GRID Systems.

The Network is thus committed to setting up this Laboratory and making it internationally recognised and sustainable. It is funded by a European grant that is assigned to the CoreGRID NoE for a duration of four years to cover the integration costs, while the network partners cover the expense required to perform the research associated with the joint program of activities. Integration is achieved by the joint execution of research projects operated through research groups within the six institutes, the sharing of a common grid testbed for research assessment, the access to a common communication infrastructure to ensure collaboration and dissemination, and a coherent management framework to encourage mobility of senior and post-doctoral researchers and PhD students. In addition to these classical integration activities, the network has a proactive role to increase the awareness of trust and security technologies among the network participants.

To guarantee that the expertise and the knowledge gained by Network partici-

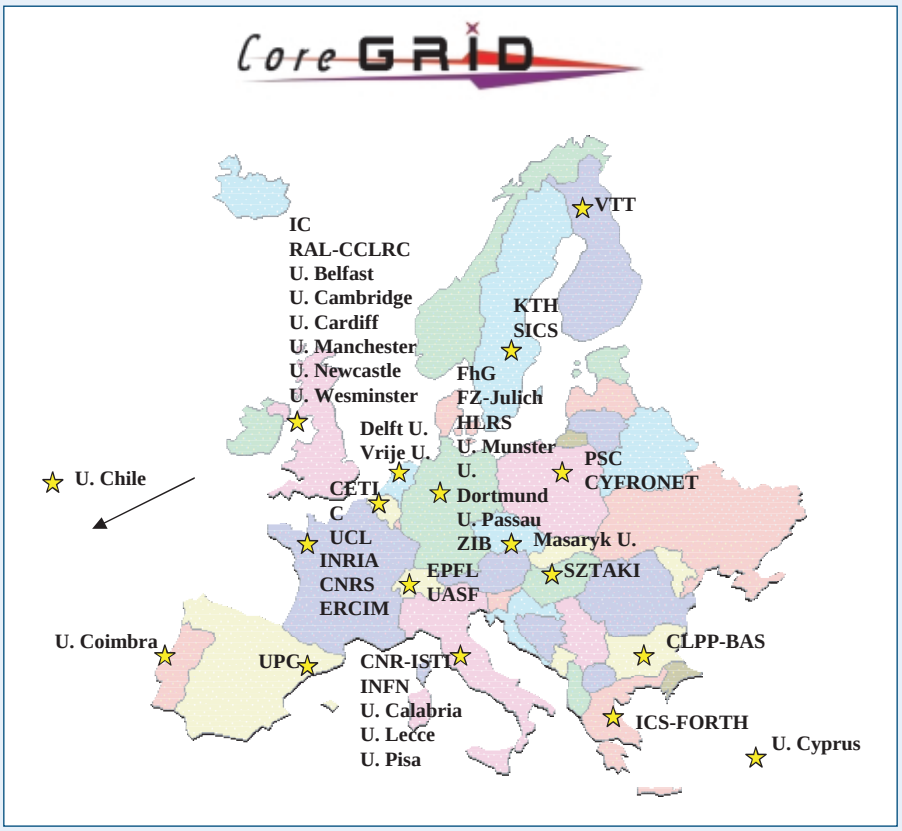


Figure 1: CoreGRID partners.

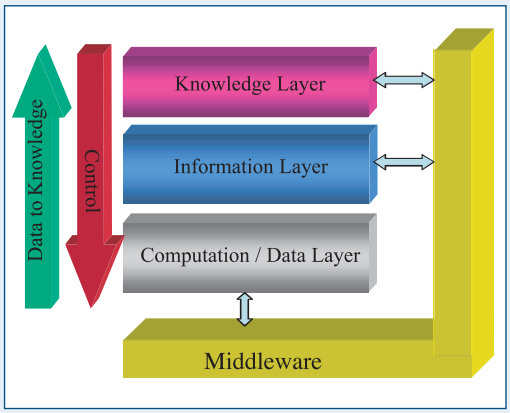


Figure 2: CoreGRID vision of the Next Generation Grid.

Figure 3: CoreGRID participants.

pants will be of mutual benefit to the European Grid community (both academic and industrial), CoreGRID will organize a set of activities to spread excellence outside the Network. A public Web site is an essential mechanism for the international Grid community to present the Network as a Research Laboratory. A set of CoreGRID publications can be accessed thanks to this Web site (technical reports, network brochure and newsletters). The Network will organize a set of thematic workshops and conferences on an annual basis. Training activities are organized to ensure that CoreGRID is a leading educational source in Europe in Grid and Peer-to-Peer technologies. A dedicated activity under the form of an Industrial Advisory Board will target mainly industry and commerce to ensure a strong interaction and involvement.

Finally, CoreGRID is committed to participating in collaboration activities that can benefit the other EC Grid projects. In particular, it will take on a leadership role in the organization of training activities and will collaborate on EU grid roadmaps, research inventories and the establishment of a repository of reference implementations and grid middleware.

Link:
http://www.coregrid.net

Please contact:
Bruno Le Dantec, ERCIM
Administrative & Financial Coordinator
Tel: +33 4 92 38 50 13
E-mail: Bruno.Le_Dantec@ercim.org

Thierry Priol, INRIA, France
Scientific Coordinator
Tel: +33 2 99 84 72 10
E-mail: Thierry.Priol@inria.fr

GridCoord – Co-ordinating Europe-wide Initiatives in Grid Research

by Marco Vanneschi

GridCoord is a Special Support Action (SSA) of the Sixth Framework Programme of the European Community. The goal is the co-ordination of European initiatives and research activities in Grid Computing in order to strengthen co-operation among the funding agencies planning future activities, to enhance collaboration between the research and user communities, and to develop visionary national and EU programmes and roadmaps enabling Europe to play a leadership role in Grid technologies and applications.

There are several Grid research initiatives, ongoing or planned, at the national and European Community level, aiming at the development of a rich set of advanced technologies, methodologies and applications. If the challenges that lie ahead are to be solved, enhanced coordination among the authorities funding these initiatives is needed in order to achieve critical mass, avoid duplication and reduce fragmentation of efforts.

From 2002 to 2006, EU funding for Grid research and deployment more than doubled, reaching 275 M in FP6. For the same period, an estimate of the funding for Grid research and deployment by a number of Member and Accession States (UK, France, Italy, The Netherlands, Germany, Hungary, Spain, Poland, Czech Republic, Sweden) is about 300-500 M .

The totality of these initiatives could provide the EU with the potential to play a leading world role in Grid technologies

and applications. National and EU collaborations have been established with other international players (in the US and Asian-Pacific area) and with international standards organisations. However, if Europe wishes to compete with leading global players, it would be sensible to attempt to better coordinate its various, fragmented efforts in order to achieve critical mass and the potential for a more visible impact at an international level. This requires:

- co-ordination among the funding authorities in order to maximise the effective use of the considerable existing and potential resources
- collaboration among individual researchers and the creation of European excellence and competence centres
- a visionary research agenda.

The first objective of GridCoord is thus to strengthen co-operation among the funding authorities in order to better co-ordinate the planning of future activities in the field of Grid research, an ERA Strategic objective. A second objective is to enhance already existing collaborations between researchers and users. Finally, it is our intention to contribute towards the development of visionary national and EU programmes and roadmaps, enabling Europe to play a leadership role in Grid technologies and applications.

Partners

The project brings together national research programme leaders and representatives from national funding authorities with the aim of ensuring national commitment to future joint activities. The following institutions are partners of GridCoord:

- Italy: University of Pisa, Department of Computer Science, and University of Genova, Department of Communication, Computer and System Science
- France: Institut National de Recherche en Informatique et Automatique (INRIA), Sophia Antipolis, and University of Nice, Sophia Antipolis
- UK: Engineering and Physical Sciences Research Council (EPSRC), and Queen's University Belfast, School of Computer Science
- Germany: University of Stuttgart, High Performance Computing Center (HLRS), and Konrad-Zuse-Zentrum für Informationstechnik (ZIB), Berlin
- The Netherlands: University of Amsterdam, Faculty of Science, Institute of Informatics
- Hungary: Computer and Automation Research Institute, Hungarian Academy of Sciences MTA SZTAKI
- Spain: Technical University of Madrid, School of Computer Science;
- Poland: Poznan Supercomputing and Networking Center
- Sweden: Swedish Agency for Innovation Systems (VINNOVA).

Objectives and Activities

To achieve the first objective — improved coordination among funding bodies, policy makers and leaders of Grid initiatives — the project will:

- develop a comprehensive compendium of national and EU initiatives including an analysis of strengths, weaknesses and gaps at EU level
- identify common interests and synergies capable of leading to cooperation between different groups, collaboration among funding authorities and eventually the planning of common initiatives to overcome fragmentation, avoid duplication of efforts and build the necessary critical mass
- provide a regular forum at which participants from national funding authorities and leaders of national Grid initiatives can meet at six-monthly intervals to establish synergy between the different projects and programmes.

To achieve the second objective — enhanced collaboration among researchers and users of Grid technology — the project will:

- organise general and specialized workshops
- promote and organise dissemination/education/training initiatives both for users and for research centres
- investigate the status of standardisation activities in Grid computing and devise a strategy for European activities in standardisation.

To achieve the third objective — development of Grid R&D programmes and road maps — the project will:

- develop a coherent and innovative Grid R&D programme vision in Europe taking into account cohesion, and considering present achievements and the multinational character of Grid R&D in Europe. This task will address the strategic research issues identified by the EU Expert Group on Next Generation Grids (June 2003) and the UK Gap Analysis report (May 2003), which have been substantially stimulated by several ongoing national and European research initiatives and/or collaborations at international level
- develop Grid R&D road maps and implementation strategies on the basis of the Grid R&D programme vision.

The multinational character of Grid research in Europe should be maintained and developed according to the intrinsic strengths of the different national programmes. This implies the continued development of both individual national and EU programmes for future Grid research initiatives at the same time as supporting multilateral exchange of experiences and plans to create a more complementary and co-ordinated approach towards Grid Research in Europe.

GridCoord is coordinated by the Dipartimento di Informatica, University of Pisa, Italy; Coordinator: Marco Vanneschi; Project Manager: Roman Tirler. The starting date was July 2004, duration 18 months, budget 960,000 , for a total cost of 1155 .

Link:
http://www.gridcoord.org

Please contact:
Marco Vanneschi, Dept. of Computer Science, Pisa University, Italy
Tel: +39 050 2212738
E-mail: vannesch@di.unipi.it

TrustCoM - A Trust and Contract Management Framework enabling Secure Collaborations in Dynamic Virtual Organisations

by Theo Dimitrakos, Michael Wilson and Santi Ristol

TrustCoM is a new European integrated project that aims to develop a framework for trust, security and contract management in dynamic Virtual Organisations. The framework will enable the secure enactment of collaborative business processes in self-managed, and dynamic value-chains of businesses and governments. The framework will leverage and extend the emerging convergence of open-standards, and in particular Web Services, and open Grid Services architectures. Validation will take place within industrial strength test-beds in collaborative engineering and electronic business. This article provides an overview of the TrustCoM project vision, objectives and anticipated results.

Recent years have seen an unprecedented acceleration in the evolution of the Internet as the technological vehicle underpinning the expansion of service provision and inter-/intra- enterprise integration in all market sectors. This brings about the prospect of ad hoc integration of systems across organisational boundaries to support collaborations that may last for a single transaction or evolve dynamically over many years. This sets new requirements for scalability, responsiveness and adaptability that necessitate the on-demand creation and self-management of dynamically evolving virtual organisations (VO) spanning national and enterprise borders, where the participating entities (enterprises or individuals) pool resources, information and knowledge in order to achieve common objectives. The objectives may be short term - eg to deliver a one-off service in response to a specific customer demand - or long-lasting. In the latter case, the VO's structure, business processes and operational infrastructure must adapt as the goals of the collaboration, the participating entities, the business context and the technologies employed, change.

TrustCOM Consortium.

Emerging ICT paradigms such as Autonomic computing, Utility computing and Grid computing are making the formation and operation of virtual organisations easier by providing dynamic management of the distribution of computational processes across available resources. Notwithstanding the major ICT breakthroughs of the last two decades, protecting one's assets while integrating services, processes and resources, remains a major ICT challenge. Overcoming such challenges requires the development of novel technology realising innovative ideas over widely acceptable interoperable platforms. The required scalability, responsiveness and adaptability for dynamic virtual organisations, makes the provision of cost effective trust and contract management solutions for VO environments, the most demanding and timely research challenge in this field. Effective solutions require interdisciplinary approaches, at a global level, integrating tools from law, cognitive and social science in addition to telecommunications and computing.

The TrustCOM Project

In response to this challenge, the European Commission and a consortium of end-users, major software vendors and telecom operators, and academic research organisations, are implementing the new Integrated Project TrustCoM, which is conducting multi-disciplinary research in order to deliver:

- a novel trust and contract management reference architecture that will enable collaborative work within on-demand created and self-managed dynamic VOs leveraging on the emerging convergence of Web Services and Grid technologies
- a set of conceptual models explaining the fundamental concepts, principles and methods underpinning the above architecture. Effectively these provide the meta-model of any new architectural constructs that may result from TrustCoM research
- a set of profiles, that bring together and potentially extend selected Web/Grid Services specifications at specific version levels, along with conventions about how they work together to support potential implementations of the TrustCoM framework
- a reference implementation of the above integrating and extending already established or emerging interoperability standards for autonomic security, trust and contract management based on Web and Grid services technology
- system and software engineering tools and methods analysis the VO lifecycle and offering a library of design patterns and generic software components implementing selected services that offer the core functionalities of the VO
- testbeds exhibiting instantiations of the above architecture and reference implementation into two classes of realistic application scenarios, namely collaborative engineering (CE) and provision of ad-hoc aggregated services (ADP)
- selected demonstrators exhibiting the business value and benefits of the TrustCoM framework in the above-mentioned application domains
- studies analysing selected aspects of the legal and socio-economic context that underpins such Virtual Organisations.

The TrustCOM Consortium

The TrustCoM consortium brings together researcher from many major middleware vendors and provides a balanced blend of academic and applied researchers, end-user organisations, and enterprises looking to utilise results in products and services. As such it is well-placed to define, conduct and exploit leading edge research that is relevant to the needs of European business, government and society.

Link:
http://www.eu-TrustCoM.com

Please contact:
Theo Dimitrakos, Scientific Coordinator, CCLRC, UK
E-mail: T.Dimitrakos@rl.ac.uk

Michael Wilson, Programme Manager, CCLRC, UK
E-mail: M.D.Wilson@rl.ac.uk

Santi Risto, Operations Manager, Cataluña Atos Origin,Spain
E-mail: Santi.Ristol@atosorigin.com

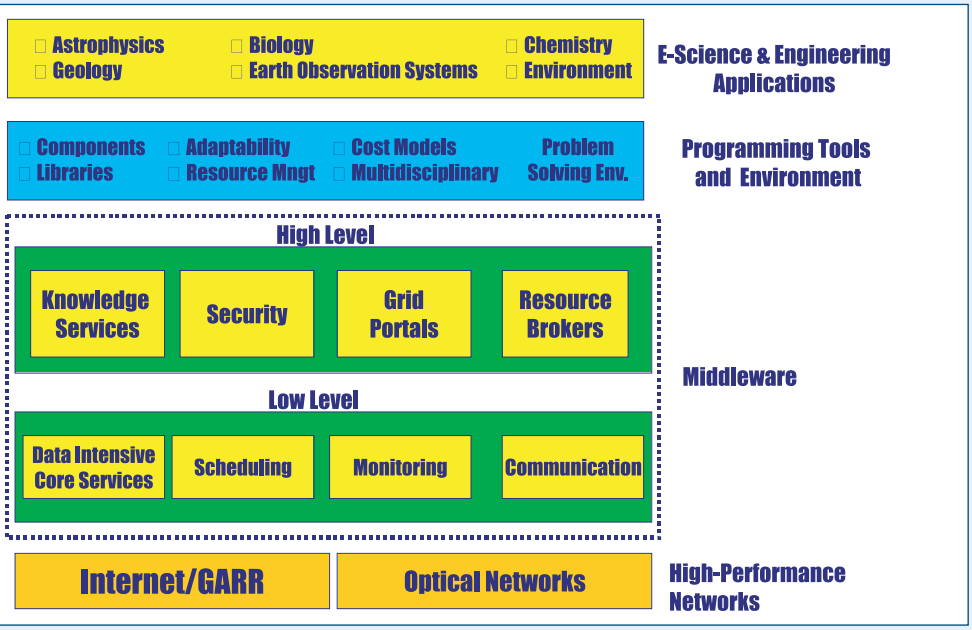
Grid.it – Next Generation Grid Platforms and their Applications

by Domenico Laforenza and Marco Vanneschi

Grid.it is a strategic project in the area of Enabling Technologies for the Information Society, coordinated by the Italian National Research Council. The project has a strong interdisciplinary character and is aimed at defining, implementing and applying innovative solutions for networked computing-enabling platforms, oriented towards scalable virtual organisations and based on the 'Grid Computing' paradigm. The research topics in the project span from high performance photonic networks to innovative middleware services, high-performance programming environments and so forth.

Moving on from the distributed platform based on a distributed infrastructure scenario, Grid.it places special emphasis on the high performance requirements of applications developed on the Grid. This means that the integration of systems and resources, as well as heterogeneity and dynamic situations management, must explicitly handle the case of Grid nodes (which in general are geographically distributed or located on private virtual networks) as high performance systems, such as parallel machine architectures or clusters. Research on high performance extends to all the platform levels, from high bandwidth networks to middleware services and, in particular, to resource management as well as tools and programming environments.

At the programming tools and environment level the high performance requirement implies that, when designing scalable virtual organisations (VOs), a unifying approach should be used. This approach takes into account both aspects related to the distribution of computations and resources, and those related to parallelism. The programming environment must therefore be characterized by a high degree of portability on different hardware-software systems (or different hardware-software combinations) in a heterogeneous and dynamic context. Portability must be guaranteed not only for code but should also ensure that the performance matches the configuration of the target system at hand. Tool interoperability and high performance application reuse are also fundamental to the programming environment.



Research Issues of Grid.it and System Levels.

At the middleware level, research on resource management includes aspects of maximum importance such as discovery, brokering, scheduling, monitoring and performance evaluation/prediction.

The study of the high speed networks needed to support enabling Grid platforms for scalable VOs is an internationally recognized 'hot' topic. Within this research activity an important role is played by experiments on very high bandwidth optical networks, based on photonic technology for Grid platforms with high performance sites that extend to the metropolitan area.

In addition to aspects concerning programming environments and resource management, the software technology studied in Grid.it includes some fundamental aspects related to middleware:

- security: secure Grid environments and cooperation among Grid environments belonging to different organizations
- data intensive services: federated database services, visualization and hierarchical management of data and meta-data according to advances and high-performance techniques
- knowledge discovery services: Grid services (data mining, search engines, etc.) that provide consistent, efficient and pervasive access to high end computational resources
- Grid portals: Grid enabled application services, eg providing the user with the possibility to submit tasks to remote jobs and collect results via Web interfaces.

The design and implementation of scientific libraries that can be used in a heterogeneous and dynamic context, such as the Grid, completes the research on programming environments.

Applications

Grid.it includes the development of some demonstrators selected within application fields that are of maximum interest, not only for their scientific value, but also as testbeds for high performance Grid platforms:

- earth observation
- geophysics
- astronomy
- biology
- computational chemistry.

In order to be able to implement and experiment the ideas and results of the project, a Grid infrastructure will be implemented on a national scale, based on the GARR network and using, in some metropolitan sites, optical fiber interconnections. This project result will provide the community with a national Grid that can be used in different computational science research sectors, and also for commerce, industrial and social service applications.

Organisation

Six Italian research units (ISTI-CNR , ISTM-CNR, ICAR-CNR, INFN, CNIT, ASI) participate in the project and also involve a large number of Italian university departments.

The principal investigator is Marco Vanneschi, Pisa University and ISTI-CNR, Pisa. The Technical Board is coordinated by Domenico Laforenza, ISTI-CNR, Pisa.

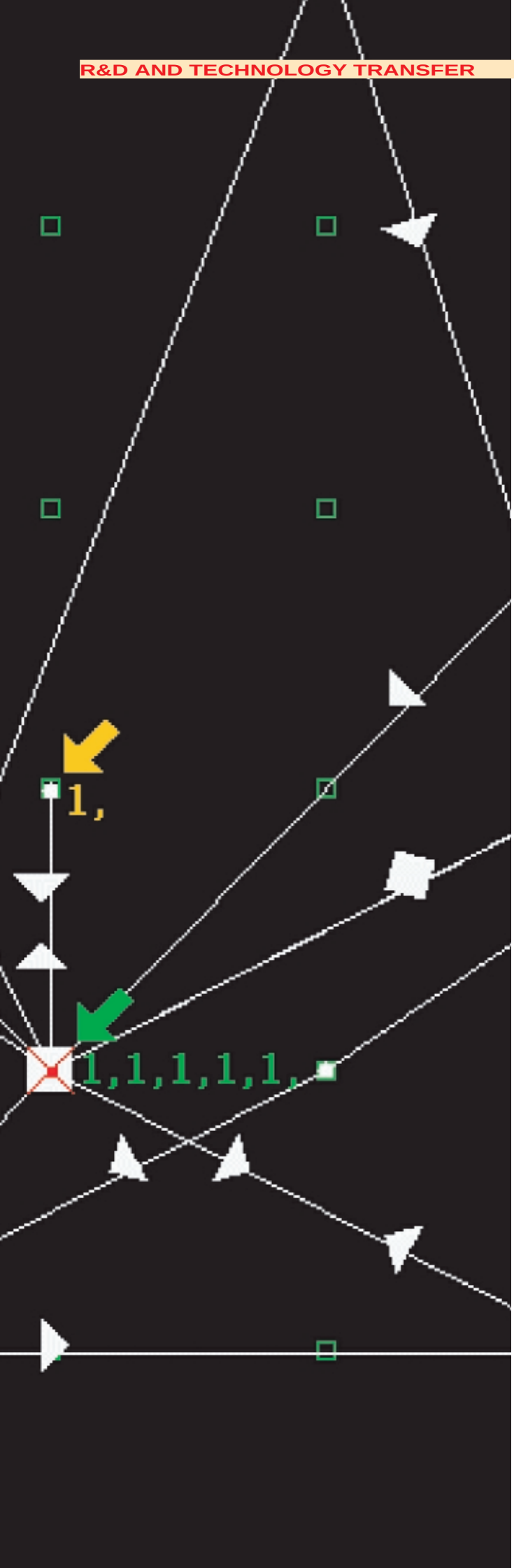
The Grid.it project aims at playing an important role in the training of highly qualified young people. A relevant part of the project budget is reserved for contracts for young researchers.

The project has funding of 8.1 M from the Italian Ministry for Education and Research for three years (November 2002 – November 2005).

Link:
http://www.grid.it

Please contact:
Domenico Laforenza, ISTI-CNR, Italy
Tel: +39 050 3152992
E-mail: domenico.laforenza@isti.cnr.it

Marco Vanneschi, Pisa University, Italy
Tel: +39 050 2212738
E-mail: vannesch@di.unipi.it



Articles in this Section

Grammar Enables Effective Multimedia Search Queries

by Fedde van der Lijn and Menzo Windhouwer

Retrieving files from a multimedia database is like finding a book in a library – without a catalogue of keywords you could search for ages. However, generating and updating such a catalogue is almost as difficult. CWI introduces so-called feature grammar systems to facilitate these tasks.

Multimedia is everywhere. Libraries and museums digitize their collections and make them available to interested parties. Cheap disk space has brought the storage of large amounts of multimedia within everyone's reach. At the same time however, it has complicated the retrieval of objects.

Not only have databases become larger, they also contain more types of media. While most people are able to find a text containing certain keywords, anyone who has ever used Google Image Search to look for a specific picture knows that finding what you need can be far from easy. The majority of search engines can only interpret textual data. They cannot 'see' what is depicted in an image or 'hear' what is on an MP3.

Annotation

One way of dealing with this problem is to annotate media objects in advance. Annotations describe particular features of the stored media objects and can be used to guide semantic search queries. For example, when MP3s are annotated for genre and background information on the performing artist, search queries like 'find me all blues songs played by guitarists from Mississippi' can give meaningful results.

Annotation can be done manually, but for large multimedia collections this quickly becomes impossible. It is therefore necessary to turn to automatic annotation using extraction algorithms. These algorithms can perform easy tasks, such as finding the length of an MP3, or complicated ones, like detecting human faces in images. Designing clever extraction algorithms is a necessary condition for an automatic annotation system.

Context Dependency

Yet high-quality extraction algorithms alone are not enough. Equally important is a system to coordinate the extraction of annotations. The main problem is that annotations depend on context, and in practice this means they depend on each other. This has consequences for the way extraction algorithms should be used. For example, feeding charts or logos to a face extractor would be a waste of system resources and time. The annotation system makes sure the extractor is only applied to images with a high chance of containing faces, like photos or drawings.

With the advent of digital cameras, DVDs and MP3s it has become very easy to compile large multimedia libraries. At the same time, it has become increasingly difficult to effectively search collections that contain a wide range of media types like text, images, movies and audio.

A lack of context also complicates incremental maintenance. Without knowledge of context and interdependencies, the entire annotation process must be rerun every time an extraction algorithm is added or replaced. The annotation system finds dependencies and determines exactly which annotations should be updated and which can be re-used.

Context dependency of annotations can also cause extraction algorithms to give ambiguous answers, especially when dealing with complex features. This can result in outputs like 'this image either contains a human or a pig'. The annotation system should be able to handle these ambiguities.

Valid Sentences

CWI solves these difficulties by combining database technology with ideas from formal language theory to form the Acoi annotation management system. Its basis is the theory of feature grammar systems. Such a system not only describes the annotations themselves, but also their dependencies and contexts. Feature grammar can be compared to grammar in natural language. Grammar determines which word classes can be combined and in what order to form a valid sentence. A feature grammar system does the same for annotations and extraction algorithms. It determines what extraction algorithms must be called to form a valid annotation 'sentence'.

Since the feature grammar system also stores each annotation's place in the network of interdependencies, incremental maintenance is possible. When updating the database, sentences can be reinterpreted to determine which extractions must be redone. Furthermore, techniques from formal language theory could be modified and used in the annotation system. Resolving ambiguities, for

example, is a classic problem in this branch of computer science.

Case Studies

Acoi has proven its capabilities in a variety of case studies. Together with a number of basic extraction algorithms it was used to create an annotation index for a collection of Web pages. Furthermore, it was used in combination with a presentation generator to unlock the digitized collection of the Rijksmuseum Amsterdam to the public. The generator, also developed at CWI, uses the annotations to automatically compose a semantically structured multimedia presentation on a user-defined subject.

CWI's feature grammar system is unique. Other annotation systems have been developed, but all lack the explicit storage of annotation context. As a result, Acoi is the only system that elegantly handles ambiguities and allows for incremental maintenance. In the near future, Acoi will be used in the MultimediaN project in which CWI participates.

Links:
http://www.cwi.nl/~acoi
http://www.windhouwer.nl/menzo/professional/index.html
http://www.cwi.nl/ins1

Please contact:
Menzo Windhouwer,
University of Amsterdam, The Netherlands
Tel: +31 20 525 3104
E-mail: M.A.Windhouwer@uva.nl

Martin Kersten, CWI, The Netherlands
Tel: +31 20 592 4066
E-mail: Martin.Kersten@cwi.nl

From Textual Use Cases to Behaviour Specifications

by Vladimir Mencl

Traditionally, use cases are written in plain English to make the requirement specifications accessible to a wide audience. It is widely accepted that automatic processing of textual use cases is not feasible in general; this is considered as a trade-off for the high readability of textual use cases. Nonetheless, there is ongoing research exploring the options to apply linguistic techniques to textual use cases. The Procasor project of the Distributed Systems Research Group at the Charles University in Prague aims to convert textual use cases into event-based behavior specifications, behavior protocols.

The functional requirements for a future system are typically documented with use cases. In the traditional approach, plain English is used for use cases. A strong benefit of this approach is that use cases are comprehendible to a wide audience. A textual use case describes the possible scenarios of the future system, by showing how a particular goal is achieved by the System under design (SuD) and its surrounding actors in a sequence of steps. A step of a textual use case describes an action that is either a request or information passed between an actor and SuD actors, or an action internally processed by SuD. Figure 1 shows a fragment of a textual use case.

An obvious drawback of writing use cases in plain English is the lack of formal tools to reason upon the requirement specifications. Also, the subsequent stages of the software development process have to start with manually processing the use cases – eg, identifying operations of future objects. Luckily, use cases are typically not written in plain natural language. The guidelines for writing use cases significantly restrict the language by asking for simple and uniform sentence structure; a use case sentence should be in active voice, following the simple pattern 'subject – verb – direct object – prepositional phrase'. The subject should be the entity (SuD or an actor) active in the task.

Readily available linguistic tools, eg, the statistical parser developed by Michael Collins at the University of Pennsylvania, permit obtaining the parse tree for a sentence. While the parsing is not 100% reliable (it is statistical parsing), the accuracy improves for simple sentences – as is the case for use cases. The simple sentence structure permits us to obtain the principal attributes of the action described by the sentence from its parse tree. With only a simple domain model (names of entities and a list of conceptual objects), we can identify the entity active in a particular step of a use case. Depending on the role of the active entity in the use case (SuD vs. an actor), it is possible to determine whether the action described by the step is receiving a request, sending a request, or an internal action of SuD.

We illustrate the approach on the parse tree of step 1 shown in Figure 2. The subject of the sentence, 'Seller', a noun (NN), is the name of an actor involved in the use case. Therefore, the step describes an event (a request) received by SuD from the entity Seller. Subsequently, we estimate the event label for representing the request. In the parse tree, we identify the verb (VBZ) 'submits' and the direct object 'item description'. Depending on conceptual objects captured in the domain model, the words constituting the direct object phrase are selected; here we acquire the event token *?SL.submitItemDescription* to represent the step 1. The step 2 describes an action internally performed by the SuD, which we represent with the event token *#validateDescription*.

Use Case: #1 Seller submits an offer

Scope: Marketplace

SuD: Marketplace Information System

Actors: Seller, Trade Commission

Main success scenario specification:

1. Seller submits item description.
2. System validates the description.

...

Figure 1: Fragment of a textual use case.

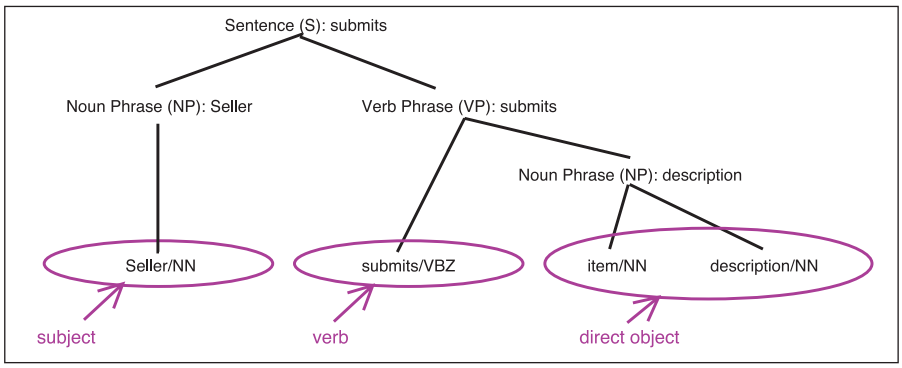


Figure 2: Parse tree of sentence "Seller submits item description".

In a subsequent stage, we convert the whole use case into a behavior protocol (IEEE Trans. Software Eng. 28(11), developed within our group), a formal method featuring, eg, a decidable compliance relation. We have implemented the conversion outlined here in the Procasor tool (Pro-case stands for Protocol use case). As the tool is employing a statistical natural language parser, the output is only an estimate of the behavior specification; nonetheless, it can prove useful in obtaining an initial design of interfaces of the future system. Our future work focuses on enhancing the transformation to produce matching event tokens for complementary actions; the long-term goal is to apply reasoning available for behavior protocols to detect inconsistencies in use case specifications.

Link:
http://nenya.ms.mff.cuni.cz/

Please contact:
Vladimir Mencl, Charles University Prague/ CRCIM,Czech Republic
Tel: +420 2 2191 4232
E-mail: vladimir.mencl@mff.cuni.cz

GHT*: A Peer-to-Peer System for Metric Data

by Michal Batko, Claudio Gennaro and Pavel Zezula

GHT* is a scalable and distributed similarity search structure that has been specifically designed to support metric space objects. The structure is based on the Peer-to-Peer (P2P) communication paradigm. It is scalable, which means that the query execution achieves practically constant search time for similarity range queries in data-sets of arbitrary size. Updates are performed locally and node splitting does not require sending multiple messages to many peers.

P2P communication paradigms are quickly gaining in popularity due to their scalability and self-organizing nature, forming the basis for building large-scale similarity search indexes at low cost. However, most of the numerous P2P search techniques proposed in recent years have focused on single-key retrieval. A good example is the Content Addressable Network (CAN), which is a distributed hash table abstraction over Cartesian space.

Our objective is to develop a distributed storage structure for similarity search in metric spaces that would scale up with (nearly) constant search time. The advantage of the metric space approach for data searching is its 'extensibility', allowing us to perform exact match, range, and similarity queries on any collection of metric objects. Since any vector space is covered by a metric space with an appropriate distance function (for example the Euclidean distance), even n-dimensional vector spaces are handled easily. Furthermore, there are numerous metric functions able to quantify similarity between complex objects, such as free text or multi-media object features that are very difficult to manage. For example, consider the edit distance defined for sequences and trees, the Hausdorff distance applied for comparing shapes, or the Jacard coefficient, which is often used to assess the similarity of sets. Much research is now focused on developing techniques to structure collections of metric objects so that search requests can be performed efficiently.

A convenient way to assess similarity between two objects is to apply metric functions to decide the closeness of objects as a distance, which can be seen as a measure of the objects 'dis-similarity'. For any distinct metric objects x, y, z, the distance must satisfy the properties of reflexivity, $d(x,x) = 0$, strict positiveness, $d(x,y) > 0$, symmetry, $d(x,y) = d(y,x)$, and triangle inequality, $d(x, y) \leq d(x, z) + d(z, y)$.

The distributed environment is composed of network nodes (peers), which hold metric objects, execute similarity queries on stored data and communicate with other peers. Every peer is uniquely denoted by its identifier PID. Peers hold data in a set of buckets. Each bucket has a unique identifier within a peer, designated as BID. Each peer also maintains a tree structure called the address search tree (AST). This is used to route similarity range queries through

the distributed network. AST is based on the metric data partitioning principle known in the literature as the Generalized Hyperplane Tree (GHT). GHT is a binary tree with metric objects stored in leaf nodes implemented as buckets of fixed capacity. Inner tree nodes contain selected pairs of objects called pivots. Respecting the metric, the objects closer to the first pivot appear in the left subtree and those closer to the second are in the right subtree.

We start by building a tree with only one root node represented by a bucket B_0 When the bucket B_0 is full we must split it: we create a new empty bucket B_1 and move some objects (half of them if possible) into it to gain space in B_0. See Figure 1. The split is done by choosing a pair of pivots p_1 and p_2 from B_0 and moving into bucket B_1 all objects o that satisfy the condition $d(p_1, o) > d(p_2, o)$.

Pivots p_1 and p_2 are associated with a new root node and thus the tree grows one level. This split algorithm can be applied on any leaf node and is an autonomous operation (no other tree nodes need to be modified).

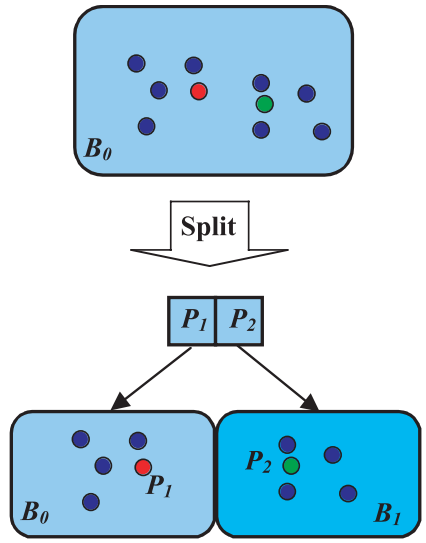


Figure 1: Split of a GHT bucket.

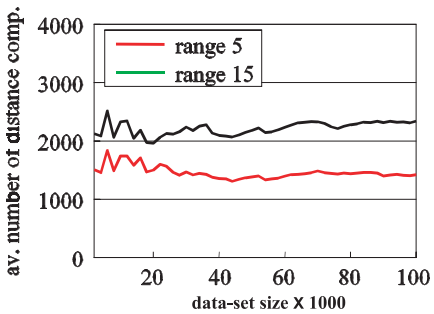


Figure 2:
Parallel response time of a range query.

The most important advantage of GHT* with respect to single-site access structures is its scalability through parallelism. As the size of a data-set grows, new server nodes are plugged in and both their storage and the computing capacity are exploited. Figure 2 shows the result of the parallel response time of a range query, determined as the maximum of the costs incurred on servers involved in the query plus the search costs of the AST. For evaluation purposes, we use the number of distance computations (both in the AST and in all the accessed buckets) as the computational costs of a query execution. This experiment shows how the parallel search time becomes practically constant for arbitrary data volume and the larger the dataset the higher the potential for interquery parallelism. This is the most important feature of GHT*.

Our future work will concentrate on strategies for updates (object deletion) and pre-splitting policies, plus more sophisticated mechanisms for organizing buckets.

Please contact:
Claudio Gennaro, ISTI-CNR, Italy
Tel: +39 050 315 2888
E-mail: claudio.gennaro@isti.cnr.it

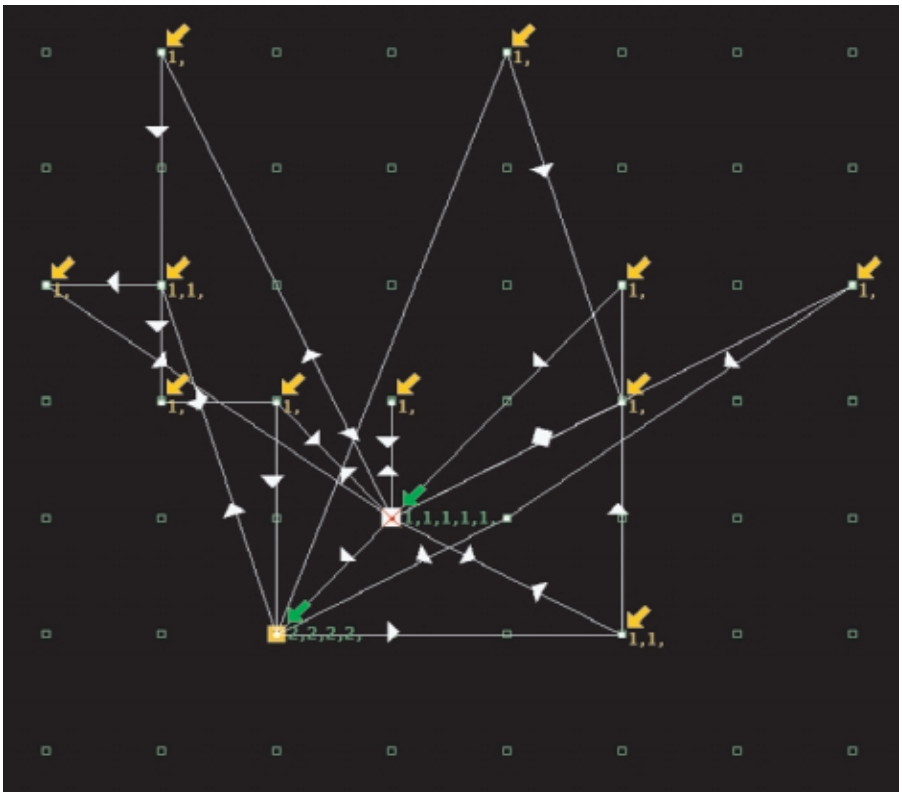
CWI's FACS Cluster - Facilitating the Advancement of Computational Science

by Stefan Blom, Sander Bohte and Niels Nes

The Centre for Mathematics and Computer Science (CWI) performs a wide range of research. Consequently, its high-performance computing facilities must meet diverse requirements. This has led to the purchase of a cluster with an asymmetric core-edge design. Under the influence of this asymmetric structure, some cluster applications are evolving into Grid applications.

Many research groups at CWI share a need for high-performance computer facilities. In the FACS (Facilitating the Advancement of Computational Science) project, three of these groups combined their resources to set up a large shared cluster. Since each user group has its own specific demands on memory, CPU power and bandwidth, the FACS cluster deviates from the more usual symmetric design. The cluster's edge serves as a compute farm. It contains several groups of machines, whose main task is to provide CPU cycles. The core consists of a small number of machines, connected by a high-bandwidth low-latency network (Infiniband) and each equipped with a large amount of memory. Currently, we have 22 (mostly dual CPU) machines in the edge and two quad AMD Opterons with 16G each in the core.

The Database Architectures and Information Access group works on database technology and needs machines with both as much memory as possible and huge, fast storage. For this type of work, the bandwidth between memory and storage is critical. Furthermore, part of the research consists of designing highly efficient systems for extremely large databases. Since the research considers typical database hardware set-



Applications of the FACS cluster include simulations of multi-agent systems such as DEAL (Distributed Engine for Advanced Logisitics). In this project, agents are used to compute efficient schedules for freight trucks.

ups, the large-memory fast-disk multi-CPU nodes of the cluster core are a good model for such set-ups.

The Evolutionary Systems and Applied Algorithmics group is working on machine-learning algorithms and agent technology. The main issue here is sufficient CPU cycles for testing variations on adaptive algorithms and for designing and simulating (large) multi-agent systems.

The third group, Specification and Analysis of Embedded Systems, works on explicit-state model checking. In model checking the task is to verify that a model of a system satisfies a property. This can be split into two steps. The first step is generating the model (a very large graph called a state space), which is a CPU-intensive task; the second step is checking if the desired property holds in the model, which is a memory-intensive task.

Originally, CWI's model-checking system was set up as a cluster application. Both steps were performed on the same types of machines. The choice of an asymmetric cluster has forced the Embedded Systems group to modify the system to a more Grid-like application, which assigns a task to the part of the network best equipped to handle it.

To achieve this, the verification step is separated into two stages. In the first stage, the state space is reduced modulo bisimulation. The second stage verifies the property on the reduced state space. The reduction step yields a state space that has the same properties as the original but is often an order of magnitude smaller. To exploit this, the Embedded Systems group has developed distributed reduction tools.

Without a distributed reduction stage, distributed state-space generation can easily produce state spaces that do not fit into the memory of a single machine. This would make validation and reduction impossible. Reduction can be useful even when using distributed model-checking tools, as the memory use of model-checking tools grows as a function of not only the size of the state space, but also the complexity of the property.

Using these tools, the work on state-space generation can be focused on the edge part of the cluster, whereas the verification step can take as much advantage as possible from the fast interconnect of the core. At the end of this evolutionary path lies a true Grid application, in which state-space generation is performed wherever there is CPU time available, and verification is performed on either a high-performance cluster or a large SMP within the grid.

The main issue currently holding back a Grid solution is the amount of data that must be transferred. In the model-checking system this can be up to half the memory. For database research the bandwidth requirements are even larger. The local storage of clusters and super-computers is usually capable of dealing with these amounts of data. However, the networks connecting the various Grid locations are typically quite slow compared to the internal networks. This can result in hours of data transfer for perhaps ten minutes of computation. DATAGRID research and infrastructure should allow model checking to work on a grid. Eventually, the internal database operations considered by the database group can become a Grid topic as well, but this a more distant possibility.

The computational tasks of research into machine learning and agent technology could be met with Grid solutions. However, the rapid prototyping typically involved makes a small cluster an efficient least-effort solution. As grids becomes more pervasive and transparent to use, a transition to Grid technologyis highly probable.

Link:
http://db.cwi.nl/projecten/project.php4?prjnr=162

Please contact:
Stefan Blom, CWI, The Netherlands
Tel: +31 20 592 4280
E-mail: Stefan.Blom@cwi.nl

Sander Bohte, CWI, The Netherlands
Tel: +31 20 592 4074
E-mail: Sander.Bohte@cwi.nl

Niels Nes, CWI, The Netherlands
Tel: +31 20 592 4098
E-mail: Niels.Nes@cwi.nl

SynDEx: System-Level CAD Software for Optimizing Distributed Real-Time Embedded Systems

by Yves Sorel

Distributed real-time embedded systems are of crucial importance in application domains such as avionics, automobiles, telecommunication and mobile robotics. Principally based on digital electronics, which includes software, this aspect of the various applications is rapidly growing. Because they must meet constraints on resource distribution and optimization, as well as time, the design of such systems is particularly complex. In order to assist the designers, scientists at INRIA are proposing a methodology called AAA (Algorithm Architecture Adequation) and its associated system-level CAD software called SynDEx. These cover the whole development cycle, from the specification of the application functions, to their implementation running in real time on a distributed architecture composed of processors and specific integrated circuits.

AAA/SynDEx provides a formal framework based on graphs and system-level CAD software. On the one hand, these specify the functions of the applications, the distributed resources in terms of processors and/or specific integrated circuit and communication media, and the non-functional requirements such as real-time performances. On the other, they assist the designer in implementing the functions onto the resources while satisfying timing requirements and, as far as possible, minimizing the resources. This is achieved through a graphical environment, which allows the user to explore manually and/or automatically the design space solutions using optimization heuristics. Exploration is mainly carried out through timing analyses and simulations. The results of these predict the real-time behaviour of the application functions executed on the various resources, ie processors, integrated circuits and communication media. This approach conforms to the typical hardware/software codesign process. Finally, for the software part of the application, code is automatically generated as a dedicated real-time executive, or as a configuration file for a resident real-time operating system such as Osek or RTlinux.

This approach will improve the design safety provided by formal models, and decrease the development cycle thanks to timing simulation and automatic code generation. Another interesting feature is the ability to easily interface AAA/SynDEx with domain-oriented specification languages such as the Synchronous languages, AIL for automobile, Scilab/Scicos for automatic control, AVS for image processing and so on. This means a link can be provided to users' most commonly used tools, and when these languages guarantee timing properties, that these will be maintained during the distributed real-time implementation. This is particularly important in the case of safety-critical applications found in the field of avionics and automobiles.

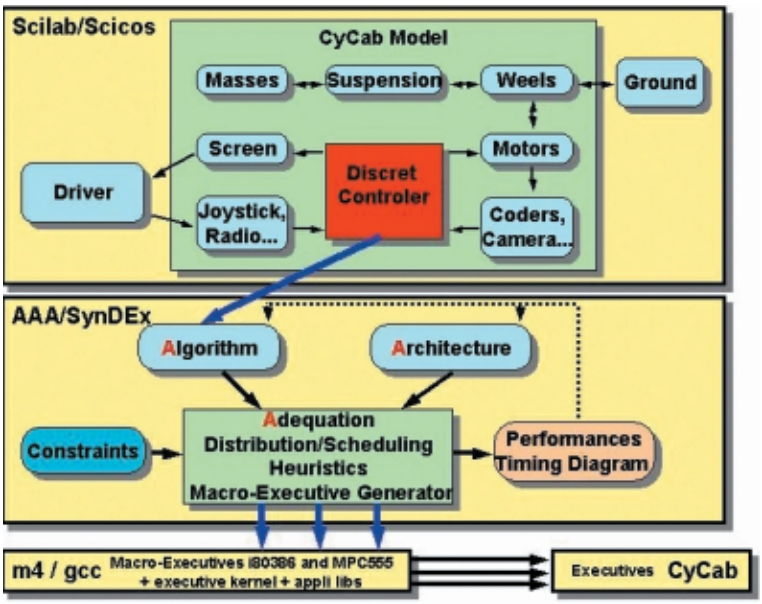


Figure 1: SynDEx interfaced with Scicos.

Figure 1 shows the principles of SynDEx interfaced with the hybrid dynamic systems modeller and simulator Scicos.

Figure 2 shows the SynDEx graphical user interface used to design a manual driving application with joystick for the CyCab modelled and simulated with Scicos. The CyCab is an intelligent and modular electric vehicle designed at INRIA Rocquencourt by the IMARA team and industrialized by Robosoft. Its architecture is based on four MPC555 microcontrollers and an embedded PC, all of which communicate through a CAN bus.

Since the early nineties, there has been a significant amount of research on AAA/SynDEx at INRIA Rocquencourt, firstly by the SOSSO, then the OSTRE, and now the AOSTE team. As time has

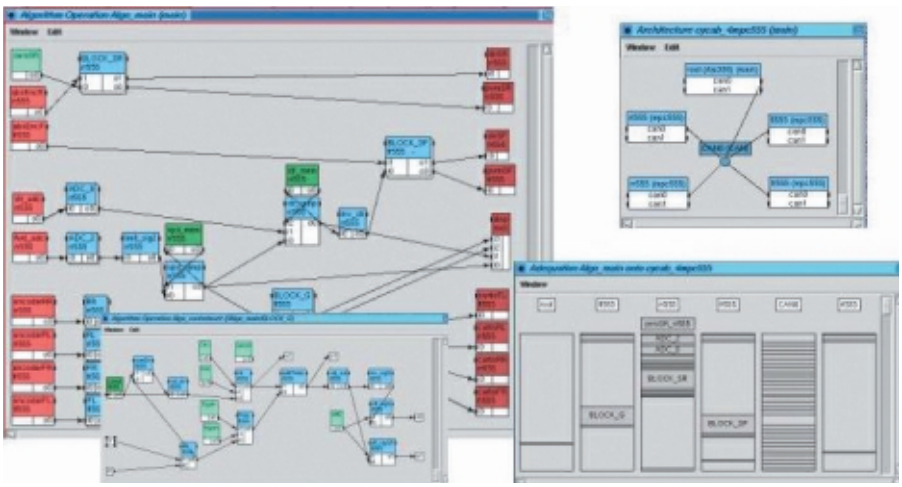


Figure 2: SynDEx GUI used to design a manual driving application with a joystick for the CyCab.

passed, it has increased in popularity with industrial users, and has therefore been evaluated by several companies. This occurred firstly in France and three years ago in Europe, through ITEA projects launched in automobile and telecommunications research. It is also currently used in real-world applications provided by Robosoft, MBDA, Mitsubishi-ITE, and PSA.

AAA/SynDEx runs under Unix/Linux, Windows and MacOS and comes with full documentation, including a reference manual, a user manual and a tutorial. It is downloadable free of charge under INRIA copyright at http://www.syndex.org.

Future work will focus on the following aspects, from high to low levels of the design flow: the integration of AAA/SynDEx with the MDA (Model Design Architecture) approach by proposing successive model transformations; the extension of currently static optimization techniques to more dynamic schemes in order to better support event handling and dynamic creation of functions; and finally, the tight coupling of system-level and circuit-level CAD tools in order to actually provide automatic hardware/software partitioning in the co-design process.

Links:
http://www.syndex.org
http://www.scilab.org
http://www-rocq.inria.fr/imara,
http://www.robosoft.fr

Please contact:
Yves Sorel, INRIA, France
Tel: +33 1 39 63 52 60
E-mail: yves.sorel@inria.fr

Large-Scale Simulations of Diffusion in Cell Biology

by Ivo F. Sbalzarini, Anna Mezzacasa, Ari Helenius and Petros Koumoutsakos

Molecular transport within the intracellular structures of live biological cells is dominated by diffusion in confined compartments with complex geometries. Quantitative evaluations of standard experimental assays in cell biology (eg to measure diffusion coefficients of proteins in vivo) require knowledge of the solution of the diffusion equation in such geometries. We present a high-performance parallel implementation of a particle method to solve the diffusion equation in 3D reconstructions of real samples obtained by fluorescence confocal microscopy.

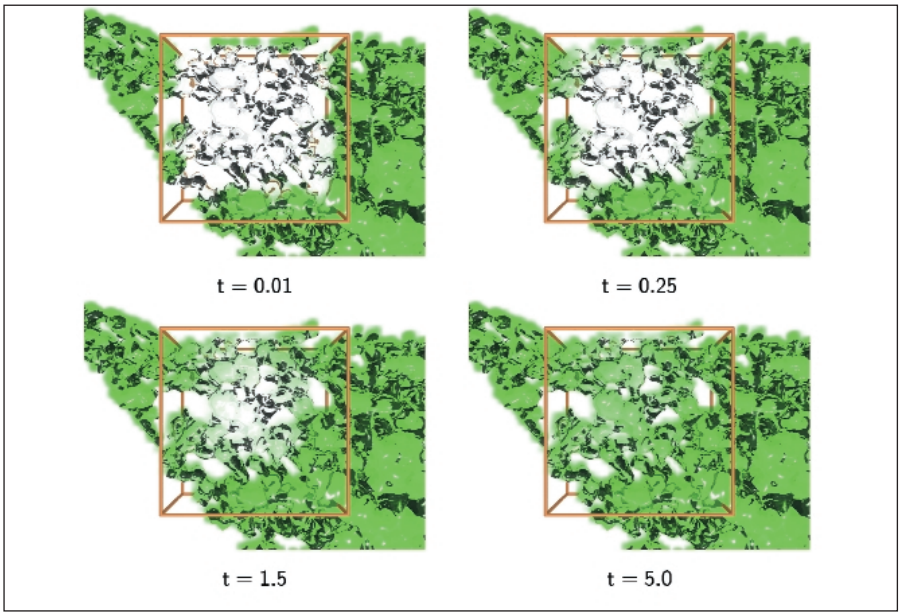
Due to the high structural organization of the intracellular space, diffusion is always restricted to the specific shape of the organelle under consideration. To determine and quantify the diffusive mobility of substances within live cells the technique of fluorescence recovery after photobleaching (FRAP) is widely employed.

From a practical point of view a method is needed to deduce molecular (microscopic) solute diffusion constants from measured fluorescence recovery data. We propose the deterministic particle method of Particle Strength Exchange (PSE) for simulations of diffusion in cellular geometries of realistic complexity. PSE was introduced as an alternative to the method of random walk, enabling higher-order simulations of convection-diffusion processes. PSE is a grid-free, deterministic particle method and thus combines the advantages of high-order convergence and geometric flexibility. It enables efficient simulations of solute diffusion in biological structures using the geometry of real samples.

Due to the complex shape of the computational domain, a large number of particles is usually needed to resolve the concentration field. We implemented an efficient parallel PSE code based on a general-purpose Parallel Particle Mesh (PPM) Library currently under development. The code provides different domain decomposition techniques, dynamic load balancing among inhomogeneous processor clusters, parallel file

I/O and efficient network communication. It was tested and optimized both on distributed memory and shared memory systems with either scalar or vector processors. Arbitrary geometries can be read in as triangulated surfaces and particles are initialized inside it. (Using a particle method, there is no need for grid generation.)

The method was successfully applied to diffusion in the Endoplasmic Reticulum (ER), which is the most prevalent example of a highly convoluted and interconnected structure in three dimensions. The geometry of several ER samples was reconstructed by recording the shape of the ER samples from live cells using stacks of serial sections obtained by confocal fluorescence light microscopy. It was checked that the reconstructed 3D geometry of the organelle was connected and exhibited all the topological properties of a real ER. It could thus be used directly as a computational domain. Figure 1 shows 3D visualizations of the evolving concentration field inside a sample ER geometry. Using reconstructions from real samples allowed direct comparisons between simulation results and experimental data from the same ER.



Snapshots of the concentration distribution from a sample PSE simulation. The ER membrane is visualized as a transparent surface, and the concentration of the protein under consideration as a volume density cloud inside it. The initial concentration field was 1 everywhere outside the cubic box shown, and 0 inside it. The region of interest around the box is enlarged.

Fitting simulated and experimentally measured diffusion curves using time stretching enabled us to determine the unknown molecular diffusion constant in live cells. By using the same geometry both for experiment and simulation, all need for modelling its effect on the apparent diffusion constant is eliminated and the only parameter left is the molecular diffusion constant itself, taking the influences of the specific geometry at hand into full account.

Please contact:
Petros Koumoutsakos,
ETH Zurich/SARIT, Switzerland
E-mail: petros@inf.ethz.ch

Real-Time MIMO Testbed for Next Generation Wireless LANs

by Andreas Burg and Helmut Bölcskei

At ETH Zurich, a MIMO-WLAN (Multiple-Input Multiple-Output Wireless Local Area Network) system has been developed and demonstrated successfully in a test setup. Four antennas at the transmitter and at the receiver boost the data rate from 54 million bits per second (Mbps) in a regular 802.11g-based WLAN to up to 216 Mbps. In addition to the higher throughput, the system provides increased link reliability and better coverage.

With the establishment of the IEEE-802.11a/b/g standards, wireless computer networks have started to gradually replace regular wired Internet access in homes, offices, and public areas, such as train stations and airports. However, as only limited bandwidth is allocated to be used by such systems, data rates need to be shared. As the number of users grows and applications become more demanding in terms of throughput and link reliability (quality of service), current standards can no longer meet the requirements.

Multiantenna (MIMO) systems are a means to increase the data rate, link reliability, and range of all kinds of wireless networks and mobile communication systems, without increasing bandwidth and/or transmit power. Three main effects are responsible for this performance gain: The 'diversity gain' stabilizes the link, as the additional antennas can be used to compensate for a weak connection on any of the other antennas. The 'array gain' allows the receiver to pick up more of the energy that was radiated from the transmitter, effectively increasing the range of the system. The

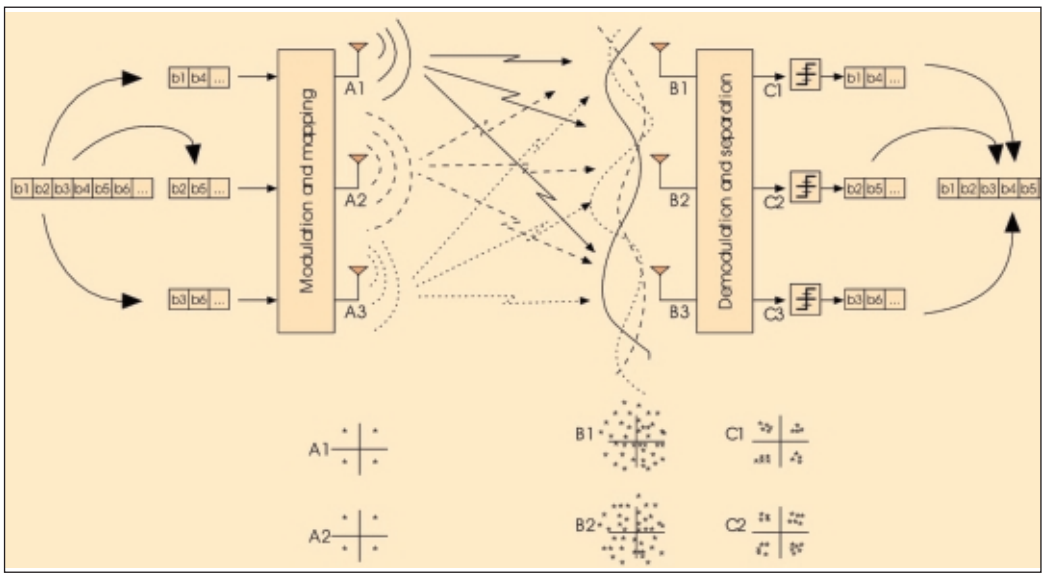


Figure 1: MEMO scheme.

'multiplexing gain' finally allows transmitting multiple data streams concurrently in the same frequency band, thereby effectively increasing the data rate. A schematic of such a system with three antennas is shown in Figure 1. The high-rate data stream is first partitioned into three lower rate streams, which are then sent simultaneously from the individual antennas, using the same frequency band. At the receiver, three antennas pick up the superposition of the streams, and a suitable decoder separates them again and reconstructs the original high-rate stream.

Researchers at the Communication Technology Laboratory and at the Integrated Systems Laboratory of the Swiss Federal Institute of Technology (ETH) Zurich are working on the development of such MIMO (Multiple-Input Multiple-Output) systems and on their integration in today's WLAN systems.

Research activities focus on both the theoretical analysis and the VLSI implementation of such systems. At the Communication Technology Laboratory, the limits of multiantenna systems are investigated, and new algorithms that approach these limits in practical systems are developed. The additional degrees of freedom induced by multiple antennas can be exploited by different algorithms, which need to be tailored to individual application scenarios. The implementation complexity of these algorithms generally exceeds the complexity of algorithms applied in single-antenna systems significantly. For this reason, researchers at the Integrated Systems Laboratory are concerned with algorithm optimization and efficient implementation in application-specific integrated circuits (ASIC). The challenge is to meet throughput requirements while maintaining a small chip area to reduce cost and to guarantee low power consumption. A recent example of this joint effort is the development of a decoder ASIC (sphere decoder) that allows the optimum separation of parallel data streams, received on multiple antennas.

Within the scope of a three years project, financed by the ETH research council, a real-time demonstrator for a MIMO-WLAN system with four antennas each at the transmitter and the receiver has been developed. The system boosts the data rate of a current single antenna IEEE-802.11a WLAN system to 216 Mbps, without the need for additional bandwidth and/or transmit power. Such initial practical experiments allow researchers to test new algorithms in real-time and in real-world scenarios. They are also indispensable contributions to standardization efforts (eg, the ongoing work in the high-throughput 802.11n working group) to identify implementation bottlenecks and other pitfalls that might appear in their hardware realization. Recently, first over-the-air experiments have been carried out, indicating that the predicted gains are in fact achievable and that the proposed system can be realized with manageable hardware complexity that will eventually allow integration into a laptop computer or personal digital assistant (PDA).

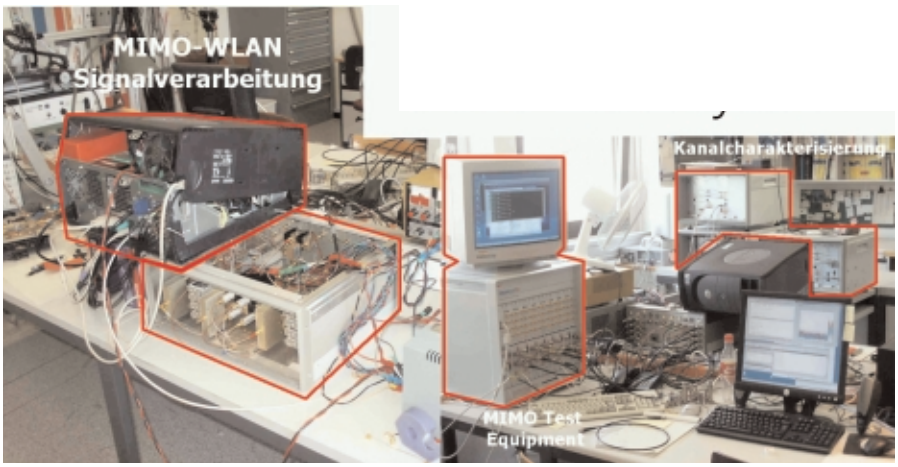


Figure 2: MIMO WLAN system experimental setup — signal processing, MIMo test equipment, channel characterization.

Please contact:
Helmut Bölcskei, ETH Zurich/SARIT
E-mail: boelcskei@nari.ee.ethz.ch

16th European Conference on Artificial Intelligence

by Vicent J. Botti

The 16th European Conference on Artificial Intelligence (ECAI 2004) was held in Valencia, Spain, from 22-27 August 2004, hosted by the Universidad Politécnica de Valencia and the Grupo de Investigación de Tecnología Informática –Inteligencia Artificial (GTI-IA). ECAI provided a public forum for researchers from academy and industry and public organizations, offering them an invaluable occasion to meet and exchange ideas, thus contributing to making real the upcoming 'Information Society'.

The conference included, from August 22 to 24, 2004, two days and a half of workshops (27 workshops) and tutorials (11 tutorials) on most advanced topics, as well as the second starting AI Researchers Symposium (STAIRS), followed by the three-day technical program of ECAI. Following an established tradition, the sub-conference on prestigious applications of intelligent systems (PAIS) ran in parallel with ECAI. The technical programme included refereed paper presentations, invited talks by prestigious speakers and poster sessions. The invited speakers were Glorianna Davenport (MIT, USA), Christian Freksa (University of Bremen, Germany), Carole Goble (University of Manchester, UK), and Seppo Laukkanen (SenseTrix, Finland).

Among 653 submissions received from 45 countries, 168 papers were accepted for oral presentation at ECAI 2004, 13 for oral presentation at PAIS 2004, and, for the first time, 87 were accepted as posters.

Although the majority of the submitted papers were from Europe (505), there has been a significant number of papers from the rest of the world, and particularly from Asia (53), North America (50) and Oceania (21). We can also notice the emergence of high quality AI research in South America (16) and Africa (8), and this is an indicator of the increasing interest in AI worldwide.

After the late registration deadline the main conference (ECAI'04) has 694 registrations, STAIRS has 53 registrations, and 436 people have registered for the 27 workshops.

More information:
https://www.dsic.upv.es/ecai2004/

Cross-Language Evaluation Forum — CLEF 2004

by Carol Peters

The results of the fifth campaign of the Cross-Language Evaluation Forum were presented at a two-and-a-half day workshop held in Bath, UK, 15-17 September, immediately following the eighth European Conference on Digital Libraries. The workshop was attended by nearly 100 researchers and system developers.

The main objectives of the Cross-Language Evaluation Forum (CLEF) are to stimulate the development of mono- and multilingual information retrieval systems for European languages and to contribute to the building of a research community in the multidisciplinary area of multilingual information access. These objectives are realised through the organisation of annual evaluation campaigns and workshops. Each campaign offers a series of evaluation tracks designed to test different aspects of mono- and cross-language system performance.

One of the principal objectives when CLEF began was to encourage developers to build multilingual retrieval systems capable of searching over collections in a number of languages. The multilingual track was thus the main track for several years and was made progressively more difficult. With CLEF 2003, where the track included a task which involved finding relevant documents in a collection containing documents in eight languages, we felt that we had achieved an important goal. We had shown that fully multilingual retrieval could be (almost) as effective as bilingual (L1 -> L2) retrieval and that systems are able to adapt and reengineer rapidly and effectively to process new languages as the need arises. For this reason, in CLEF 2004 we decided to reduce the multilingual document retrieval activity to leave more space for

other types of cross-language information retrieval experiments.

Six tracks were offered to evaluate the performance of systems for:

- mono-, bi- and multilingual document retrieval on news collections (Ad-hoc)
- mono- and cross-language domain-specific retrieval (GIRT)
- interactive cross-language retrieval (iCLEF)
- multiple language question answering (QA@CLEF)
- cross-language retrieval on image collections (ImageCLEF)
- cross-language spoken document retrieval (CL-SDR).

CLEF 2004 thus marked a breaking point with respect to previous campaigns. The focus was no longer concentrated on multilingual document retrieval but was diversified to include different kinds of text retrieval across languages (exact answers in the question-answering track) and retrieval on different kinds of media (not just plain text but collections containing image and speech as well). In addition, increasing attention was given to issues that regard system usability and user satisfaction with tasks to measure the effectiveness of interactive systems or system components being included in both the QA and the ImageCLEF tracks with the collaboration of the coordinators of iCLEF (the interactive track).

In order to cover all these activities, the CLEF test collection has been expanded considerably: the main multilingual comparable corpus now contains almost 2 million news documents in ten languages. A secondary collection used to test domain-specific system performance consists of the GIRT-4 collection of English and German social science documents. ImageCLEF used two distinct collections: a collection of historic photographs provided by St Andrews University, Scotland, and a collection of medical images with French and English case notes made available by the University Hospitals, Geneva. Finally, the cross-language spoken document retrieval track (CL-SDR) used speech transcriptions in English from the TREC-8 and TREC-9 SDR tracks, supplied by the National Institute of Standards and Technology (NIST), USA.

The response from the information retrieval community was very encouraging. Participation in this year's campaign was considerably up with respect to the previous year with 55 groups submitting results for one or more of the different tracks: 36 from Europe, 13 from N.America; 4 from Asia and one mixed European/Asian group. As in previous years, participants consisted of a nice mix of new-comers and veteran groups. The success of the question answering and image retrieval tracks had a big impact on participation in CLEF 2004, not just with respect to the numbers but also regarding the skills and expertise involved. The popularity of question answering has meant that a growing number of participants have a natural language processing background while the image and spoken document retrieval tasks have brought in groups with experience in diverse areas including speech recognition, image processing and medical informatics – making CLEF an increasingly multidisciplinary forum.

The campaign culminated in the workshop held in Bath, UK, 15-17 September. In addition to presentations by participants in the campaign, talks included reports on the activities of the NTCIR evaluation initiative for Asian languages, and on industrial experience in building cross-language applications. The final session consisted of a panel in which panellists attempted to analyse the current organisation of the CLEF campaigns in depth, considering whether we are working on the right problems, choosing our investments wisely, and giving sufficient attention to the user perspective. Tracks taken into consideration for the CLEF 2005 campaign include multilingual Web retrieval and a cross-language Geographic Information Retrieval track.

The presentations given at the CLEF Workshops and detailed reports on the experiments of CLEF 2004 and previous years can be found on the CLEF website at http://www.clef-campaign.org/

CLEF workshop participants.

CLEF is an activity of the DELOS Network of Excellence for Digital Libraries.

Link:
http://www.clef-campaign.org

Please contact:
Carol Peters, ISTI-CNR, Italy
CLEF Coordinator
Tel: +39 050 3152897
E-mail: carol.peters@isti.cnr.it

CALL FOR PAPERS

CAiSE 2005 — The 17th Conference on Advanced Information Systems Engineering

Porto, Portugal, 13-17 June 2005

Since the late 1980's, the CAiSE conferences provide a forum for the presentation and exchange of research results and practical experiences within the field of Information Systems Engineering. CAiSE'05 aims at bringing together researchers, users, and practitioners in the field of information systems engineering. The conference programme will feature paper presentations, workshops, tutorials, and interactive panel sessions.

Conference Theme

The Internet has been changing the society and the economy, and the way institutions and businesses operate has evolved rapidly. All citizens are now expected to interact directly with technology based systems without direct human intermediation, using different interface appliances. Moreover, many information systems interact with each other with limited human supervision. The Internet and its infrastructure play a critical role in this evolution, and it is of the utmost importance that semantic aspects are taken into consideration on a sound basis. Improved communication and understanding between people, the final objective of advanced information systems, requires improved communication and understanding between systems and its developers. Advanced information systems engineering therefore needs to keep improving its methods and technologies in order to support the proper development of the economy and of our society.

Relevant Topics

In addition to the special theme, topics relevant for submissions to CAiSE 2005 include, but are not limited to, the following:

- Methodologies, Models, and Tools for IS Development
- Requirements Engineering for IS
- Model Driven Architectures
- Enterprise Modelling Methods and Tools
- Service Oriented and Mobile Computing
- E-government Strategies and Applications
- Knowledge, Information and Data Quality
- Web Content Management and Distribution
- Workflow Systems
- Knowledge Management
- Model and Software Reusability
- Data Warehousing & OLAP
- Metadata and Ontologies
- Support for Co-operative Work
- Innovative Database Technologies
- Distributed, Mobile, and Open Architectures
- Agent-Oriented Technologies
- Languages and Protocols for IS
- Component-Ware and IS
- IS Reengineering
- IS Usability and Interfaces to IS
- Simulation in IS Development
- Semantic Web.

Case studies and experience reports related to the above topics are also welcome

Important Dates

- Paper Submission: 30 November 2004
- Notification of Acceptance: 12 February 2005
- Final Paper Submission: 12 March 2005

The conference is organized by FEUP - Faculdade de Engenharia da Universidade do Porto (www.fe.up.pt) and DSIC-UPV - Departament de Sistemes Informàtics i Computació, Universitat Politècnica de València (www.dsic.upv.es) and sponsored by ERCIM.

More Information
http://www.fe.up.pt/caise2005/

CALL FOR PAPERS

Mirage 2005

Rocquencourt (near Versailles), France, 1-2 March 2005

MIRAGE 2005 is an international conference with focus on Computer Vision/Computer Graphics collaboration techniques involving image analysis/synthesis approaches.

In the domain of Computer Vision, this analysis by synthesis collaboration may take the form of model-based approaches for which the analysis, recognition or understanding process does not rely on the data only but is linked to the optimization of a parametric model introduced A PRIORI. In the field of Computer Graphics, this thematic refers to image-based approaches which found the modeling, animation, lighting or rendering process on the analysis of real pre-existing data in order to gain in productivity or in realism. In both cases, this extraction and synthesis collaboration can be iterative and conducted in a loop: the synthesized information laying the foundations for a better characterization that will in turn enable the synthesis step. Augmented Reality is a direct applition of this collaboration .

Authors are encouraged to submit papers on theoretical, computational, experimental or industrial aspects of model-based image analysis and image-based model synthesis.

Topics

Topics include (but are not limited to) :

- model-based imaging and analysis
- iImage-based modeling and 3D reconstruction
- data driven animation
- image and video-based lighting and rendering
- model-based vision approaches
- model-based indexing and database retrieval
- model-based object tracking in image sequences
- model-based image and shape analysis
- model-based video compression techniques.

With applications in the field of :

- human/Computer interfaces

- video-games and entertainment industry
- mMedia production from film, broadcast and games
- post-production, computer animation, virtual special effects
- realistic 3D simulation , virtual prototyping
- multimedia applications, multimedia databases classification
- virtual and augmented reality
- medical and biomedical applications.

Important Dates

- 15 November 15, 2004: Full paper submission
- 20 December 2004: Notification of acceptance
- 15 January 2005: Camera-ready papers due, early registration deadline
- 1 February 2005: Registration deadline for authors of accepted papers
- 15 February 2005: Late registration deadline
- 1-2 March 2005.

More information:
http://acivs.org/mirage2005/

CALL FOR PARTICIPATION

EGC2005 — European Grid Conference 2005

Amsterdam, The Netherlands, 14-16 February 2005

The European Grid Conference EGC2005 is the premier event on Grid Computing in Europe. It will focus on all aspects of Grid computing in Europe and a such will bring together participants from research and industry. EGC 2005 is a follow-up of the AcrossGrids Conferences held in Santiago de Compostela (2003) and in Nicosia (2004).

EGC2005 will have three main tracks:
- a scientific track;
- an industrial track;
- a special events track.

Together they aim to cover the complete field of Grid developments in Europe.

More information:
http://genias.biz/egc2005

CALL FOR PAPERS

HCI International 2005

Las Vegas, Nevada, USA, 22-27 July 2005

The 11th International Conference on Human-Computer Interaction is jointly held under one management and one registration with the Symposium on Human Interface (Japan) 2005, the 6th International Conference on Engineering Psychology & Cognitive Ergonomics, the 3rd International Conference on Universal Access in Human-Computer Interaction, the 1st International Conference on Virtual Reality, the 1st International Conference on Usability and Internationalization, the 1st International Conference on Online Communities and Social Computing, and the 1st International Conference on Augmented Cognition.

HCI International 2005 and the affiliated conferences will be held at the Caesars Palace Hotel, one of the most prestigious resorts in the world, under the auspices of 9 distinguished international boards of 202 members from 31 countries!

The conference objective is to provide an international forum for the dissemination and exchange of scientific information on theoretical, generic, and applied areas of HCI, usability, internationalization, virtual reality, universal access and cognitive ergonomics. This will be accomplished through the following six modes of communication: plenary presentation, parallel sessions, demonstration and poster sessions, tutorials, exhibitions and meetings of special interest groups. The six-day conference will start with three days of tutorials. The tutorials will begin on Friday, 22 July 2005. A total of 18 tutorials will be offered (both half-day and full-day) at introductory, intermediate, and advanced levels covering the entire spectrum of the 9 tracks and 7 conferences.

HCI International conference gathers over a thousand scientists, academics and professionals from research institutions, universities, and companies from all over the world. The HCI International 2005 exhibits will provide state-of-the-art HCI, VR, psychology and computer products and services for the users, professionals, and researchers in the field. Attendees will be able to examine state-of-the-art HCI technology and interact with manufacturing representatives, vendors, publishers, and potential employers. Exhibitors who commit to exhibit by November 20, 2004, apart from the fact that they will have a better opportunity in booth location selection and a very attractive pricing, they will be listed in the Advance Program, Final Program, and the Conference Proceedings, which will be distributed worldwide. Exhibitors who commit to exhibit by May 1, 2005 will be listed in the Final Program.

Thematic Areas

- Ergonomics and Health Aspects of Work with Computers;
  Program Chair: Pascale Carayon
- Human Interface and the Management of Information;
  Program Chair: Michael J. Smith
- Human-Computer Interaction;
  Program Chair: Julie Jacko
- Engineering Psychology and Cognitive Ergonomics;
  Program Chair: Don Harris
- Universal Access in Human-Computer Interaction;
  Program Chair: Constantine Stephanidis
- Virtual Reality;
  Program Chairs: Kay Stanney & Michael Zyda
- Usability and Internationalization;
  Program Chair: Nuray Aykin
- Online Communities and Social Computing;
  Program Chair: Jennifer Preece
- Augmented Cognition;
  Program Chair: Dylan D. Schmorrow.

Important Dates

- Paper Presentations - Deadline for Abstract Receipt: 1 October 2004
- Posters/demonstrations - Deadline for Abstract Receipt: 1 April 2005
- Special Interest Groups - Deadline for Abstract Receipt: 20 October 2004
- Tutorials - Deadline for Abstract Receipt: 20 October 2004
- Early Registration: 20 January 2005
- Mid Registration: 21 January - 20 May 2005
- Late Registration: After 20 May 2005

More information:
http://www.hci-international.org/

CALL FOR PAPERS

ECDL 2005 — European Conference on Digital Libraries

Vienna, Austria,
18-23 September 2005

ECDL 2005 is the 9th conference in the series of European Digital Library conferences. ECDL has become the major European conference on digital libraries, and associated technical, practical, and social issues, bringing together researchers, developers, content providers and users in the field. ECDL 2005 is jointly organized by the Vienna University of Technology (VUT), the Austrian National Library (ÖNB), and the Austrian Computer Society (OCG). The conference will take place in the Vienna Technical University.

Important Dates:

- workshop proposal deadline: 28 January 2005
- workshop acceptance notification: 27 February 2005
- paper/tutorial/panel submission deadline: 1 March 2005
- acceptance notifications: 15 May 2005
- final version of papers: 3June 2005.

Topics:

Topics of contributions include (but are not limited to):

- concepts of digital libraries and digital documents
- system architectures, integration and interoperability
- information organization, search and usage
- user studies and system evaluation
- digital preservation
- digital library applications.

More information:
http://www.ecdl2005.org

CALL FOR PARTICIPATION

Virtual Reality 2005

Bonn, Germany, 12-16 March 2005

IEEE VR 2005 is the premier international conference and exhibition in virtual reality, and provides a truly unique opportunity to interact with leading experts in VR and closely related fields such as augmented reality, mixed reality and 3D interaction. Test your own work and educate yourself through exposure to the research of your peers from around the world. And of course, there will be time to renew friendships, make new ones and experience the European cuisine and culture in Bonn, the former German capital.

IEEE VR is an excellent opportunity to present:

- leading edge research contributions (papers)
- innovative applications and experience reports (application sketches)
- communicate the state of the art (survey reports)
- present and discuss solutions (panels)
- introduce to your VR research (lab presentations).

The main conference starts with tutorials and workshops on March 12 and 13. They will introduce to specific topics of the field and give a compact view of ongoing research activities.

The industrial exhibition and the research lab demonstrations will provide practical insights into solutions and prototypes.

IEEE VR2005 is cohosted by:

- caesar: center for advanced European studies and research
- University of Applied Sciences Bonn-Rhein-Sieg
- Fraunhofer Institute for Applied Information Technology
- Rheinische Friedrich-Wilhelms-Universitaet Bonn.

The conference is sponsored by ERCIM

More information:
http://www.vr2005.org

CALL FOR PAPERS

INTERACT 2005 — Tenth IFIP TC13 International Conference on Human-Computer Interaction

Rome, Italy,
12-16 September 2005

The Interact '05 conference will highlight to both the academic and industrial world the importance of the Human-Computer Interaction area and its most recent breakthroughs on current applications.

Topics:

Suggested topics include but are not limited to:

- Multi-modal Interfaces
- Context-dependent Systems
- End User Development
- Intelligent Environments
- Tangible Interfaces
- Novel User Interfaces
- Usability Evaluation
- Location-aware Interaction
- Model-based Design of Interactive Systems
- Multi-user Interaction
- Safety Issues and Human error
- Cognitive Models
- Visualisation Techniques
- Web Design and Usability
- Mobile and Ubiquitous Computing
- 3D and Virtual Reality Interfaces
- Adaptive Interfaces
- Web Accessibility
- Usability & Accessibility.

Important Dates:

Full papers, tutorials, workshops, doctoral consortium:

- Submission deadline: 31 January 2005
- Notification of acceptance: 15 March 2005
- Final version due: 30 April 2005.

Short papers, special interest groups, panels, posters, demos:

- Submission deadline: 30 March 2005
- Notification of acceptance: 10 May 2005
- Final version due: 30 May 2005.

The conference is sponsored by ERCIM.

More information:
http://www.interact2005.org/

CALL FOR PAPERS

ECIR'05 — 27th European Conference on Information Retrieval

Santiago de Compostela, Spain, 21-23 March 2005

The annual European Conference on Information Retrieval Research, organized by the Information Retrieval Specialist Group of the British Computer Society (BCS-IRSG), is the main European forum for the presentation of new research results in the field of information retrieval.

The organisers encourage the submission of high-quality research papers reporting new, unpublished, and innovative research results within information retrieval. Poster submissions addressing any of the areas identified in the conference topics are also invited. Authors are encouraged to demonstrate work in progress and late-breaking research results. The conference intends to cover all aspects of accessing digital information without explicitly specified semantics. Papers whose sole or main author is an MSc, PhD or postdoctoral student are especially welcome.

Topics

Topics of interest include but are not limited to the following areas:

- IR models, techniques, and algorithms
- Users, society, and IR
- IR applications
- IR system architectures
- Content representation, and processing
- Test and evaluation methods for IR systems
- Multimedia and cross-media IR.

Important Dates

- 1 November 2004: paper submission
- 8 November 2004: poster submission
- 9 December 2004: notification for all submissions
- 22 December 2004: Final copy due.

The conference is sponsored by ERCIM.

More information:
http://www-gsi.dec.usc.es/ecir05/

SPONSORED BY ERCIM

Workshop on Construction and Analysis of Safe, Secure and Interoperable Smart Devices

Nice, France, 7-11 March 2005

The CASSIS International Workshop will be held in the week of March 7th to March 11th 2005 in Nice and will consist of keynote and invited presentations by leading academics and industrials in the field of security and analysis of smart devices.

The workshop will focus on:

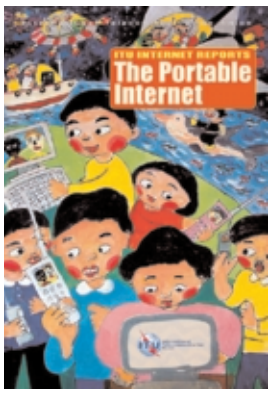
- Research trends in smart devices
- Web services
- Virtual machine technology
- Security
- Proof-Carrying Code
- Validation and formal methods
- Embedded devices.

The workshop is organized by the INRIA's Everest team.

More information:
http://www-sop.inria.fr/everest/events/cassis05/

BOOK REVIEW

The International Telecommunication Union (ITU) has just (September 2004) released the sixth book in its Internet Reports series entitled "The Portable Internet". The book examines the possible impact of high-speed wireless communication used together with portable terminals. The conclusion is that this combination could well be a "disruptive technology" with wide sweeping impact on global communication.



The book is nicely self contained with a glossary of terms included. This will be important for some readers when the myriad of available technologies is discussed in the second chapter.

The third chapter deals with market trends. Here, for example, it is interesting to see that while almost half of the world's broadband subscribers is already in the Asia-Pacific region, the potential here is enormous. Future broadband development is highly dependent on national policy and regulation, the subject of the fourth chapter. Case histories show dramatic results of deregulation. The impact that the portable Internet might have on the “digital divide” is covered in the fifth chapter. This is one of the world's largest problems but as the cost of the portable Internet decreases perhaps the enormous potential of the Internet can be realized in developing countries. Future market applications are covered in the next chapter. Existing applications in Korea let mobile telephone users pay for many items and services using their mobile telephone — examples are food and transportation. The last chapter looks at the potential impact on society and includes mention of the medical concerns that some people have for widespread wireless use.

There is an Annex at the end of the book which gives all kinds of statistical information for some 200 countries around the world. Penetration of the Internet, wide band access and mobile telephony, as well as costs, are covered. Such statistics could be rather dry but these are so interesting that it is easy to study them for quite some time. For example, the country with the highest broadband penetration is the Republic of Korea with somewhat more than 23%.

The book, over 200 pages, may be purchased via the ITU Web site: http://www.itu.int/osg/spu or http://www.itu.int/publications/bookshop or http://www.itu.int/portableinternet/

Harry Rudin,
Consultant and Swiss Local Editor for ERCIM News

ERCIM NEWS

ERCIM News is the magazine of ERCIM. Published quarterly, the newsletter reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of over 9000 copies.

ERCIM News online edition is available at http://www.ercim.org/publication/ERCIM_News/

ERCIM News is published by ERCIM EEIG, BP 93, F-06902 Sophia-Antipolis Cedex
Tel: +33 4 9238 5010, E-mail: office@ercim.org
ISSN 0926-4981

Director: Michel Cosnard

Central Editor:
Peter Kunz (peter.kunz@ercim.org)

Local Editors:

AARIT:	n.a.
CCLRC:	Martin Prime M.J.Prime@rl.ac.uk
CRCIM:	Michal Haindl haindl@utia.cas.cz
CWI:	Annette Kik Annette.Kik@cwi.nl
CNR:	Carol Peters carol.peters@isti.cnr.it
FORTH:	Eleftheria Katsouli ekat@admin.forth.gr>
Fraunhofer ICT Group:	Michael Krapp michael.krapp@scai.fraunhofer.de
FNR:	Patrik Hitzelberger hitzelbe@crpgl.lu
FWO/FNRS:	Benoit MICHEL michel@tele.ucl.ac.be
INRIA:	Bernard Hidoine bernard.hidoine@inria.fr
Irish Universities Consortium:	Ray Walshe ray@computing.dcu.ie
NTNU:	Truls Gjestland truls.gjestland@ime.ntnu.no
SARIT:	Harry Rudin hrudin@smile.ch
SICS:	Kersti Hedman kersti@sics.se
SpaRCIM:	Salvador Lucas slucas@dsic.upv.es
SRCIM:	Gabriela Andrejkova andrejk@kosice.upjs.sk
SZTAKI:	Erzsébet Csuhaj-Varjú csuhaj@sztaki.hu
VTT:	Pia-Maria Linden-Linna pia-maria.linden-linna@vtt.fi
W3C:	Marie-Claire Forgue mcf@w3.org

Free subscription

You can subscribe to ERCIM News free of charge by:

- sending e-mail to your local editor
- contacting the ERCIM office (see address above)
- filling out the form at the ERCIM website at http://www.ercim.org/

News about legal information relating to Information Technology from European directives, and pan-European legal requirements and regulations.

On the Impact of Broadband Internet

For most European internet users in 2000, broadband internet was an unknown form of access.

Fast internet did exist, but it was almost exclusively used by companies and universities. Only consumers (aka punters) with cable internet could surf the net at speeds well over 128Kbps (ISDN). However, due to the specific architecture of coax-internet, this kind of access was not common in Europe round 2000.

Some reports, mainly based on research done in the US, were not very positive about the growth perspectives for telecom operators of Broadband. There were two main reason for this:

- one renowned research company briefed to its customers : "Broadband is perceived as a luxury; until service providers make it a necessity, consumers' interest will remain limited".
- as most European punters had learned a few years before that there was something like 'free access' not many those days were willing to pay monthly fees over Euro 50.

Nevertheless during 2002/2003 all over Europe the number of broadband users surged. More Cable television cable-networks were upgraded to make internet access possible. As a result, cable providers in countries like Belgium, Austria and the Netherlands became leaders in broadband and in almost every EU-member state DSL was no longer a service offered exclusively by the incumbent telecom operator.

The number of broadband suppliers grew and all were eager to introduce new access forms to increase their market share. As a result of this in all EU member states the monthly fees for broadband access dropped.

So the number of users grew, but why? Was growth only caused by lower prices?

One thing is for sure, no access provider had unique content that that could cause a rapid growth in subscriber numbers.

Most Internet Service Providers (ISPs) now acknowledge that broadband access has rapidly become a commodity, like electricity, that enables services and content to users that are not owned in any way by their own ISPs.

Services like online banking and p2p file sharing are the main reasons why broadband access is such an success. It is somehow strange to realize, that suppliers of bank services do profit from broadband and they are even in a large degree responsible for its success, but they are not responsible for investments or maintenance of the copper infrastructure.

The way p2p filesharing lead to the success of broadband Internet is also an interesting case. Despite the fact ISPs in Europe are not banks and are not napster-like super nodes, banking and filesharing have become very important issues for ISPs in Europe.

The ever increasing demand for a safer internet and the calls from music and movie industry to do more to respect their rights, claims that are partly backed by politicians from all over Europe, causes the access providers to become more aware of the possible impact of new access technologies.

Back in 2000 some analysts already predicted that fast internet and the number of users would change over the next years. Some predicted that applications like online gaming or digital photo sharing sites could trigger rapid chances. But no one predicted that broadband use would lead to world wide discussions about DRM, eSecurity and copy-protection schemes.

by Rashid Niamat, NLIP (Dutch Internet Service Provider Association)
Editor: Heather Weaver, CCLRC, UK

ISTI-CNR — **Paolo Cignoni** is the first recipient of the Eurographics 'Young Researcher' award in recognition of his outstanding contributions to the development and use of advanced multiresolution techniques in computer graphics and scientific visualization.

The scope of Cignoni's interests ranges from scientific visualization to LOD techniques. He has published more than 60 research papers in international refereed journals and conferences, including six papers at Eurographics, and three papers at Siggraph'04. In addition, his work has found important applications in 3D scanning and cultural heritage.

Among his more recent results there is a novel approach to variable resolution rendering based on a 'GPU-friendly' data structure that is one or two orders of magnitude faster than previous solutions and has been proposed for the efficient rendering of planet-sized digital terrains and huge 3D models.

Paolo Cignoni is one of the founding members of the Visual Computing Lab at ISTI-CNR and a lecturer at the Computer Science Dept. of the University of Pisa.

INRIA — **Gilles Kahn**, Chairman of INRIA and INRIA's representative on ERCIM's board of directors was appointed member at the French National Strategic Advisory Board on Information Technologies on 27 September 2004. The Strategic Advisory Board on Information Technologies (CSTI — Conseil Stratégique des Technologies de l'Information) is chaired by the French Prime Minister and composed of leading entrepreneurs from industry and R&D, selected for their own competence. CSTI is responsible for presenting recommendations to Government related to strategic orientations in innovation and research and development, concerning information technologies.

SpaRCIM — **The First Spanish Conference on Informatics** (CEDI 2005) is to take place between 13-16 September, 2005. While national conferences are traditionally organized in several specific topics, in this case, all are to take place simultaneously and in the same location. The main goal is to join the Spanish informatics research community in order to demonstrate scientific advances, discuss specific problems and improve the visibility of this area in Spain, while emphasizing its role in the age of the information society.

CEDI will be organized in a 'federated conference' format, joining several more specific symposia in areas including Artificial Intelligence, Software Engineering and Databases, Programming Languages, Parallelism and Compute Architecture, Computer Graphics, Concurrency, Soft-Computing, Pattern Recognition, Natural Language Processing, Ubiquitous Computing and Human-Computer Interaction.

In addition to the scientific activities of each symposium, the conference will organize a number of invited talks, round tables, social events and a gala dinner, at which several awards will be announced, recognizing the efforts of some colleagues in promoting the growth of the area in Spain.

More than 1000 researchers are expected to attend the conference. The event will be funded by the Spanish Ministry of Education and Science, and is sponsored by SpaRCIM.

The conference is organized by the Department of Computer Science at the University of Granada. It will be chaired by Alberto Prieto from the University of Granada, with Juan José Moreno Navarro (UPM and SpaRCIM director) as Program Chair. Further details can be found at http://cedi2005.ugr.es/

CWI — **CWI has started a new line of research on fundamental and practice-oriented cryptology**. The institute is already known for the Number Field Sieve project for factoring large integers, which attracted world-wide attention with the factorization of the RSA-512 internet security code. The new Cryptology and Information Security group, led by prof. Ronald Cramer, includes this project. Additionally research is initiated in computational number theory with relevance for cryptology, public key cryptography, information theoretically secure cryptography, quantum cryptography, formal security analysis, and applied information security. One of the group's major focal points is secure computation. This area deals with two or more parties who wish to achieve a joint task securely even though they are mutually distrustful and wish to keep sensitive, private information secret from each other. This is sometimes called multilateral security, as opposed to unilateral security in the case of secure communications. Practical applications include profile matching, joint database comparison, electronic voting or auctions, and threshold cryptography. More information can be found on http://www.cwi.nl/pna5

FNR — **The National Research Fund of Luxembourg launched the second call of its multiannual research programme** 'Security and efficiency of new practices in e-commerce for all socio-economic actors' (SECOM) on April 1st 2004. It has been adopted by the Council of Government in the year 2000 and covers a total budget of EUR 1.2 million over the period 2005 to 2007. Five project proposals were submitted by the deadline of July 15th 2004 and are now being evaluated by independent experts. The selected projects will start early 2005.

ERCIM – The European Research Consortium for Informatics and Mathematics is an organisation dedicated to the advancement of European research and development, in information technology and applied mathematics. Its national member institutions aim to foster collaborative work within the European research community and to increase co-operation with European industry.

ERCIM is the European Host of the World Wide Web Consortium.

Austrian Association for Research in IT
c/o Österreichische Computer Gesellschaft
Wollzeile 1-3, A-1010 Wien, Austria
Tel: +43 1 512 02 35 0, Fax: +43 1 512 02 35 9
http://www.aarit.at/

Council for the Central Laboratory of the Research Councils, Rutherford Appleton Laboratory
Chilton, Didcot, Oxfordshire OX11 0QX, United Kingdom
Tel: +44 1235 82 1900, Fax: +44 1235 44 5385
http://www.cclrc.ac.uk/

Consiglio Nazionale delle Ricerche, ISTI-CNR
Area della Ricerca CNR di Pisa,
Via G. Moruzzi 1, 56124 Pisa, Italy
Tel: +39 050 315 2878, Fax: +39 050 315 2810
http://www.isti.cnr.it/

Czech Research Consortium
for Informatics and Mathematics
FI MU, Botanicka 68a, CZ-602 00 Brno, Czech Republic
Tel: +420 2 688 4669, Fax: +420 2 688 4903
http://www.utia.cas.cz/CRCIM/home.html

Centrum voor Wiskunde en Informatica
Kruislaan 413, NL-1098 SJ Amsterdam,
The Netherlands
Tel: +31 20 592 9333, Fax: +31 20 592 4199
http://www.cwi.nl/

Fonds National de la Recherche
6, rue Antoine de Saint-Exupéry, B.P. 1777
L-1017 Luxembourg-Kirchberg
Tel. +352 26 19 25-1, Fax +352 26 1925 35
http:// www.fnr.lu

FWO
Egmontstraat 5
B-1000 Brussels, Belgium
Tel: +32 2 512.9110
http://www.fwo.be/

FNRS
rue d'Egmont 5
B-1000 Brussels, Belgium
Tel: +32 2 504 92 11
http://www.fnrs.be/

Foundation for Research and Technology – Hellas
Institute of Computer Science
P.O. Box 1385, GR-71110 Heraklion, Crete, Greece
Tel: +30 2810 39 16 00, Fax: +30 2810 39 16 01
http://www.ics.forth.gr/

Fraunhofer ICT Group
Friedrichstr. 60
10117 Berlin, Germany
Tel: +49 30 726 15 66 0, Fax: ++49 30 726 15 66 19
http://www.iuk.fraunhofer.de

Institut National de Recherche en Informatique
et en Automatique
B.P. 105, F-78153 Le Chesnay, France
Tel: +33 1 3963 5511, Fax: +33 1 3963 5330
http://www.inria.fr/

Norwegian University of Science and Technology
Faculty of Information Technology, Mathematics and
Electrical Engineering, N 7491 Trondheim, Norway
Tel: +47 73 59 80 35, Fax: +47 73 59 36 28
http://www.ntnu.no/

Spanish Research Consortium for Informatics
and Mathematics
Esperanza Marcos, Rey Juan Carlos University,
C/ Tulipan s/n, 28933-Móstoles, Madrid, Spain,
Tel: +34 91 664 74 91, Fax: 34 91 664 74 90

Swedish Institute of Computer Science
Box 1263,
SE-164 29 Kista, Sweden
Tel: +46 8 633 1500, Fax: +46 8 751 72 30
http://www.sics.se/

Swiss Association for Research in Information Technology
c/o Prof. Dr Alfred Strohmeier, EPFL-IC-LGL,
CH-1015 Lausanne, Switzerland
Tel +41 21 693 4231, Fax +41 21 693 5079
http://www.sarit.ch/

Slovak Research Consortium for Informatics and
Mathematics, Comenius University, Dept.of Computer
Science, Mlynska Dolina M, SK-84248 Bratislava, Slovakia
Tel: +421 2 654 266 35, Fax: 421 2 654 270 41
http://www.srcim.sk

Magyar Tudományos Akadémia
Számítástechnikai és Automatizálási Kutató Intézet
P.O. Box 63, H-1518 Budapest, Hungary
Tel: +36 1 279 6000, Fax: + 36 1 466 7503
http://www.sztaki.hu/

Trinity College
Department of Computer Science,
Dublin 2, Ireland
Tel: +353 1 608 1765, Fax: 353 1 677 2204
http://www.cs.tcd.ie/ERCIM

Technical Research Centre of Finland
P.O. Box 1200
FIN-02044 VTT, Finland
Tel:+358 9 456 6041, Fax :+358 9 456 6027
http://www.vtt.fi/tte

Order Form

If you wish to subscribe to ERCIM News ***free of charge*** *or if you know of a colleague who would like to receive regular copies of ERCIM News, please fill in this form and we will add you/them to the mailing list.*

Send, fax or email this form to:

ERCIM NEWS
2004 route des Lucioles
BP 93
F-06902 Sophia Antipolis Cedex
Fax: +33 4 9238 5011
E-mail: office@ercim.org

Data from this form will be held on a computer database. By giving your email address, you allow ERCIM to send you email

I wish to subscribe to the

❑ ***printed edtion*** ❑ ***online edition (email required)***

Name:

Organisation/Company:

Address:

Post Code:

City:

Country

E-mail:

You can also subscribe to ERCIM News and order back copies by filling out the form at the ERCIM website at http://www.ercim.org/publication/Ercim_News/